Economics of wildfire suppression:
Estimation of drivers of suppression expenditure
and Risk preference experiments with wildfire management

by

Michael Shucheng Huang

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Agricultural and Resource Economics

Department of Resource Economics and Environmental Sociology

University of Alberta

Abstract

Public wildfire management agencies are presented with a momentous responsibility: to protect life, property and infrastructure from the devastation of wildland fire, while operating at a level of expenditure justifiable to taxpayers. At a time when climate change drives more extreme fire behaviour, agencies such as Alberta Wildfire must also be prepared to respond despite uncertain budgets. This thesis contributes by offering new empirical insights on wildfire suppression through two directions. Firstly, Chapter 2 focuses on the drivers of wildfire suppression costs. An empirical model seeks to explain how costs are affected by a series of environmental factors, such as time-variant weather variables and time-invariant landscape characteristics, as well as by operational policy variables based on Alberta Wildfire's organizational capacity and priorities. Results from regression analysis and machine learning show that while policy decisions have measurable impacts on abating costs, the bulk of expenditures is driven by environmental factors. Chapter 3 studies risk aversion of Alberta Wildfire Incident Commanders (ICs). Through laboratory economic experiments, I seek to determine whether ICs, who are in a risky profession, exhibit risk preferences in laboratory experiments that are different from a control group. Results show that ICs' experiment choices are not significantly different from those of typical experimental subjects across all risk elicitation tasks. However, among their colleagues, ICs with additional operational deployment experience tend to exhibit significantly lower levels of risk aversion. Findings from this study motivate further research into wildland firefighters' risk preferences that will help decisionmakers better understand how individual risk perceptions impact resource allocation, and by extension, costs. Taken together, the novel insights generated from this thesis contribute to the multi-disciplinary field of wildfire suppression research.

Preface

Experimental research reported in this thesis was approved by the University of Alberta Research Ethics Board:

- Chapter 3: “Risk Elicitation Economics Experiments”, Pro00106176, Feb. 01, 2021

Financial support for this thesis was provided by Alberta Agriculture and Forestry, Wildfire Management Branch (Alberta Wildfire), through the Canadian Partnership for Wildland Fire Science (Canada Wildfire).

Acknowledgements

I would like to extend my gratitude to Alberta Wildfire for the funding and expertise contributed to this project, and to the Incident Commanders who took the time to complete the risk elicitation experiment. A special thanks to Cordy Tymstra, Dave Schroeder and Lynn Ducharme, who took the time to explain to us the mechanics and data of wildfire suppression. All references to “us” and “we” throughout this thesis include my supervisor, Bruno Wichmann, who has guided me through the research journey. From experiments to machine learning, Bruno has continually encouraged me to perform my absolute best. Muito obrigado. A big thanks to Vic Adamowicz and his research group, whose advice and feedback helped motivate me through the most challenging parts of this project. My friends and colleagues in REES, thank you for your comradery through coursework, socially distanced walks and ice creams breaks, and for the second set of eyes on coding conundrums that inevitably arose from time to time. Finally, I am particularly thankful to my family for their continued support and patience throughout the course of this degree.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

In 2019, Alberta Wildfire faced a particularly severe wildfire season in which nearly 1,000 wildfires burned over a total of 880 thousand hectares. Expenditures incurred in fighting the flames totalled $570 million (MNP LLP, 2020). As climate change contributes to new fire weather extremes, wildfire management agencies like Alberta Wildfire are also embattled with challenges to provide suppression coverage while dealing with increased public scrutiny (Canton-Thompson et al., 2008) and uncertain budgets (Tymstra et al., 2020). In order to understand the drivers of wildfire suppression expenditures, I perform empirical analyses to discover how policy and environmental factors influence expenditures. In addition, I undertake an experiment that seeks to uncover the risk preferences of wildfire managers, who play a key role in providing suppression services, and incurring the expenses resulting from suppression.

Chapter 2 focuses on “Wildfire suppression expenditures”, in which regression analysis and machine learning tools are applied to an expenditure model. While regression analysis uncovers the linear relationships between costs and various environmental and policy factors, ML allows us to estimate flexible functions that map environmental factors to key discretionary policy variables and expenditure. In Chapter 3, “Risk aversion in wildfire management” is explored through a series of risk elicitation methods common in the economic literature (Multiple Price List, Single Choice List, Certainty Equivalent Method and Investment Game), plus a self-evaluated risk survey. Through these revealed and stated risk preference tasks, I seek to discover how risk preferences differ between Alberta Wildfire Incident Commanders and control participants, as well as how risk preferences differ among Incident Commanders due to operational experience.

Chapter 2. Wildfire suppression expenditure

Understanding the impacts of environmental and policy variables on wildfire expenditures is critical for wildfire management agencies, who must adapt to a challenging ecological future, often equipped with similarly challenging budgets. In this chapter, I apply regression analysis and machine learning methods to a dataset of over 5,000 wildfires suppressed by Alberta's Wildfire Management Branch between 2015 and 2020. Results show that, while certain operational changes have meaningful (albeit modest) impacts on suppression costs, the bulk of wildfire suppression expenditures is driven by environmental factors such as weather and landscape.

2.1. Introduction

Research in wildland fire science in Canada has flourished over the past 50 years, focusing on fire behaviour, risk analysis, forest management, and more recently, climate change impacts (for a review, see Coogan et al., 2021). In the recent 20 years, as suppression costs rise with increasingly severe fire seasons in Canada, there has been growing awareness among the public on the impacts and costs of wildfires (Owen, 2021; Popyk, 2021), and greater motivation amongst researchers to account and forecast wildfire suppression expenditures (Hope et al., 2016; Tymstra et al., 2020).

In Alberta, the devastating 2016 Fort McMurray fire serves as a reminder of wildfire's disruptive nature, which is captured in part by a large economic toll. Estimates of total damage range from the $3.7 billion as paid out by insurance (KPMG, 2017), to over $10 billion when accounting for costs incurred by the private and public sector, plus values of ecosystem loss (Alam et al., 2017). The Province of Alberta footed $400 million in responding to this wildfire through suppression efforts by the Ministry of Agriculture and Forestry's Wildfire Management Branch (WMB or "Alberta Wildfire"), with additional funds for community recovery (KPMG, 2017; Lamoureux and Bellefontaine, 2016).

Prior to the fire events of 2016, Alberta's yearly wildfire expenditures were already experiencing a steady increasing trend (MNP LLP, 2016; Stocks and Martell, 2016). While Alberta Wildfire's total yearly expenditures rarely exceeded $100 million prior to 2000, from 2000 to 2016, annual costs regularly exceeded $200 million, and in 2015, exceeded $400 million[1] (MNP LLP, 2016). Faced with the challenges presented by rising expenditures and uncertain budgets (Tymstra et al., 2020), as well as an increasingly severe and unpredictable wildfire future due to

[1] All costs mentioned in this paragraph are in 2016 dollars.

climate change impacts (Flannigan et al., 2000; Robinne et al., 2016; Tymstra et al., 2021; Wotton et al., 2017), Alberta Wildfire is supporting research studies that address a pressing policy question: "How can wildland fire operations be made more efficient and effective?"

In this chapter of the thesis, I discuss the effects of environmental and policy variables on fire-level suppression costs. Regression modelling uncovers the extent to which suppression costs are impacted by environmental variables such as weather, geography and location, and also by discretionary decisions taken under the guidance of Alberta Wildfire policy, including strategic delay, resource allocation under fire competition, and the prioritization of "values-at-risk"[2]. The next sections include a brief overview of related literature on wildfire suppression expenditure analysis, data and analytical applications, the empirical expenditure model, and results, followed by a discussion on the policy implications of these findings. However, firstly I provide some background information on wildfire management in Alberta and Canada, with additional focus on Alberta Wildfire's operations and objectives.

Background on wildfire management in Alberta and Canada

Across present-day Alberta and throughout forested areas of North America, wildland fires, or wildfires, have historically been an integral part of ecosystems by reducing fuel buildup and promoting vegetative diversity (Groesch et al., 1992; Pausas and Keeley, 2019). Prior to European settlement, indigenous peoples played an active role within such ecosystems through their regimented use of prescribed fire to maintain an environment adaptable for hunting and other land-based activities (Kimmerer and Lake, 2001; Lake and Christianson, 2019).

[2] Values-at-risk are protected by Alberta Wildfire in descending order of priority: Human Life, Communities, Watershed and Sensitive Soils, Natural Resources, Infrastructure (Alberta Wildfire, 2017).

In the contemporary era, public authorities responsible for wildfire management have largely favoured full suppression of wildfires in all circumstances. In Canada, full suppression is prevalent in most jurisdictions, especially where human communities and infrastructure border fire-prone forests, in what is termed the “wildland-urban interface”, or WUI (Pyne, 2008). However, there remains a lack of consensus among wildfire researchers on whether full suppression serves as the best method to protect communities (Houtman et al., 2013; Johnson and Miyanishi, 2001; McFarlane et al., 2011; Riley et al., 2018). One concern in particular is fuel buildup: when a full suppression program deprives wildfire of its role in removing low-lying vegetation, future fires will be able to catch on to these sources of fuel, and are more likely to grow into larger, more devastating wildfires (Brown, 1983; Johnson et al., 2001).

In Alberta, the provincial government, through Alberta Agriculture and Forestry’s Wildfire Management Branch, is responsible for suppressing all wildfires in the designated Forest Protection Area (FPA) that extends from the eastern slopes of the Rocky Mountains to the boreal forest of northern Alberta (see Figure A.1). The FPA is divided into 10 Forest Areas (FAs) with an Area Fire Centre in each FA. Wildfire detection and large-scale resource deployment is centrally coordinated from the Alberta Wildfire Coordination Centre (AWCC) in Edmonton. Municipal governments and federal governments are responsible for fire management within their own jurisdictions, however, there is mechanism for inter-agency collaboration for fires that cross boundaries: the Canadian Interagency Forest Fire Centre (CIFFC), jointly owned by federal, provincial and territorial agencies, coordinates inter-provincial resource deployments and information sharing (CIFFC, 2021).

Within Alberta’s FPA are many residential and industrial infrastructures bordering the WUI. 370 thousand people are year-long residents in the FPA, and populations in these regions are forecasted to increase substantially into the future (Government of Alberta, 2021). Even more people work and recreate within Alberta Wildfire's jurisdiction, including up to 30 thousand rotational upstream energy workers (PetroLMI and ENFORM, 2015) in the boreal area, as well as recreationalists in the eastern slopes who make yearly visits in the millions (Colgan, 2021).

Wildfire preparedness and response in Alberta

While full suppression response is the most prominent and well-known approach of wildfire management, it is crucial to recognize it is only one of four pillars: *mitigation, preparedness, response* and *recovery* (Canadian Wildland Fire Strategy Assistant Deputy Ministers Task Group, 2005). A hallmark mitigation/preparedness strategy is CIFFC's FireSmart program which encourages WUI communities to build their resilience to wildfires through landscaping changes that mitigate the likelihood of flames reaching homes (FireSmart Canada, 2021). The adoption of FireSmart practices has proven effective in reducing wildfire damages both in simulated environments and in WUI communities across Canada (Ergibi and Hesseln, 2020; Labossière and McGee, 2017; Schroeder, 2010). Nevertheless, precautionary measures are not always sufficient by themselves in preventing the risks posed by forest fires, and response through suppression remains most commonly used and most funded approach by Canadian wildfire agencies (McGee et al., 2015).

Wildfire suppression operations in Alberta

Wildfires in the Forest Protection Area (FPA) are detected by Alberta Wildfire through air, ground, and water patrols and a network of manned lookout towers. Members of the public area also encouraged to report wildfires to the agency through a hotline.

While the yearly budget for pre-suppression preparedness hovers from $100 million to $200 million (MNP LLP, 2016), Alberta Wildfire is authorized to access a contingency fund for emergency response, incurring costs as high as $750 million during a particularly intensive fire season, such as in 2019 (MacVicar, 2019). With the financial backing of the provincial government, Alberta Wildfire aims to meet two operational performance objectives: 1) to initiate suppression action before the wildfire reaches 2 hectares, and 2) to contain wildfire spread by 10:00 hours the following day. In order to meet these objectives, Alberta Wildfire engages nearly 660 permanent and seasonal wildland firefighters during the annual fire season, March 1 to October 31. Frontline firefighters across the FPA are supported by logistic coordinators of the Alberta Wildfire Coordination Centre (AWCC) in Edmonton. The response protocol is summarized in Figure 2.1.

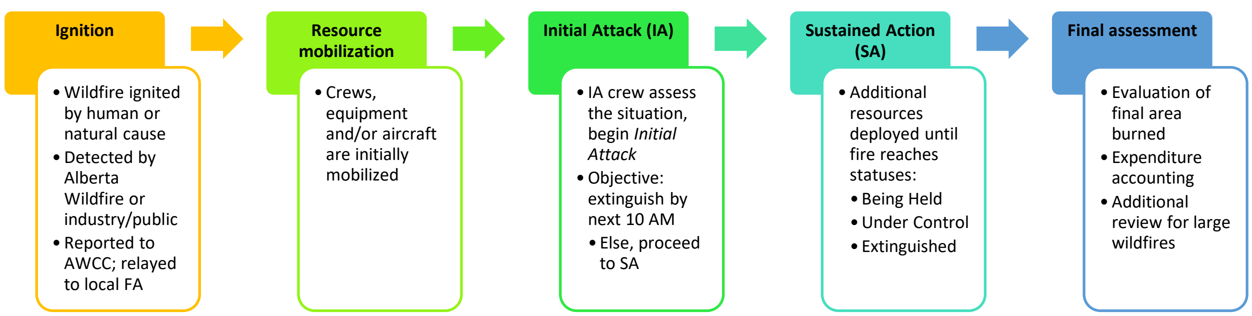

Figure 2.1 Flowchart of Alberta Wildfire's response protocol.

A detected wildfire is first reported to AWCC, which relays notification to the local Forest Area (FA) in which the fire is detected. Typically, four-person *Helitack* crews are the first to respond to a wildfire, arriving by truck, fire engine or helicopter. *Initial Attack* will generally provide the fastest and most cost-effective form of wildfire suppression, due to the small size and intensity of a nascent fire. However, if the fire has developed to an extent that it cannot be contained by *Initial Attack*, additional crews of 8 to 20 members are deployed for *Sustained Action*. (See Figure 2.1 for a chronological flowchart of suppression phases.) In addition to using manual tools, personnel on the ground also employ heavy equipment such as dozers, excavators and water trucks. Airtankers and other aircraft are the most prominent tools in the wildfire management arsenal. While air deployments incur substantial costs, they are crucial for suppressing fires in locations deemed too remote, or with conditions too dangerous, for firefighters' direct engagement. Nevertheless, crew members form the backbone of Alberta Wildfire's wildfire response; every fire is attended by personnel on the ground to ensure full extinguishment.

Wildfires are classified by size class based on their burn area: A (≤ 0.1 ha); B (> 0.1 ha to 4 ha); C (> 4 ha to 40 ha); D (> 40 ha to 200 ha); E (>400 ha). As a fire develops in size, Alberta Wildfire protocols dictate that certain minimum resource levels must be deployed at each size class, per resource availability (Alberta Wildfire internal documents; personal communication). During a busy fire season, Alberta Wildfire may recruit emergency contractors on an as-needed basis, as well as request importation of manpower, equipment, and/or aircraft from other Canadian and international jurisdictions.

Within our dataset of wildfires in the Alberta FPA from 2015 to 2020, we observe that yearly expenditures on wildfire suppression fluctuate between a low of $6.5 million to over $280 million, while the yearly number of wildfires has gradually declined (see Figure 2.2 Yearly counts and yearly expenditures by size class in Section 2.3, Data, below). In this project, I discover how expenditures incurred during a wildfire suppression mission are influenced by environmental factors like environmental variables, as well as highlighting how decisions made by Alberta Wildfire managers are associated with expenditures.

2.2. Related literature

2.2.1. Suppression expenditure regression analysis

The wildfire issue in Canada has been studied rigorously by forest science researchers (Beverly et al., 2011; Flannigan et al., 2005, 2000; Hanes et al., 2019; Stocks et al., 1998), including work on capturing *ex-post* wildfire expenditures as well as forecasting future socio-economic costs (Hope et al., 2016; Stocks and Martell, 2016; Tymstra et al., 2020). Additionally, sociologists have studied the social effects of particularly devastating wildfires notably the 2016 Fort McMurray wildfire (Boulianne et al., 2018; Drolet et al., 2021; Verstraeten et al., 2021). To the best of our knowledge, empirical economic research on wildfire suppression expenditures has been missing in the Canadian literature.

We look abroad for existing literature on expenditure analysis, and find in Hand et al. (2014) a thorough review of suppression expenditure modelling in the United States. Empirical research into wildfire expenditures was first motivated by a 2005 federal directive to establish performance measures for the United States Department of Agriculture's Forest Service (USFS) and other federal agencies' wildland fire suppression programs. This directive included a mandate on reporting the percentage of fires that exceeded a "Stratified Cost Index" (SCI), originally specified as expenditures per acre over energy release component, and later capturing fire intensity and size, regional variability, proximity to communities and other factors influencing suppression expenditures. Since 2006, SCI has become a catch-all term for all expenditure regression models within USFS.

SCI has been incorporated in interagency federal wildfire framework as an indicator of cost effectiveness, although not necessarily one of performance. Instead, within the broader scope of the mission to reduce the likelihood and occurrence of large, costly fires, expenditure models "highlight the budgetary consequences of reducing the incident of large fires" (Hand et al., 2014, p. 11).

Previous to the 2005 directive, regression analysis had already been applied to expenditure modelling on other occasions, albeit with much fewer independent variables and smaller sample sizes (Donovan et al., 2004; Gonzalez-Caban, 1984). In response to the 2005 directive, Gebert et al. (2007) develop a regression model that explores how suppression expenditure per area burned is influenced by factors including: area burned, environment, values-at-risk, resource availability, initial suppression strategy and delay. Gebert et al. characterize fire environment by the aspect, slope and fuel type (grass, brush and timber) at the ignition point, a fire weather index from the nearest weather station to ignition, and a measure of fire intensity at the start of the burning period. Based on previous research into role of population in forested areas in influencing expenditures (Snyder, 1999), Gebert et al. define values-at-risk as the distance between ignition and nearest population centre as well as total housing value within certain radii from ignition. In addition, they control for designated wilderness lands, which are areas where USFS have different suppression protocols.

In order to include the effect of operational decisions, Gebert et al. also add the delay time between ignition and discovery, as well as dummy variables that capture suppression strategy (*confine* and *contain*, referenced against *control*). To control for availability of suppression resources, Gebert et al. calculate the difference between the count of fires burning during the same time as fire i and the average count of regional fires in the same period during previous years.

This model is applied to a sample of 1,550 large fires[3] suppressed by the USFS from 1995-2004, for which expenditures were in constant 2004 dollars. The dependent variable is expenditure over area, as Gebert et al. assert that "fire managers are accustomed to thinking in terms of cost per acre" (p. 189); cost per area is log form, to reduce heteroskedasticity among residuals due to the variation in wildfire size and costs. To account for geographic differences, western and eastern United States regions are analyzed separately. Endogeneity is an issue of concern in this model due to the possible two-way causality between cost/area and burned area. As suppression costs increase with area burned, theoretically, burned area has a negative relationship with suppression costs. Gebert et al. acknowledge this concern, but they argue endogeneity is mitigated by the heterogeneity of selected large fire events, which by definition are so large that they resist initial suppression efforts, the variable burned area is more likely to vary as a function of fire complexity rather than suppression effort.

For both western and eastern models, area burned has a negative effect on cost/acre, while other environmental variables like slope and timber fuel have positive effects. As expected, both fire weather severity and initial fire intensity correlate positively and significantly with suppression expenditure. Gebert et al. express surprise in discovering, in the eastern model, that costs increase with distance to communities, because they expected fewer values-at-risk in remote areas; they consider it is possible that the further a fire is from town, the costlier it is to mobilize resources. Another surprise in the eastern model is the discovery that the initial strategy of confinement or containment over control cause costs per acre to double; Gebert et al. had expected the most aggressive strategy of *control* would cost more. In the western model, the effect of resource

[3] Gebert et al. (2007) define large fires as those that exceeded the USFS "escaped" fire limit, being 100 ac (40 ha) prior to 2003 and 300 ac (121 ha) since 2003. They also exclude fires missing identifiable expenditure information (e.g. in some large fire complexes) or other explanatory variables.

availability is marginally significant in its modest negative impact on cost per acre, a finding that is consistent with the authors' hypothesis that fewer resources (and thus lower expenditures) are available for each individual fire during a busy fire season.

Using these estimators, Gebert et al. predict costs of a selection of out-of-sample fires that took place in the last year of the study period, 2004. Comparing these estimates to real expenditure values, they note the R-squared values for out-of-sample predictions (western model: 0.33; eastern model: 0.18) were significantly lower than those in the original model (0.45; 0.46). Gebert et al. acknowledge a key limitation in model is that it does not capture political and jurisdiction influences on suppression strategies, such as the pressure to aggressively fight fires beyond operational standards. Further, noting the large confidence intervals of the predicted 2004 out-of-sample fire costs, they caution against implementing reforms based on the interpretation of results of expenditure regression models. Instead, Gebert et al. believe that regression analysis on expenditures is most beneficial in its ability to identify outliers, leading to further in-depth reviews on the fires that are significantly more, or less, expensive than expected.

Following this article, additional research introduces variations of the suppression expenditure regression model by focusing on specific sets of policy and environmental variables that explain variations in expenditure. The following is a review of the suppression expenditure literature that follow Gebert et al. (2007).

Effects of environmental variables

Liang et al. (2008) investigate the impacts of solely environmental factors on USFS expenditures over a similar time period, although within a smaller sample set of 100 large wildfires (over 121 ha) in the Rocky Mountains region of Idaho and Montana. The dependent variable is suppression expenditure in log costs. Among the environmental factors, which also include infrastructure value, distance to WUI zones, fuel conditions, topography and location, Liang et al. discover that only fire size and proportion of private land have strong positive effects, accounting for more than half of the variation in expenditures.

Effects of values-at-risk

Throughout the wildfire expenditure literature, researchers generally only account for residential properties as values-at-risk (Bayham and Yoder, 2020; Clark et al., 2016; Gebert et al., 2007; Liang et al., 2008; Yoder and Gebert, 2012). This valuation is indicative of wildfire management agencies' prioritization of human life and property, but it is also due to the difficulty of quantifying nonmarket resource values (Calkin et al., 2005; Venn and Calkin, 2011).

While Gebert et al. (2007) and Liang et al. (2008) account for property values in their expenditure analysis, Clark et al. (2016) argue this variable is unnecessary because wildfire managers are more concerned about protecting a structure for its purpose as a primary residence over its assessed value. In their model specification, Clark et al. find that the property value effect is insignificant; instead, it is the spatial pattern of residential development in the affected fire area that explains expenditure variation. While an additional house within the fire area increases suppression expenditure by $100, the marginal cost suppression increases further with distance from the house to structures; as such, the marginal suppression cost for a very remote, rural residence reaches $225 thousand. Considering these findings, Clark et al. encourage Wildland-Urbane Interface planners to disincentivize remote residential development, towards reducing public expenditure on wildfire suppression.

The significance of the property value effect is inconsistent across the expenditure literature. Donovan et al. (2011) find that media coverage and political pressure are both significantly positive in their effects on suppression costs, while the value of protected structures are insignificant. However, outside of the expenditure framework, Bayham and Yoder (2020) discover a significant relationship between aircraft deployment, an expensive resource, and the property value of threatened homes.

Effects of suppression strategy decision

Gebert and Black (2012) analyze 1,330 fires in the United States from 2006-2008, characterizing suppression strategies, in order of decreasing intensity, from *direct* to *modified* to *limited*. On average, direct suppression does result in lower total expenditure, because fires are kept small and last shorter. Yet when variables like area burned are held constant in a generalized linear mixed model, regression analysis demonstrates that *limited* suppression leads to lower costs in the form of expenditure per acre. Gebert and Black acknowledge that analysis trough the expenditure framework is limited in its scope, as it disregards other performance measures of wildfire suppression, particularly in a time when the wildfire research community is becoming more concerned about long-term implications of full suppression.

Incorporating spatially and temporally descriptive data

While most of the existing expenditure analysis literature focus on the effect of static fire conditions on cost per area of a fire (Donovan et al., 2011; Gebert et al., 2007; Liang et al., 2008), Hand et al. (2016) investigate how spatially and temporally descriptive environmental variables affect total expenditures. The authors argue that models that use ignition-point conditions (time and place of ignition) do not capture the management decisions and inherent expenditures that are made in response to a fire with space/time-variable conditions. As well, Hand et al. contend that total cost per fire is a superior variable over cost per area because the value of a fire's total cost is more pertinent to wildfire managers.

Towards building their model, Hand et al. create spatially and/or temporally descriptive iterations of environmental variables. For instance, while predominant fuel type is captured in the ignition-point model as a dummy variable, Hand et al. create spatially descriptive variables which are the proportions of burned area made up of brush/timber/slash fuel types; while ignition-point models use a single weather index at the start of the fire, Hand et al. calculate the standard deviation of this index throughout the duration of the fire, in order to capture spatial/temporal variation in weather. In addition, Hand et al. address the possibility of endogeneity of fire size (due to reverse causality of expenditure values) by instrumenting fire size on calendar year, seasonality, and dummy for an outlier region.

When comparing the performance of their spatial-temporal heterogenous model to that of the ignition-point model on large USFS wildfires, Hand et al. find that the former captures variation modestly better[4], while accounting for fire size endogeneity in model specification does not improve fit. Hand et al. promote the spatially descriptive model for its unique insights into the degree to which environmental variables impact expenditures, however, they acknowledge that general interpretations of these effects are consistent between spatially descriptive and ignition-point models.

2.2.2. Heterogeneity in resource allocation between firefighters

To investigate the influence of heterogeneity among firefighters in wildfire suppression efficiency, Hand et al. (2017) study a set of 89 incident management teams (IMT) on large wildfire incidents in the US from 2007 to 2011. They estimate a linear model in which the dependent variable is the fire-line production capacity, a measure of suppression effort, of (I_{ijt}) of IMT i for fire j on day

[4] R-squared for the models in Hand et al. (2016): Spatially descriptive model: 0.669 (fire size exogenous); 0.648 (fire size endogenous); Ignition-point model: 0.636 (fire size exogenous); 0.634 (fire size endogenous)

t. Having controlled for geography, seasonality, fire conditions and jurisdiction, Hand et al. find that 14% of variation in resource use is accounted by differences between IMTs. In all, 17 (14) of the 89 IMTs are significantly more (less) productive than the median level *($p < 0.05$)*. While this study offers preliminary insights into the IMT influence on heterogeneity in resource allocation, Hand et al. acknowledge it is limited in its ability to account for differences in management objectives, nor does it account for the efficacy of resource deployment.

2.2.3. Suppression strategy choice experiments

Beyond analysis of real-life wildfires, economists also use experiments in order to better understand the impact of risk behaviour on suppression strategy selection. Wibbenmeyer et al. (2013) theorize that an expected economic loss minimizer, the behaviour purportedly characterized by USFS fire management policy, makes rational decisions after having considered the conditions of the fire and the likelihood of its endangerment to values-at-risk. The authors develop a choice experiment for wildfire managers in which respondents are presented with three hypothetical wildfire scenarios, and tasked to select a preferred strategy for each. Each scenario includes unique characteristics: fire conditions, as probability of a certain area being burned, and values-at-risk, as the relative location of homes and watersheds to a potentially burned area. For each scenario, the choice set includes alternative strategies with varying levels of resource value-at-risk protection, resource allocation and personnel exposure, probability of success, and program cost. Respondents select both their personal choice, as well as the choice which they believe their agency would choose.

Applying the Kahneman and Tversky (1979) prospect theory model, Wibbenmeyer et al. discover that their wildfire manager respondents exhibit non-expected utility in risk behaviour. When homes are at risk, respondents are more sensitive to the differences across choice sets in the probability of suppression success rather than that of burn probability. To Wibbenmeyer et al., this behaviour is indicative that wildfire managers, who feel like they have no control over fire scenarios, wish to be able to control the probability of success through their choice.

Using the same dataset, Wibbenmeyer and Calkin collaborate again in Calkin et al. (2013), in which they focus on the differences between respondents' preferred choice and the one they expect the agency would choose. Controlling for the effects of fire condition, values-at-risk, and the probability of suppression success, a wildfire manager is less likely to personally prefer a strategy when its cost increases. However, choices by respondents indicate that expenditure values are not significantly impactful to decisions taken by their upper management. These results lead Calkin et al. to discuss how management agencies are often pressured to prioritize social-political considerations over economic ones. This is particularly pertinent if a fire were to break through containment, as a wildfire management agency that has already incurred great expenses through a disproportionately high level of resource allocation is in a better position to contend that they are not liable for damages.

Wibbenmeyer and Hand collaborate once again in Hand et al. (2015). In this paper, the researchers elicit risk behaviour from wildfire managers using a multi-attribute lottery experiment in the context of hypothetical wildfire scenarios. The experiment, inspired by the Holt and Laury (2002) method[5], tasks participants with choosing between "safe" and "risky" wildfire suppression strategies. Each strategy also includes two possible outcomes: a "good" outcome has low levels of

---

[5] The Holt and Laury (2002) method is a lottery choice game with a set of 10 paired lotteries. On each line, two lotteries are represented as *Option A* and *Option B*. *Option A* is a "safe" lottery in which the difference between its high/low payoffs is smaller than that of the "risky" *Option B*. Down the list, the high/low payoff values of *Option A* and *Option B* remain constant, while probabilities change such that an expected utility maximizer is induced to switch from *Option A* to *Option B*. Switching further down the line is indicative of higher risk aversion.

This method is applied in Chapter 3. Risk aversion in wildfire management, and an example of this method can be found in *Task "Orange" Decision*, Appendix I: Instructions for the Risk Elicitation Economics Experiment.

firefighter fatality, property damage, and suppression costs, and a "bad" outcome with high levels. "Safe" strategies have moderate differences between good/bad attribute values, while "risky" strategies have large differences. As in Holt and Laury (2002), probabilities associated with good/bad outcomes change throughout the decision list.

Experiment choices reveal that that wildfire managers are overwhelmingly risk averse in their suppression strategy choice, opting for a safe strategy even when the expected losses in the risky strategy are lower. In addition, through applying Prospect Theory, Hand et al. discover that participants tend to overweight low probabilities of outcomes and underweight high probabilities. Hand et al. conclude that wildfire managers are biased towards risk aversion in suppression management, a result that supports previous theoretical (Maguire and Albright, 2005) and empirical (Wilson et al., 2011) work.

2.2.4. Qualitative research on organizational drivers of expenditures

The reviewed literature, above, all speculate on the socio-political motivations for expenditure variation in wildfire management scenarios (real and hypothetical). However, not all factors influencing decision-making and suppression costs can be captured by quantitative data. Unobserved human factors, from the individual to the organizational level, can create a wildfire management environment in which large fire suppression expenditures seem to be an inevitable and insurmountable problem.

To shed light on the role of individual/organizational effects on wildfire expenditures, Canton-Thompson et al. (2008) carry out in-depth interviews with 48 United States Forestry Service (USFS) firefighters across all levels of seniority. Senior firefighters lament that individualist attitudes among younger colleagues have led to less cohesion in wildfire management, and more wasted time during decision-making. Further, as a changing organizational culture leads agency units to become silos, USFS is neglecting to properly fund pre-suppression programs, including the recruitment and training of professional firefighters. Consequently, supplementary costs are incurred at the end of fire season as suppression expenditures due to the hiring of contractor firefighters.

Changes in society-at-large also contribute to increased wildfire expenditures. Interviewees point to increasing population in the wildland-urban interface (WUI) as a pressure for intensive wildfire suppression; additionally, new residents in the WUI lack appreciation for the complexity of wildfires. Consequently, such residents often exert political pressure to obtain intensive fire suppression, even in situations where high resource allocation is excessively costly and unnecessary. Yet at the end of such a fire season, USFS faces the opposite form of political pressure: being criticized for the additional suppression expenditures incurred in meeting politically-driven suppression strategies.

2.3. Data

To determine the impact of policy and environmental variables on wildfire expenditure, we primarily use Alberta Wildfire's internal datasets on operations and expenditure from calendar years 2015 to 2020[6], made available to us for the purpose of this project. We combine fire-level

[6] Alberta Wildfire organizes operational and expenditure data by "fire year". We received data on wildfires from January 2015 – December 2020. From January 2015 – December 2018, Alberta Wildfire defined the fire year as April 1 to March 31. Starting in January 2019, the fire year follows the calendar year (January 1 to December 31). As 88% of wildfires in our dataset take place April – August, we believe that it is appropriate to reference wildfires by calendar year, and to incorporate calendar year instead of fire year in analysis.

observations of operation and expenditure data with geospatial data from public databases (Altalis, Government of Alberta GENESIS, Statistics Canada) to create a high-dimensional final dataset.

Within Alberta Wildfire, fires are classified by their burn area, from size class A (<0.1 ha) to E (>400 ha); at each size class, operational protocols dictate that certain minimum resource levels must be met, per resource availability (Alberta Wildfire internal documents; personal communication). A cursory look at our Alberta Wildfire operational dataset, Table 2.1 , reveals that 65% of wildfires in Alberta's Forest Protection Area (FPA) during the study period were suppressed before exceeding size class A, and nearly 90% before surpassing size class B.

Table 2.1 Count of wildfires responded by size class, calendar years 2015 to 2020

Calendar Year	Size class					Total
	A	B	C	D	E	
2015†	1089	514	107	44	64	1818
2016	930	414	62	19	11	1436
2017	852	296	52	24	20	1244
2018	824	349	87	18	21	1299
2019	635	272	61	16	21	1005
2020	574	129	16	1	3	723
Total	4904	1974	385	122	140	7525

† Excluding fires from Jan 1 to Mar 31, 2015, for which we did not request expenditure data (n = 38)
Size classes: A: 0 to 0.1 ha; B: >0.1 ha to 4 ha; C: >4 ha to 40 ha; D: >40 ha to 200 ha; E: >200 ha

Table 2.2 shows summary statistics on wildfires. While the average total cost per fire is \$101,292, the cost distribution is very wide and total cost standard deviation is \$2.2 million. 11% of fires reported no costs. The average burned area is nearly 250 hectares, however, this variable also varies widely, with a standard deviation of 7,477 ha. Most fires responded by Alberta Wildfire take place on provincial land (72%) and a large majority of fires receive immediate suppression (94%). The reporting delay, measured as the hours between fire ignition and reporting to Alberta

Wildfire also ranges widely; however, a considerable number of time entries in the datasets contain likely transcription errors in the dataset[7].

Table 2.2 Summary statistics on wildfires, 2015 to 2020

	N	Mean	Std Dev	Min	Max
Total cost per fire ($, 2020)	7525	101,292	2,234,273	0.00	142,670,231
Fires with no cost	7525	0.11	0.31	0.00	1.00
Burned area (ha)	7525	248.85	7,477.54	0.01	485,124
Provincial land	7525	0.72	0.45	0.00	1.00
Immediate suppression	7524	0.94	0.24	0.00	1.00
Reporting delay (hours)	7331	140.04	1,442.25	0.00	78,888.27

Figure 2.2 shows that, a) while most fires have been suppressed as small fires at low cost, b) a significant proportion of yearly expenditures are attributed to large wildfires of class E, exceeding 200 hectares. Moreover, in some years, a large share of Alberta Wildfire's operation expenditures can be accrued by just a handful of these large fires. In 2019, the total cost for all fires in our sample reach $282 million; half of this amount is accounted by one individual wildfire. Combined with the next four most expensive events, these wildfire operations make up 86% of the suppression expenditure in 2019 (Figure 2.3).

[7] For a detailed explanation of these data errors and our resolution of these errors, see *Exclusions of fire observations*, below.

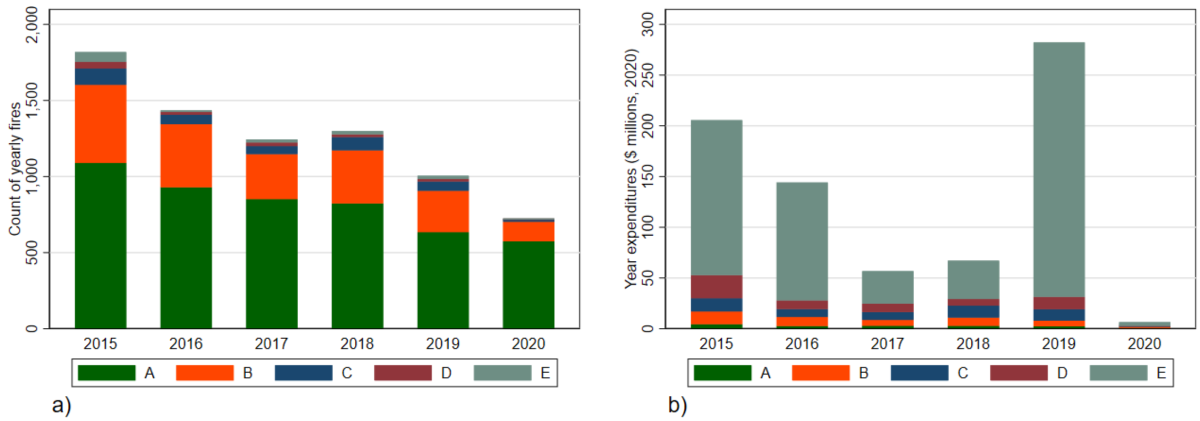

Figure 2.2 Yearly counts and yearly expenditures by size class
a) Count of yearly Alberta Wildfire fires by size class
b) Yearly suppression operation expenditures by size class (millions of 2020 dollars)

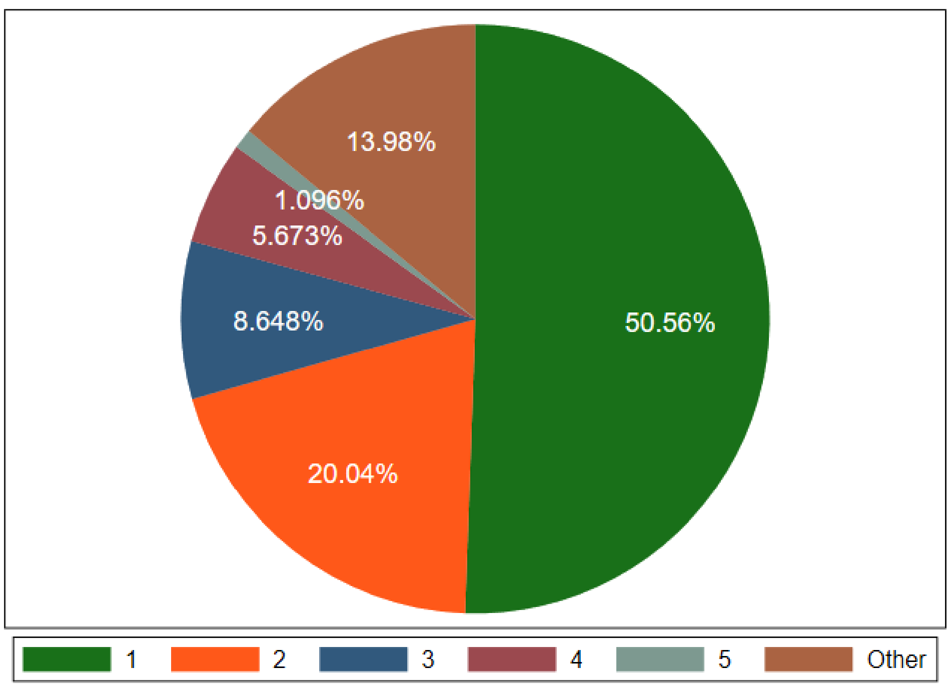

Figure 2.3 Proportion of suppression expenditure incurred by Alberta Wildfire for the five largest fires in 2019 ($282 million total)

Delay in wildfire detection and reporting also affects suppression costs, particularly for small size class A and B wildfires, under 4 hectares. *Reporting delay* measures the hours between the time a wildfire starts (or is predicted to have started) to the time it is reported to the Alberta Wildfire Coordination Centre. Figure 2.4 shows that, there is a positive correlation between log suppression cost and *Reporting delay* in most subsets of wildfires[8]. However, this relationship is only significant size in class A and B *($p < 0.01$)*[9].

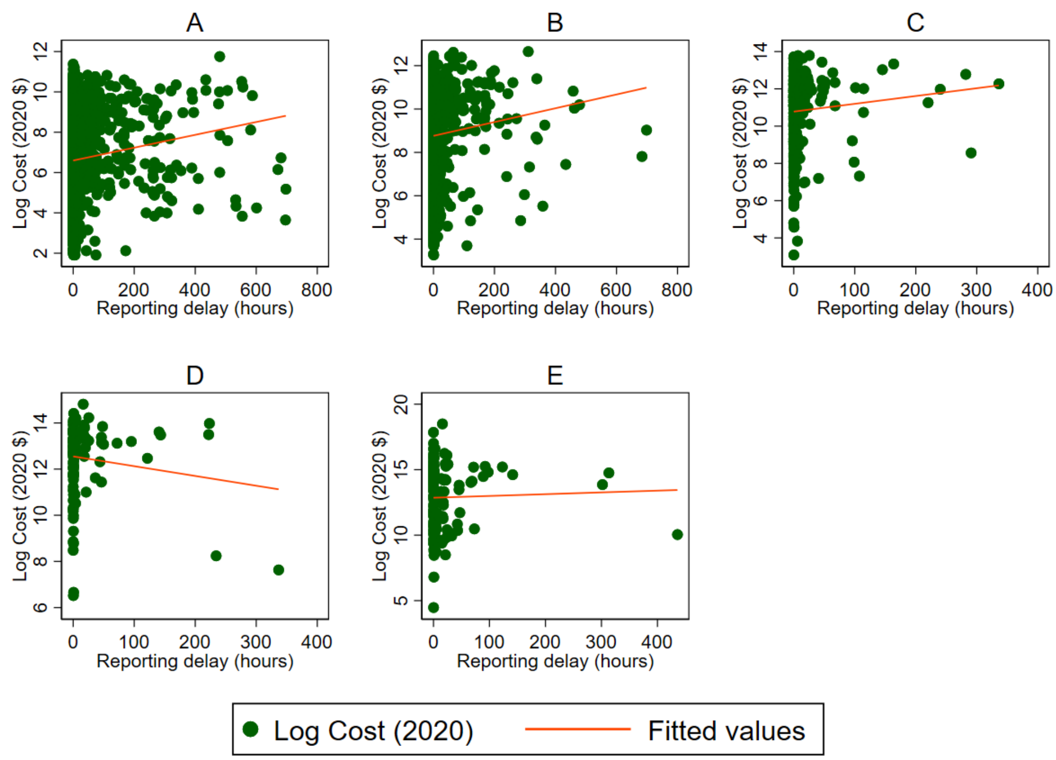

Figure 2.4 *Log cost* and *Reporting delay* by size class.
(excluding observations with ≥ 30-day Reporting delay*)*

[8] We exclude observations in which *Reporting delay* exceeds 30 days, due to concerns with transcription errors in the dataset, as well as to mitigate outliers. (See *Exclusions of fire observations*, below)
[9] For a report on the significance of the effect of *Reporting delay* on *Log* Cost, see Table C.1 in Appendix C: Auxiliary regressions

Having considered how one variable, *Reporting delay*, impacts suppression expenditures, we continue onto the next section detailing how the fire-level variables are transformed or calculated from original datasets.

Creating variables

Firstly, we collapse itemized Alberta Wildfire expenses to the fire level in order to generate the total cost of suppression per fire, adjusted to 2020 dollars. Among the reviewed wildfire expenditure literature, we see some papers specifying cost per area as the dependent variable (Donovan et al., 2011; Gebert et al., 2007; Yoder and Gebert, 2012), some others using total cost (Clark et al., 2016; Hand et al., 2016; Hand et al. 2014, chap. 4; Liang et al., 2008), or both (Gebert and Black, 2012). We choose to focus on total suppression expenditure per fire, because we believe, as argued by Hand et al. (2016) in the US context, that fire-level total costs are much more salient to a wildfire management agency.. Hand et al. contend that suppression effort: i) is applied at a portion of the free-burning perimeter, and not on a per-unit area basis; ii) is affected by the size, shape and location of a wildfire relative to landscape and values-at-risk; iii) can be influenced by resource availability and unit cost of resources; and iv) can be adjusted in response to spatial and temporal variation in fire conditions. As such, Hand et al. assert that suppression effort is better reflected through total cost, rather than cost per area. Furthermore, we notice in summary reports (MNP LLP, 2020, 2017, 2016) and through our personal communication, Alberta Wildfire also considers expenditures as fire-level costs. Thus, we choose to define the dependent variable of our empirical model is the (log of) total cost of suppression on the fire-level, adjusted to 2020 dollars.

Explanatory variables can be categorized into three categories: *Fire environment*, *Operation*, *Values-at-risk* (Table 2.3). Most variables are transformed or calculated from original datasets using Stata, and data from geospatial sources has been manipulated with ArcGIS and Stata. All variables are specified at the fire-level.

Most variables are sourced from Alberta Wildfire's proprietary Fire Information Resource Evaluation System database (FIRES). Using original variables from FIRES, we create dummy variables for *Fuel type*, *Fire type*, and *Jurisdiction*. Some observations of *Fire type* are manually categorized because they were recorded in FIRES as string values rather than categorical values. Additionally, we use the recorded times for suppression phases[10] to calculate two continuous variables: *Other fires*, a measure of resource availability, and *Reporting Delay*. Gebert et al. (2007) calculate a resource availability as the average number of fires in the same region in previous years; as we will control for the fixed effects of *year*, our variable is made of the number of concurrently burning fires on the assessment date of fire i, when the lead wildfire Incident Commander must consider what resources to order for initial suppression. When there are more fires burning across the province, we speculate that there are less resources available for immediately deployment to fire i[11]. *Reporting Delay* is the number of hours between the time of wildfire ignition (or predicted time of ignition) and the time when the fire is reported to Alberta Wildfire Coordination Centre. *Reporting Delay* is a key policy variable, as it serves as a measure of Alberta Wildfire's detection and reporting apparatus.

Alberta Wildfire also provided a dataset with daily weather observations for 499 weather stations throughout the FPA, across the timeframe of our study period. While some papers use a fire weather index to control for weather conditions (Clark et al., 2016; Gebert et al., 2007; Yoder

[10] Wildfire suppression phases: *start, report, arrival, assessment, being held, under control, extinguished, current* (final evaluation). See Figure 2.1 for a visual display of the all chronological phases.
[11] As our *Other fires* variable is calculated from assessment date +/- 2 days, so imported and newly recruited resources should not meaningfully impact total resource availability.

and Gebert, 2012), following Bayham and Yoder (2020), we elect to calculate *Temperature, Wind speed, Rain* and *Relative Humidity* from raw values in the Alberta Wildfire weather dataset, focusing on values that capture the max/min/total weather conditions on assessment date plus/minus two days which would drive fire behaviour: maximum temperature, maximum wind speed, total rain fall and minimum relative humidity. These weather variables are created from mean observations of every weather station within 50 km from the coordinates of the fire ignition point[12]. We recognize there is the possibility of two-way causal relationship between fire weather and suppression costs, particularly in large, long-lasting fires. In such scenarios, it is possible that weather conditions drive fire behaviour, which in turn affects suppression costs, while simultaneously, costs reflect suppression efforts that will alter fire growth and behaviour, which in turn can impact regional weather conditions. To mitigate endogeneity, we focus on weather observations that drive a fire in its initial period, which we define as five days centered on the assessment date.

Following Gebert et al. (2007), we control for the influence of topography on fire behaviour using calculated aspect and elevation variables. These variables are sourced from Altalis, a public-private provider of geospatial data. Using the 100-metre raster projection of the Alberta Provincial Digital Elevation Model, we create dummy variables of *South aspect* and *High elevation* (over 1250 m, which is approximately where the eastern slopes of the Rocky Mountains begin), and we account for elevation variation within the landscape as *Elevation difference* between the highest and lowest points within the "pseudo-burned area" of the fire. As geospatial data on wildfire

[12] The range and accuracy of wildfire weather stations is highly dependent on local conditions, and after consulting Alberta Wildfire we were unable to determine range precision for stations in our dataset (personal communication). At the 50 km buffer, stations cover 7,522 of 7,525 ignition points in our dataset. For more research into fire weather specification, see Cai et al. (2019); Lawson and Armitage (2008); National Wildfire Coordinating Group (2019).

perimeter is not available for all fires, we define the pseudo-burn area as a circle originating from the ignition point with an area equal to the final burned area recorded in FIRES.

Values-at-risk are the residential, industrial and infrastructure assets that are threatened by wildfire. In this model, *Values-at-risk* are dummy variables that indicate the presence of an asset within close distance to a wildfire. We define these assets using a variety of data sources: *Community* (Alberta Wildfire, internal dataset on wildfire community locations), national or provincial *Park* (Government of Alberta's GENeric Enterprise Spatial Information Services, GENESIS), *Power generation* (GENESIS) and *Road* (Statistics Canada, National Road Networks). We select these assets because they represent some of the values-at-risk mostly commonly defined in the reviewed literature (Donovan et al., 2011; Gebert et al., 2007; Liang et al., 2008), and because communities, parks, power generating stations and roads have been present and stationary in the Forest Protection Area throughout our study period, from 2015 to 2020. These *Values-at-risk* variables also represent some of Alberta Wildfire's protection priorities, which are in decreasing order of importance: Human Life, Communities, Watershed and Sensitive Soils, Natural Resources, Infrastructure (Alberta Wildfire, 2017). We could expect that the coefficient estimates on *Values-at-risk* to reflect Alberta Wildfire priorities (i.e. "important" values-at-risk receive more suppression expenditure), however we are cautious to interpret *Values-at-risk* solely through this lens. *Values-at-risk* may also capture unobserved attributes that can be a source of heterogeneity between individual wildfire suppression operations. Table 2.3 offers an overview of all variables used in the empirical model.

Table 2.3 Variables used in development of regression equation.
Dependent variable: *Wildfire suppression expenditures* (2020 dollars, log form)

Fire characteristics	Variable definition	Source	Calculated
Fire environment			
Weather	*Weather variables are max/min/total on assessment day +/- 2 days*		
Temperature	Maximum temperature (°C)	FIRES	yes
Wind speed	Maximum wind speed (km/h)	FIRES	yes
Rain	Total rainfall (mm)	FIRES	yes
Humidity	Minimum relative humidity (%)	FIRES	yes
Fuel type	Dummy variable* for *Timber* or *Slash* fuel types (reference: *Open* (grass, peat, moss) or *Manmade*)	FIRES	original
Fire type	Dummy variable for *Crown* fire type (reference: *Ground* or *Surface*)	FIRES	original
High elevation	Dummy variable for elevation over 1250 m	Altalis	yes
Elevation difference	Difference in elevation between highest and lowest elevation points of the pseudo-burned area†	Altalis	yes
South aspect	Dummy variable for mean aspect between 157.5-202.5°	Altalis	yes
Lake/River	Dummy variable for lakes/rivers (within 3km). Surface water can serve as natural boundaries, and can be used to suppress fire.	GENESIS	yes
Operation			
Other fires	Number of wildfires across Alberta on assessment day	FIRES	yes
Reporting delay	Hours of delay between fire ignition and report	FIRES	yes
Assessment result	Dummy variable for the assessment decision for *Delayed action* (reference: *Immediate Action*) ††	FIRES	original
Jurisdiction	Dummy variables for land on which the fire started: provincial, indigenous (reference: private land)	FIRES	original
Values-at-risk	*(Presence of value-at-risk within 3km)*		
Community	Dummy variable for town/hamlet/village/ settlement	Alberta Wildfire	yes
Park	Dummy variable for Provincial or National Park	GENESIS	yes
Oil and gas	Dummy variable for oil and gas facility	GENESIS	yes
Power generation	Dummy variable for power generating station	GENESIS	yes
Road	Dummy variable for road	StatCan	yes

* For variables that are dummy variables, 1 = presence of the attribute; 0 = absence.
† The pseudo-burn area is a circle centered on the ignition point, with the same area as the final burn area recorded in FIRES
†† *Delayed action* includes categories: “Beyond Resources Capability”, “Delayed Action - Lower Priority”, “Delayed Action - No Resources Available”

Exclusions of fire observations

The identifier of an observation in our empirical models is a fire. We focus on wildfires in which the Wildfire Management Branch responded from April 1, 2015 to December 31, 2020 (n = 7,525). Due to some gaps in data, as well as some outlier wildfire events, we make a series of exclusions so that the regression model has explanatory power.

i. *Missing weather observations*: no Alberta Wildfire fire weather observations within 50 km of ignition point on all assessment +/- 2 days (n = 290)
ii. *Missing fuel type or fire type*: values not recorded in FIRES (n = 921)
iii. *Missing Assessment delay*: value not recorded in FIRES (n = 1)
iv. *Missing topographic data*: missing from GENESIS elevation map (n = 2)
v. *Missing ignition / report times*: values not recorded in FIRES (n = 194)
vi. *Reporting delay greater than 30 days*: Exclusion of these observations (n = 446) serve two purposes:
    a. Wildfires that were unreported within a month of ignition may have some unobservable characteristics that make them unpredictable to suppress, and will also incur unexpected expenses,
    b. We notice a considerable number of possible transcription errors in the dataset, in which it appears month or year had been incorrectly entered (e.g., the ignition time is recorded as *2015-07-01 12:12*, and reported time is *2015-08-01 12:45* or *2016-07-01 12:45*). Most of these errors were found in entries of smaller wildfires (size class A/B). As we are unable to discern erroneous entries from real observations, removing all observations of 30+ day delay will mitigate effects of both outliers and reporting error.

vii. *Wildfire complexes*: Generally, researchers acknowledge fire-level expenditures are difficult to determine within fire complexes (Gebert et al., 2007). In personal communication with Alberta Wildfire, we were informed that expenditure identification issues in fire complexes may exist within their database. (n = 101)

viii. *Turned-over:* We exclude fires that had been turned over from Alberta Wildfire to another agency, as the total cost of suppression may not reflected in Alberta Wildfire's expenditures dataset. (n = 668)

ix. *Military land:* A set of observations of wildfires on Department of National Defense were excluded, as high costs were associated with these fires, likely due to the political and safety concerns for wildfire spread to military equipment such as munitions. (n = 30)

Additionally, we exclude *Fires with zero cost* in the main OLS model (n = 839). We specify cost in log form and the natural logarithm is only defined for variables greater than zero. After these exclusions, the final number of observations used in the main empirical model is 5,098.

Table 2.4, below, is a series of summary statistics tables for the variables used in the main empirical model, separated by size class.

Table 2.4 Summary statistics by wildfire size class

	N	Mean	Std Dev	Min	Max
Size class A (0 to 0.1ha)					
Temperature (°C)	3440	20.45	5.50	-9.10	33.00
Wind speed (km/h)	3440	16.96	6.29	3.50	50.00
Rain (mm)	3440	1.55	3.18	0.00	40.27
Relative Humidity (%)	3436	36.02	10.59	10.00	89.00
Fuel type: Timberslash	3054	0.56	0.50	0.00	1.00
Fire type: Crown fire	3063	0.02	0.13	0.00	1.00
South aspect	3518	0.22	0.41	0.00	1.00
High elevation	3518	0.24	0.42	0.00	1.00
Elevation difference (m)	3518	0.00	0.00	0.00	0.00
Lake/River within 3km	3518	0.29	0.45	0.00	1.00
Other fires	3518	27.26	28.14	0.00	125.00
Reporting delay (hr)	3518	11.69	49.34	0.00	697.34
Delayed suppression	3518	0.05	0.22	0.00	1.00
Provincial land	3518	0.77	0.42	0.00	1.00
Indigenous land	3518	0.14	0.35	0.00	1.00
Community	3518	0.20	0.40	0.00	1.00
Park	3518	0.05	0.22	0.00	1.00
Power generation	3518	0.01	0.09	0.00	1.00
Road	3518	0.74	0.44	0.00	1.00
Size class B (>0.1 to 4ha)					
Temperature (°C)	1654	20.78	5.90	-2.50	33.40
Wind speed (km/h)	1654	17.35	6.43	0.00	47.00
Rain (mm)	1654	1.05	2.50	0.00	27.93
Relative Humidity (%)	1646	34.91	10.21	11.00	100.00
Fuel type: Timberslash	1726	0.60	0.49	0.00	1.00
Fire type: Crown fire	1726	0.05	0.23	0.00	1.00
South aspect	1725	0.21	0.41	0.00	1.00
High elevation	1726	0.05	0.22	0.00	1.00
Elevation difference (m)	1726	0.38	1.93	0.00	33.72
Lake/River within 3km	1726	0.31	0.46	0.00	1.00
Other fires	1726	26.55	27.70	0.00	125.00
Reporting delay (hr)	1726	10.82	42.94	0.00	698.69
Delayed suppression	1726	0.05	0.23	0.00	1.00
Provincial land	1726	0.70	0.46	0.00	1.00
Indigenous land	1726	0.23	0.42	0.00	1.00
Community	1726	0.23	0.42	0.00	1.00
Park	1726	0.01	0.12	0.00	1.00
Power generation	1726	0.01	0.08	0.00	1.00
Road	1726	0.57	0.50	0.00	1.00

Note: See Table 2.3, above, for a detailed description of all variables.

Table 2.4 Summary statistics by wildfire size class *(continued i.)*

	N	Mean	Std Dev	Min	Max
Size class C (> 4 to 40 ha)					
Temperature (°C)	325	22.20	5.95	0.70	31.70
Wind speed (km/h)	325	17.69	6.57	5.00	43.00
Rain (mm)	325	1.04	2.76	0.00	25.23
Relative Humidity (%)	324	33.97	9.62	12.31	59.67
Fuel type: Timberslash	333	0.71	0.45	0.00	1.00
Fire type: Crown fire	333	0.21	0.41	0.00	1.00
South aspect	333	0.24	0.43	0.00	1.00
High elevation	333	0.07	0.25	0.00	1.00
Elevation difference (m)	333	7.02	17.04	0.06	230.07
Lake/River within 3km	333	0.19	0.39	0.00	1.00
Other fires	333	31.11	30.19	0.00	125.00
Reporting delay (hr)	333	10.23	35.30	0.00	336.31
Delayed suppression	333	0.07	0.26	0.00	1.00
Provincial land	333	0.78	0.41	0.00	1.00
Indigenous land	333	0.14	0.34	0.00	1.00
Community	333	0.10	0.30	0.00	1.00
Park	333	0.01	0.09	0.00	1.00
Power generation	333	0.00	0.05	0.00	1.00
Road	333	0.44	0.50	0.00	1.00
Size class D (> 40 to 200 ha)					
Temperature (°C)	89	24.38	3.96	10.00	30.50
Wind speed (km/h)	89	17.48	6.41	7.00	38.00
Rain (mm)	89	0.69	1.75	0.00	9.45
Relative Humidity (%)	89	33.37	8.81	17.00	61.81
Fuel type: Timberslash	95	0.92	0.28	0.00	1.00
Fire type: Crown fire	95	0.48	0.50	0.00	1.00
South aspect	95	0.22	0.42	0.00	1.00
High elevation	95	0.04	0.20	0.00	1.00
Elevation difference (m)	95	13.39	20.32	0.61	132.79
Lake/River within 3km	95	0.17	0.38	0.00	1.00
Other fires	95	38.73	31.53	3.00	125.00
Reporting delay (hr)	95	13.65	35.19	0.00	234.36
Delayed suppression	95	0.18	0.39	0.00	1.00
Provincial land	95	0.91	0.29	0.00	1.00
Indigenous land	95	0.07	0.26	0.00	1.00
Community	95	0.02	0.14	0.00	1.00
Park	95	0.02	0.14	0.00	1.00
Power generation	95	0.01	0.10	0.00	1.00
Road	95	0.29	0.46	0.00	1.00

Table 2.4 Summary statistics by wildfire size class *(continued ii.)*

	N	Mean	Std Dev	Min	Max
Size class E (> 200 ha)					
Temperature (°C)	85	25.18	4.25	6.30	32.60
Wind speed (km/h)	85	18.61	7.38	4.00	44.00
Rain (mm)	85	0.82	3.21	0.00	24.70
Relative Humidity (%)	85	31.80	8.77	12.31	58.35
Fuel type: Timberslash	94	0.91	0.28	0.00	1.00
Fire type: Crown fire	94	0.36	0.48	0.00	1.00
South aspect	94	0.27	0.44	0.00	1.00
High elevation	94	0.09	0.28	0.00	1.00
Elevation difference (m)	94	116.91	246.20	2.41	1465.13
Lake/River within 3km	94	0.16	0.37	0.00	1.00
Other fires	94	40.49	33.08	1.00	125.00
Reporting delay (hr)	94	24.48	65.81	0.00	435.81
Delayed suppression	94	0.30	0.46	0.00	1.00
Provincial land	94	0.99	0.10	0.00	1.00
Indigenous land	94	0.00	0.00	0.00	0.00
Community	94	0.02	0.15	0.00	1.00
Park	94	0.02	0.15	0.00	1.00
Power generation	94	0.00	0.00	0.00	0.00
Road	94	0.10	0.30	0.00	1.00

2.4. Empirical model

To determine the impacts of policy and environmental factors on wildfire expenditures, we estimate the following model:

$$y_{irtm} = \beta_W \cdot W_{irtm} + \beta_X \cdot X_{irtm} + \beta_V \cdot V_{irtm} + \lambda_r + \delta_t + \mu_m + \varepsilon_{irtm}$$

in which y_{irtm} is the log of the total suppression expenditure (in 2020 dollars) incurred on fire i, in Forest Area region r, in year t, in the month of the year m. The log specification allows us to interpret the marginal effects of each independent variable as its impact to a percentage of cost. Categories of explanatory variables include: *Fire environment* (W_{irtm}); *Operational* variables, including reporting delay, jurisdiction and other concurrent fires (X_{irtm}); and *Values-at-risk* to be protected from wildfire, including communities, national and provincial parks and roads (V_{irtm}).

Forest Area, Alberta Wildfire's administrative subregion unit, year, and month are fixed effects represented by $\lambda_r,\ \delta_t,\ \mu_m$ (Descriptions of variables are in Table 2.3).

In our specification we absorb the fixed effects of *Forest Area* region, year and month on wildfire suppression. We apply the a high-dimensional fixed effects regression function[13] to split samples of wildfires. Wildfires are split by size class (from A to E), so that we control for any unobserved heterogeneity due to size class-specific standard operating procedures (Alberta Wildfire internal documents; personal communication). Separation of wildfires by size classes also serves to control the effects of wildfire size while also omitting *Burned area*, mitigating the potential for the area variable to be endogenous with our dependent variable, suppression expenditure (Gebert et al., 2007; Hand et al., 2016).

2.5. Overview of empirical analysis

Having controlled for *Temperature, Wind speed*, and other environmental factors, we anticipate that the model will discern the effect of discretionary policy choices made during wildfire suppression operations.

We expect that to observe that an additional number of fires concurrently burning across the province creates complexity in suppression resource allocation, as reflected in a change in the level of suppression costs per fire. During a busy part of the fire season, "competition" for suppression resources may lead to either higher or lower suppression costs per fire. Higher costs can be incurred if a) existing resources are utilized beyond normal capacity (e.g. overtime

---

[13] We employ the Stata package *reghdfe* (Gaure, 2010; Guimarães and Portugal, 2010), a linear Ordinary Least Squares regression function with high-dimensional fixed effects.

payments for staff, additional fuel costs for aircraft), or b) additional resources are made available to Alberta Wildfire (e.g. through emergency contracting, or borrowing resources from other jurisdictions). Lower costs per fire could also be observed due to: a) economies of scale in firefighting, whereby each additional concurrently burning wildfire is less expensive to supress because resources are distributed more efficiently across multiple wildfires, or b) lack of sufficient resources, such that each wildfire is allocated fewer resources than usual.

In addition, this empirical model can uncover the relationship between wildfire detection times and suppression expenditures. A wildfire that has been promptly detected and reported is smaller in size and intensity. Such wildfires can be suppressed by *Initial Attack* resources in a quick and cost-effective manner. Delayed reporting means that a wildfire has the potential to develop in size and intensity in the absence of intervention. Large, fast-burning fires will require more personnel, equipment and/or aircraft for control and suppression, thus incurring a higher level of expenditure.

Through our empirical model, we are also interested in finding how the operational decision to delay immediate suppression will influence total suppression costs. Gebert and Black (2012) propose that direct suppression results in lower total expenditure, reasoning that a wildfire that does not receive immediate suppression may develop in a manner that will require more resources later on.

In our model it will be possible to observe that wildfires in provincial and indigenous lands incur higher expenses than those on private lands, because of Alberta Wildfire's jurisdictional mandate. With the inclusion of values-at-risks (communities, parks, power generating stations, roads), the model can demonstrate how priority infrastructure are protected, as demonstrated through heightened levels of suppression expenditure. The proximity to values-at-risk can have an ambiguous effect on expenditures. Suppression of fires close to communities, parks and power generating stations may receive more expenditure, reflecting Alberta Wildfire's mission to protect these values-at-risk. However, the effect of nearby roads on costs remains uncertain because they serve both as access for suppression units, thus reducing operational costs, but as key infrastructure requiring protection, may incurring higher costs. It is therefore difficult to interpret estimates of coefficients of values-at-risk variables given that our data does not have the variation to allow us to disentangle the effects of Alberta Wildfire's mission (that is constant across the sampling period) from those of other expenditures confounding factors.

2.6. Results

Table 2.5 presents the coefficients of our empirical model estimated in a split sample approach where each split is determined by fire class (from A to E). Each column shows the results of one empirical model (or one split sample). For example, column A shows the estimates of the 2,965 class A fires in our sample. The estimates represent the effect of a factor on the log expenditure in each wildfire suppression operation.

Table 2.5 OLS regression models on Log Expenditure (2020 dollars), by wildfire size class

	A	B	C	D	E
Fire environment					
Temperature (°C)	0.030**	0.034**	0.059**	0.080	0.380***
	(0.010)	(0.012)	(0.022)	(0.114)	(0.063)
Wind speed (km/h)	0.001	0.017*	0.004	0.054**	0.000
	(0.006)	(0.008)	(0.009)	(0.020)	(0.026)
Rain (mm)	-0.024*	-0.005	-0.053**	0.007	-0.138***
	(0.012)	(0.014)	(0.017)	(0.093)	(0.025)
Relative Humidity (%)	-0.007	-0.007	0.035***	0.049	0.063
	(0.007)	(0.005)	(0.009)	(0.034)	(0.061)
Fuel type: Timberslash	1.275***	1.164***	1.215***	0.602*	-0.669
	(0.156)	(0.071)	(0.354)	(0.275)	(1.354)
Fire type: Crown fire	-0.173	0.482***	0.094	0.211	0.248
	(0.277)	(0.130)	(0.174)	(0.112)	(0.376)
South Aspect (true south)	0.023	0.170*	0.087	-0.048	1.110*
	(0.130)	(0.077)	(0.244)	(0.394)	(0.534)
High elevation	0.085	0.210	0.537	0.000	0.000
	(0.212)	(0.155)	(0.365)	(.)	(.)
Elevation difference (m)	0.000	0.147***	0.014	-0.005	0.007
	(.)	(0.025)	(0.008)	(0.003)	(0.004)
Lake/River within 3km	-0.136	-0.047	0.115	0.248	0.184
	(0.092)	(0.068)	(0.231)	(0.415)	(1.087)
Operation					
Other fires	-0.000	-0.002	-0.005	-0.001	-0.024**
	(0.002)	(0.002)	(0.005)	(0.008)	(0.009)
Reporting delay (hr)	0.002***	0.002***	0.001	-0.010	0.006
	(0.000)	(0.000)	(0.002)	(0.007)	(0.005)
Strategic delay	-0.223*	-0.465***	-1.532**	-1.409***	-0.914
	(0.102)	(0.094)	(0.601)	(0.115)	(0.722)
Provincial land	0.276	0.401*	0.759*	0.049	0.000
	(0.200)	(0.194)	(0.359)	(0.156)	(.)
Indigenous land	0.315	0.001	0.100	-0.662	0.000
	(0.172)	(0.171)	(0.442)	(0.532)	(.)
Values-at-risk (within 3 km)					
Community	-0.449***	-0.578***	0.332	1.725***	1.671
	(0.086)	(0.172)	(0.484)	(0.204)	(1.071)
Park	0.174	0.112	1.155	0.000	0.874
	(0.193)	(0.243)	(0.632)	(.)	(1.567)
Power generation	0.393	-0.194	1.178***	3.268***	0.000
	(0.287)	(0.418)	(0.331)	(0.365)	(.)
Road	-0.573***	-0.318***	-0.398**	-0.126	0.227
	(0.076)	(0.088)	(0.145)	(0.226)	(0.875)
N	2,965	1,644	323	85	81
R-squared	0.431	0.522	0.536	0.617	0.564

All models include corporate region, year, and month of the year fixed effects. Standard errors clustered at the corporate region are in parenthesis. Size classes: A: 0 to 0.1 ha; B: >0.1 ha to 4 ha; C: > 4 ha to 40 ha; D: >40 ha to 200 ha; E: >200 ha.
* p<0.10, ** p<0.05, *** p<0.01

Results of these models are outlined below; a detailed discussion of these results follow in Section 2.8, below. In general, coefficients on *Fire environment* variables have the expected signs. Higher temperature and windspeed during the initial days of the fire exacerbate conditions, making fires more expensive to fight, while additional rainfall reduces costs. For every additional Celsius degree increase in temperature, the cost of suppressing class A and B fires increases by 3%; however, an additional degree makes class E fires (200 ha+) 38% more expensive. On average, an additional millimetre of rain during the initial fire period will reduce expenditures for a class E wildfire close to 14%. The significantly positive effect of relative humidity on class C fires is unexpected, because low relative humidity correlates with extreme fire weather (MNP LLP, 2017). However, given that relative humidity is a determinant of fire ignition (Adab et al., 2013; Plucinski et al., 2014; Vasilakos et al., 2009) and our observations are those of already ignited fires, it is possible that this variable is not as pertinent to determining fire behaviour and suppression costs. Additionally, it is possible that the variable *South aspect,* which identifies landscapes receiving more sunlight, may be capturing some of the variation that would be identified by *Relative humidity*. Indeed, suppression of class B and E fires on a south aspect is marginally significant in additional cost.

Compared to the costs of fighting fires on grass or manmade fuels, suppression of class A, B and C wildfires are more than twice as expensive when flames catch on standing or slashed trees. As well, class B fires are 48% more expensive when flames reach tree crowns. The positive correlation between timber/slash fuel and higher suppression costs can be expected, as fires will tend to burn more intensely on these fuels (Clark et al., 2016; Gebert et al., 2007; Gebert and Black, 2012). While fires above 1250 m elevation (mainly located in the eastern slopes and some parts of the Grande Prairie region) are not significantly more expensive, expenditures for class B fires increase with landscape variation in the order of 15% per additional metre of elevation difference between the highest and lowest points. Lakes or rivers act as natural barriers or water sources for suppression; however, the results show that the presence of water bodies anywhere within a 3km radius of a fire is insignificant.

Coefficients on *Operations* variables demonstrate the impact of Alberta Wildfire's protocols and response strategies. For every additional fire that is concurrently burning in the province, Alberta Wildfire will spend less resources on each individual wildfire, although the impact of *Other fires* is only significant in large class E fires (>200 ha). Delay between fire ignition and reporting to Alberta Wildfire also significantly and positively impacts suppression costs of class A and B fires (up to 4 ha), though it is a relatively small impact of 0.2% for every additional hour.

Results indicate that the strategic decision by the initial response team to delay immediate suppression action does not cause wildfire suppression to be more expensive. Compared to fires that receive immediate suppression, fires in which resource allocation is delayed (due to overwhelming fire conditions, resource non-availability or low suppression priority) have significantly reduced expenditures, by order of 47% for class B fires, 153% for class C fires and 140% for class D fires.

Alberta Wildfire's jurisdiction is limited to provincial public lands in the Forest Protection Area, as well as select Government of Canada jurisdictions (including Indigenous Services, National Defense). Compared to suppression efforts on private land, class B and C wildfires on Crown land are marginally significant in receiving additional suppression expenditures (respectively, 40%, 76%). Suppression costs on indigenous lands (incl. First Nations and also Metis settlements) are not significantly different from costs on private land; however, it is possible that some degree of variation may be captured by the *Values-at-risk* variable *Community*, which also includes communities on indigenous lands. While we exclude *Military land* from our main analysis due to relatively few observations (n = 30), Table C.3 shows that size class B and C fires on National Defense bases receive significantly more additional suppression costs, compared to fires with otherwise similar conditions off-base.

The presence of certain *Values-at-risk* tend to also influence suppression expenditures. The effect of *Community*, the presence of a community within 3 km of ignition, reduces costs by 45% and 58% for size class A or B fires, respectively. However, if a wildfire close to a community grows to size class D levels (>40 ha to 200 ha), its suppression becomes 173% more expensive. The observed effect may be attributed to better detection/access for small fires close to communities, as well as an impetus to suppress large wildfires that threaten lives and property that are concentrated in communities. The cause of this effect is discussed further in Section 2.8, below.

Larger fires also require additional suppression costs when a power generating station is within its 3 km vicinity; class C fires will 118% more costly, and class D, 327%. Access to the wildfire, as captured by *Road* within 3 km, makes suppression of fires up to 40 ha significantly less expensive, reducing costs by 57%, 32% and 40% respectively for classes A, B, and C. The presence of a national or provincial park within 3 km can induce more expenditures in suppressing class A, B, C and E fires, although this relationship is not significant.

Estimating components of expenditure

We estimate simplified expenditure models in order to determine what the proportions of expenditure variation can be attributed to environmental and policy variables:

$$y_{irtm} = \beta_W \cdot W_{irtm} + \beta_r \cdot F_r + \delta_t + \mu_m + \varepsilon_{irtm}$$

where W is a vector of *Fire environment* variables and F represents *Forest Area* dummy variables. We compute mean predicted expenditure per fire of each size class (mean $\hat{y}$) by making partial predictions at the mean of *Fire environment* $\hat{\beta}_W\overline{W}$ and *Forest Area* effects $\hat{\beta}_r\overline{F}$. While the effects on *Fire environment* variables are time and region specific, the estimates on *Forest Area* represent effects that are time-invariant. The *Forest Area* effect captures a wide array of policy impacts that are not captured in our dataset, including the influence of fixed regional managers, resource capacity and operational preparedness. Estimates from *Forest Area* demonstrate how operational conditions at the regional level impact costs; however, we recognize these estimates could also be capturing the impact of other unobserved environmental factors that are fixed over time. To make more accurate predictions, we exclude observations of wildfire that were above the 95th expenditure percentile of each size class.

Table 2.6 Estimates of expenditure proportion (Environment and Forest Area)

	Size class			
	A		B	
Count	3042		1582	
	$	%	$	%
Mean total cost per fire	3040	*100%*	16542	*100%*
Environment	2402	*79%*	12886	*78%*
Forest Area	755	*25%*	104	*1%*
Residuals	-117	*-4%*	3552	*21%*

Observations above 95% percentile of expenditure level per size class have been excluded. Year and month have been absorbed.

As costs predictions are more robust in large samples without extreme values, in Table 2.6 we report on the expenditure component estimates for size classes A and B (with, 3,042; 1,582 observations, respectively). In the postestimation models for classes A and B, we observe *Fire environment* variables making up a large proportion of expenditures in suppressing fires under 4 ha. For instance, Alberta Wildfire spends an average of $16,542 to suppress a class B (>0.1 to 4 ha) wildfire; within this cost, $12,886 (78%) can be attributed to conditions that are beyond the agency’s control, while $104 (1%) can be attributed to internal regional organization or unobserved time-invariant environmental factors.

2.7. Evaluating impacts of wildfire detection with Machine Learning

Thus far, we have focused on the impacts of wildfire suppression operations, and particularly, on how these impacts are reflected in expenditures. In this section, we highlight a critical policy variable that drives the momentum of every wildfire suppression mission: *Reporting delay*.

Alberta Wildfire’s response framework begins as soon as a wildfire is detected in the Forest Protection Area (FPA), reported to the Alberta Wildfire Coordination Centre (AWCC) which relays the incident to its respective Forest Area region[14]. Only then are resources mobilized to the fire for assessment and suppression action, ending the response process at extinguishment. Within this strategic framework, the preliminary stages of detection and reporting are critical, because the early and precise detection of wildfires allows decisionmakers the necessary time and information to implement an appropriate response.

Currently, wildfire detection depends largely on regular patrols by crews on the ground, on water, and in the air, as well as on public reporting. Alberta is also one of the last jurisdictions in Canada to keep an extensive network of manned lookout towers, which Alberta Wildfire maintains is critical for precise wildfire detection in a populous wildland-urban interface that covers much of the FPA (Cheek, 2021). In 2021, Alberta Wildfire has invested over $4.3 million in piloting the use of tools such as cameras and drones (CBC News, 2021; Cheek, 2021). This technological

[14] see Appendix A: Maps of the FPA and FAs for maps of FPA and FA regions

renewal seeks to improve detection capacity, as well as to help the agency in its resilience to budget shocks that have impacted staffing in recent fire seasons (MacVicar, 2019).

Existing empirical research has proven that early detection is critical for two measures of wildfire suppression outcome: total cost and fire duration. Early detection is proven to be essential for capping suppression costs (Steele and Stier, 1998), and response time correlates with the duration of suppression (Arienti et al., 2006; Hirsch et al., 2004). Through our initial reduced form exploration assuming a linear relationship between reporting delay and log costs, we observe that small wildfires (class A and B) are significantly less expensive to suppress when they were reported sooner. Figure 2.4 in Section 2.3: Data shows that, there is a positive correlation between log suppression cost and *Reporting delay* in most subsets of wildfires. Likewise, earlier reporting tends to diminish the time it takes to extinguish small and intermediate sized fires (see Figure B.1 in Appendix B: Auxiliary figures).

However, the true functional form of the relationship between reporting delay and measures of suppression outcome remains unclear. For example, we do not know the real parametric functional form of the influence of Temperature, Wind speed, Elevation, etc. on the suppression outcome; if the form is misspecified, our estimated coefficients on the effect of the policy variable on the suppression outcome may be biased and inconsistent. Thus, in addition to Ordinary Least Squares (OLS) modelling, we also apply a Machine Learning (ML) method in analyzing the effect of a key policy variable on wildfire suppression outcome. We are interested in how ML explores non-linear relationships between explanatory variables and the outcome of interest (costs or duration), as well as how ML captures the influence of confounding variables (in our case, those of the *Fire environment*) on our key policy variable, *Reporting delay*. In summary, the application of ML overcomes a prominent shortcoming in OLS models: the assumption of a linear relationship between explanatory variables and outcome.

However, typical supervised ML methods can also be too flexible, leading to both regularization and overfitting biases. To address these issues, Chernozhukov et al. (2018) propose the double/debiased machine learning (DML) method, in which regularization and overfitting are addressed via orthogonalization and cross-fitting methods[15]. In addition to these benefits, DML is also favourable for our dataset because this method is well suited to analyze data with relatively high dimensionality and small observation size.

We apply DML to estimate a simplified model: the impact of *Reporting delay*, a measure of wildfire policy (P) on wildfire suppression outcome (Y), controlling for *Fire environment* factors (X) and year effects (δ). Two measures of wildfire suppression outcome (Y) are investigated: 1) *Log cost* per fire, the same dependent variable as our main empirical (Table 2.5), and 2) *Fire duration* between report status to extinguished status. DML is applied to the following partially linear model:

$$Y_{it} = \beta P_{it} + g(X_{it}, \delta_t) + \varepsilon_{it} \quad (1)$$

$$P_{it} = m(X_{it}, \delta_t) + \mu_{it} \quad (2)$$

in which $g(\cdot)$ and $m(\cdot)$ are nuisance parameters and β is the main parameter of interest, indicating the effect of policy variable *Reporting delay* (P_{it}) on the wildfire suppression outcome of fire i at

[15] DML has been applied in various field settings, allowing flexible functions to capture the influence of a policy variable (P) on an outcome variable (Y), as well as the effect of confounding variables (X) on both P and Y. For example: temperature on an energy consumption policy variable and energy efficiency outcome (Burlig et al., 2020); weather on a public health policy variable and COVID-19 social distancing outcomes (Holtz et al., 2020); Airbnb rental property attributes on a professional/amateur renter policy variable and revenue outcomes (Casamatta et al., 2022).

year t. Equation (1) models the process determining wildfire suppression outcomes, including policy variable P_{it}, while also allowing environmental variables (X_{it}) and year (δ_t) influence outcomes through $g(\cdot)$. Equation (2), though not of main interest, plays an important role in modeling of X_{it} and δ_t on policy variable *Reporting delay* (P_{it}). Both functions $g(\cdot)$ and $m(\cdot)$ are flexible in functional form and allow X_{it} and δ_t to potentially influence both outcome and policy variable.

We apply 400 iterations of the DML algorithm to our set of wildfires. In each iteration b the dataset is randomly split, and an orthogonalization process finds the residuals of P_{it} and Y_{it} from X_{it} and δ_t, in order to compute $\hat{\beta}_b$.[16] Lastly, the algorithm averages $\hat{\beta}_1 \ldots \hat{\beta}_{400}$ to obtain the empirical distribution $\hat{\beta}$ which approximates its true distribution. To control for size effects and unobserved heterogeneity in environmental and policy factors due to size class, DML is applied separately across subsets of size class. Due to small sample sizes (n < 90), we do not estimate DML models for size classes D and E.

Table 2.7 DML estimates of *Reporting delay* effect (β), by size class

	A	B	C
Panel A: Log cost			
$\hat{\beta}$	0.00256***	0.00243***	0.00284***
	(5.2423e-06)	(9.4944e-06)	(3.1415e-05)
N	2,965	1,645	324
Panel B: Fire duration			
$\hat{\beta}$	0.11390***	0.11707***	0.71173***
	(2.8150e-04)	(7.3472e-04)	(8.3587e-03)
N	3,247	1,668	327

Size classes: A: 0 to 0.1 ha; B: >0.1 ha to 4 ha; C: > 4 ha to 40 ha.
Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01

[16] Letting residual 1 be the residuals from a ML prediction (using random forests) of Y from X and δ, and residual 2 are the residuals from the random forest prediction of P from X and δ, the orthogonalization is complete with when the estimate of β is obtained by the OLS regression of residual 1 on residual 2 (Chernozhukov et al. (2018)).

In Table 2.7, we observe that every additional hour of *Reporting delay* increases suppression costs by a statistically significant measure of 0.24% or more across size classes A to C. Wildfire suppression operations also last longer when reporting is delayed. The estimates of *Reporting delay* can be interpreted as: for every additional hour of *Reporting delay*, class A fires will burn an additional 6 minutes 50 seconds ($p < 0.01$), class B for 7:01 ($p < 0.01$), class C for 42:42 ($p < 0.01$).

Overall, the estimates of $\hat{\beta}$ in DML are similar to OLS estimates (for a comparison, see Table C.4 and Table C.5 in Appendix C: Auxiliary regressions). In addition to being robust to flexible functional forms, DML predicts the relationship between *Reporting delay* and suppression outcomes with greater precision, while also capturing the potential impact of *Fire environment* variables on *Reporting delay*.

2.8. Discussion

Results from a linear regression model on wildfire expenditures shed some light on how discretionary policy choices taken by Alberta Wildfire influences costs in suppressing wildfires in Alberta's Forest Protection Area. Using a postestimation model, we find that 78%+ of wildfire expenditure can be attributed to the effects of *Fire environment* variables which are beyond the operational control of the Alberta Wildfire Management Branch.

Operations

During a busy period in the fire season, suppression costs tend to be lower per individual fire tend. For every additional fire that is concurrently burning in the province, suppression of a large class E fire (> 200 ha) receives 2.4% less in expenditure. This result can be interpreted in two manners:

Firstly, it is possible that Alberta Wildfire operations benefit from economies of scale, as the marginal cost of every additional resource is less than the previous one. Alternatively, it is possible that with more demand for resources across the entire Forest Protection Area, wildfire managers for each class E fire simply need to make do with less available resources.

Delay between fire ignition and reporting to Alberta Wildfire is highly significant in modest costs to suppressing small fires. Every additional hour of delay makes class A (≤ 0.1 ha) and B fires (> 0.1 to 4 ha) 0.2% more expensive to suppress. While the additional cost is modest, considering that 90% of the 5,098 fires in our study sample are class A and B fires, such an impact can certainly add up throughout a fire season. This result demonstrate the importance of Alberta Wildfire's detection protocol (Alberta Wildfire, 2019). Through a network of detection agents (air, water, ground patrol, and lookout towers) as well as engaged industry and citizen stakeholders, Alberta Wildfire's array of instruments for wildfire detection is critical for keeping small both burned areas and expenditures.

In addition, we observe a significant impact of the initial wildfire responder's assessment decision between immediate suppression or delayed suppression. Results from the regression show that when suppression is delayed, either due to lower prioritization or resource capacity limits, this decision does not increase the total suppression cost. This strategy, which we call "strategic delay", is highly significant in reducing suppression costs by 47% for class B fires, and up to 141% for larger class D fires (>40 ha to 200 ha). Thus, we are inclined to interpret this finding as one that demonstrates wildfire managers' expertise in judgement, through which they are able to prioritize operational objectives while being cognizant of resource and expenditure limits. When a wildfire is evaluated as not requiring immediate attention, Alberta Wildfire will exercise a form of strategic delay in order to prioritize other operational objectives.

Alberta Wildfire's costs of suppressing wildfires on provincial Crown land (public land and parks) is higher than those of fighting comparable fires on private land, but this effect is only marginally significant in class B and C. Suppression on indigenous land is not significantly different than on private land, however, it is likely that much of the expenditure variation is captured by the *Community* variable, which includes communities on First Nations and Métis settlements.

Values-at-risk

We use variables for *Community, Park, Power generation,* and *Road* to measure how Alberta Wildfire prioritizes values-at-risk through the level of expenditures incurred in their protection. In Table 2.5, we note that the presence of communities within 3 km of a small fire (class A and B, under 4 ha) makes suppression roughly half as expensive when compared to situations where there are no nearby communities. However, as wildfires increase in size, suppression costs increase due to the presence of communities within 3 km, a relationship that is significant for class D fires (> 40 ha to 200 ha). We believe this is because small fires near communities are more likely to be noticed and reported by residents, and can thus receive prompt suppression. In contrast, when large fires are threatening communities, higher levels of suppression costs are incurred to protect human lives and residential infrastructure.

Likewise, sizable wildfires ranging from 4 ha to 200 ha (class C and D) are two to four times more expensive when they encroach on power generating stations. Understandably, Alberta Wildfire would allocate more resources towards ensuring key energy infrastructure are secure. As discussed above, there is uncertainty on the effect of roads on expenditures, as they are utilized as access for lower costs, yet they are also critical infrastructure that merit protection. In Table 2.5, we see that the presence of a road within 3 km of a wildfire makes it significantly less costly to suppress when compared to remote fires with similar conditions. Cost reductions attributed to proximity to roads are: 57% for size class A fires *($p < 0.01$)*, 32% for class B *($p < 0.01$)* and 40% for class C *($p < 0.05$)*. However, the cost reduction effect of roads drops out of significance for larger size classes, though it does become modestly positive in the model of size class E fires. These results demonstrate that while roads can facilitate access to resource mobilization for smaller wildfires, their impact on costs falls out for large wildfires. It is possible that suppression of large wildfires (class D, E) do not rely on access by roads, as those fire conditions necessitate additional air resources.

While results from our main empirical model demonstrate how policy choices and value-at-risk protection priorities impact costs, the postestimation and debiased machine learning (DML) analyses provide additional insight into the effects of wildfire suppression policy on cost variation. Postestimation results show that unobserved organization effects play a modest role in explaining cost variation, while environmental factors drive much of wildfire expenditures. Results from DML show that prompt notification through the Alberta Wildfire's detection and reporting apparatus is critical in reducing expenditures in size A to C, as well as duration of these fires.

Policy implications

Findings from this research are generally consistent with existing expenditure literature in estimates on fire environment properties, reporting delay protection of communities (Donovan et al., 2011; Gebert et al., 2007; Hand et al., 2016; Liang et al., 2008). Focusing on the interpretation of policy and environmental factor effects on wildfire expenditure, we highlight some key points of consideration for Alberta Wildfire.

Firstly, we must recognize that the large proportion of expenses incurred in wildfire suppression is likely beyond the control of Alberta Wildfire policy. When compared to the impact of unobserved Forest Area management effects, environmental factors such as temperature, wind speed, and elevation will drive the overwhelming majority of costs (78%+). Timber/slash fuel type is also a driver of higher expenditures, doubling the costs associated with wildfires up to 40 ha. While Alberta Wildfire can have an impact in reducing the flammable fuels in vulnerable areas, through fuel management and encouraging FireSmart in wildland-urban interface communities, extensive landscape changes could also spur negative environmental consequences (Rhemtulla et al., 2011).

Secondly, we consider the impact of multiple concurrent wildfires, which also serves as a proxy for resource availability. While wildfire managers would appreciate having a wide array of resources and suppression strategies available at hand, it is possible that having an unlimited arsenal (and accompanying war chest) may not be optimal operationally, or financially. On one hand, additional resources will allow firefighters to undertake early and intense action through direct suppression and line construction to prevent wildfire spread. Yet on the other hand, availability of resources may also encourage excessive resource use beyond an optimal level of operational efficiency (Donovan and Brown, 2005; Gebert et al., 2007).

In our main empirical model, we adapted Gebert et al. (2007)'s specification of resource availability (average number of wildfires in the region over previous years), by calculating the number of concurrent fires during fire i's assessment day, +/- 2 days. In this model, we find that each fire incurs marginally fewer expenditures when additional fires are concurrently burning throughout the Forest Protection Area. This effect is significant *($p < 0.05$)* for class E fires: for every additional wildfire concurrently burning, 2.4% fewer expenditures are incurred for each class E fire. While it is possible that Alberta Wildfire gains some form of economies-of-scale in suppression effort when scaling up resource deployment, particularly in fighting small wildfires, it is more likely that the effect of resource availability on costs in class E fires may instead be indicative of resource strain. Thus, this finding would drive home the importance of keeping wildfires at a manageable size, such that additional resources can be made available when large wildfires need to be suppressed.

Thirdly, we would like to bring attention to the effect of reporting delay. For every additional hour of delay between fire ignition and reporting to Alberta Wildfire, a small fire (class A, B) become significantly more expensive by 0.2% *($p < 0.01$)*. This impact is modest, but considering that nearly 9 in 10 wildfires belong to a small size class, the effect of reporting delay can contribute to considerably higher overall expenditures through a fire season. Thus, this finding encourages Alberta Wildfire's continued investment into advanced detection methods that will reduce reporting delay.

Finally, we highlight the effect of the decision to delay initial wildfire suppression, which we term "strategic delay". We find evidence that strategic delay does not increase total costs per fire for fires of any size class. This finding leads us to believe that effective operational decision-making in Alberta Wildfire helps to reduce wildfire costs, by prioritizing resource allocation to the wildfires that require it the most. However, we would be reluctant to interpret this estimate as certain proof of operational effectiveness, considering the limitations of an expenditure model, as discussed above. While strategic delay is associated with lower suppression costs in wildfires up to 200 ha (size classes A through D), expenditure analysis does not take into account how this decision impacts the achievement of certain operational objectives (e.g. to protect a certain value-at-risk in the area), nor do we establish what a counterfactual scenario might have been if immediate action had been taken.

2.9. Conclusion

Understanding the key impacts of environmental and policy variables on wildfire expenditures is critical for wildfire management agencies, who must adapt to a challenging ecological future, often equipped with similarly challenging budgets. Findings from this chapter can inform Alberta's Wildfire Management Branch how expenditures incurred over a six-year period of wildfire suppression can be explained by both the factors that largely lie outside of Alberta Wildfire's control, including weather and other fire environment variables, as well as by Alberta Wildfire's operational decisions.

The main empirical models and accompanying postestimation models show that a large proportion of variation in fire costs can be attributed to environmental factors. As climate change continues to shape a fire landscape that is increasingly severe and unpredictable (Flannigan et al., 2000; Robinne et al., 2016; Tymstra et al., 2021; Wotton et al., 2017), Alberta Wildfire can expect that extreme weather conditions, particularly that of temperature, will continue to drive a large component of suppression expenditure. Nonetheless, operational decisions to reduce *Report Delay* and practice *Strategic delay* can deliver significant, albeit modest, cost reductions. Machine learning analysis also supports the causal relationship between reporting delay and costs, while accounting for the possibility that environmental factors affect both explanatory variable and dependent variable.

As acknowledged by expenditure modelling pioneers Gebert et al. (2007), expenditure analysis is limited in its ability to fully evaluate wildfire management programs. We recognize that, while our models shed some light on the causes of variation in suppression expenditure variation, their results alone do not provide holistic insight into the efficacy of suppression procedures. Nonetheless, the findings from regression, postestimation and double/debiased machine learning analyses provide some novel insights into the impact of environmental and policy variables on wildfire suppression costs in the Alberta context.

Into the future, we hope that expenditure modelling can be applied to datasets in other Canadian jurisdictions in order to further the understanding of wildfire operations across Canada. Considering that provincial wildfire agencies often share resources during extreme regional fire conditions, expenditure models could be improved by including the effects of inter-provincial (and occasionally international) resource collaboration.

Chapter 3. Risk preference experiments with wildfire management

Wildfire Incident Commanders (ICs) are familiar with making risky choices in their professional lives, and the choices they make can have profound impacts. In this chapter I measure Alberta Wildfire Incident Commanders' risk preferences through a series of risk elicitation methods typically used in the economics and psychology literatures. Results indicate that ICs, although engaged in a risky profession, do not differ significantly from the Control group in risk aversion across all elicitation methods; however, we observe that ICs' revealed and stated risk preferences can change between methods and across contexts. Further, we find that Incident Commanders' levels of revealed risk aversion inversely correlates with their level of recent operational experience, a result that suggests a link between risk preferences in the lab and in the field.

3.1. Introduction

From rappelling off helicopters to working over 12 non-stop hours in a day in thick smoke and heat, wildland firefighters face elevated occupational health risks (Adetona et al., 2016; Government of Alberta, 2018). Fighting forest fires is a physically and mentally challenging career, due to the inherent safety risks involved, as well as the gravity of the work in protecting lives, communities and infrastructure from wildfire in Alberta's Forest Protection Area. An identity with the arduous nature of this work is embraced by Alberta Wildfire, as exemplified in their recruitment campaigns (Alberta Wildfire, 2015).

Wildland firefighters' risk preferences have been assessed in the risk literature through a stated preference survey on preferred suppression strategies for hypothetical wildfire scenarios (Hand et al., 2015). However, there remains a lack of empirical research on assessing wildland firefighters' risk preferences through their behaviour in incentivized tasks, and through their self-evaluation. While wildland firefighters may be perceived by the public as risk-seekers (Desmond, 2009), in our research, we find some evidence that firefighters can be in fact comparatively risk-averse in certain contexts, both in revealed risk elicitation methods and in self-evaluated questionnaires. This chapter addresses the following questions:

- *Do wildland firefighters exhibit more or less risk aversion, when compared to the general public?*, and,
- *What factors make some wildland firefighters riskier than others?*

The study adds to the risk elicitation literature in evaluating the merits of four commonly used risk elicitation methods (Abdellaoui et al., 2008; Eckel and Grossman, 2008; Gneezy and Potters, 1997; Holt and Laury, 2002) in combination with a self-reported risk tool (Nicholson et al., 2005). As well, this research investigates how operational experience plays a role in shaping the risk attitudes of Incident Commanders, Alberta's professional wildland firefighters. The study of wildland firefighters' risk preferences can be extended in future research in investigating how individual risk preferences shape operational decisions, and consequently, affect expenditures incurred in wildfire suppression.

3.2. Related literature

The concept of risk is considered fundamental to economic theory (Arrow, 1965; Bernoulli, 1738; Markowitz, 1952), and various elicitation methods (EMs) have been developed to elicit risk preferences, both within the field of economics and beyond. One method of classifying risk aversion research methods is by separating them into: experimental and non-experimental measures. In general, experimental economists have sought to design EMs that reveal subjects' "true" risk preferences (see review: Charness, Gneezy, and Imas 2013). Psychologists have produced questionnaires seeking subjects' self-evaluated risk tolerance on Likert scales, both in the general context (SOEP, Wagner et al. 2007) and in domain-specific contexts, such as in Financial and Recreational domains (Blais and Weber, 2006; Nicholson et al., 2005; Weber et al., 2002). In addition, researchers across both disciplines have used experimental and non-experimental measures to test EM instrument validity, as well as to measure across-method consistency (Frey et al., 2017; Pedroni et al., 2017).

A supplementary review of other forms of risk behaviour research, which have not been used in my research, is included in Appendix J.

Below, I first introduce experimental and non-experimental methods most commonly used in the risk aversion literature, with particular attention on the instruments that were operationalized in my project. Afterwards, I provide a summary of literature that has aimed to compare measures of risk preference within economics, between economics and psychology, and briefly introduce some research in the interdisciplinary field of cognitive science.

3.2.1. Revealed risk preference in the economics laboratory

The Becker-DeGroot-Marschak (BDM) method for willingness to pay elicitation is a pioneering revealed-preference mechanism which incorporates elements of risk aversion; since its 1964 publication, this method has been employed in a wide variety of contexts (Glaeser et al., 2000; Harrison and List, 2004). However, lottery lists are the first format of elicitation methods designed specifically as a tool to measure risk aversion, starting with Binswanger's field experiment in rural India (1980). Since then, a variety of lottery list elicitation methods (EMs) have been developed for revealed risk elicitation (Charness and Gneezy, 2012; Filippin and Crosetto, 2016; Harrison and List, 2004; Mata et al., 2018), though few have been as widely applied as the methods of Holt and Laury (2002) and Eckel and Grossman (2008). For a review of the most commonly used economic measures risk preferences, see Charness, Gneezy, and Imas (2013).

The following are introductions to commonly used lottery-style EMs, and a brief overview of other methods applied in economics. Firstly, I present a primer on the terminology commonly used in risk elicitation literature and in the descriptions below:

3.2.1.1. Lottery game terminology

In a lottery elicitation method (EM), subjects are tasked to make a single lottery choice (among a set of lotteries), or multiple choices down a list of paired lotteries. In each lottery, there is an inherent gamble between a high (x_H) and low (x_L) payoff with probabilities associated with each payoff: p_H ; p_L. Under the Expected Utility Theory (EUT) framework, subjects have some form of utility for each payoff value, $u(x)$. The expected utility that an individual derives from a lottery, $E[u]$, depends on both the lottery's probabilities as well as the utility that the individual attains from payoffs. Expected utility is represented by:

$$E[u] = p_H \cdot u(x_H) + p_L \cdot u(x_L)$$

Generally, experimental economists have designed tasks such that lotteries in which expected utility is higher also tend to have larger difference between high and low payoff values; lotteries that have a smaller difference between high and low payoff values have smaller expected utility. Subjects who are averse to risk will opt for a lottery in which difference between high/low values are modest, while those who are less risk averse opt for lotteries in which they can receive either a very large payout, or a very small payout, depending on the probabilities associated with these payoffs.

In a lottery experiment, the experimenter presents subjects with a choice of lotteries, from which subjects are tasked to choose one for play. This is often referred to as a "game". Games are played over real dollar values. The final payoff of the game is determined through a randomized selection of outcomes (e.g. dice, bingo balls, or a computerized random number generator), hence referred as a "lottery". If multiple choices were elicited during the lottery choice task, one of the choices is typically randomly selected for payment of experimental subjects. Randomization will determine payout outcome in the selected lottery, based on the probabilities associated with high/low payoffs.

3.2.1.2. Multiple Price List

Holt and Laury (2002) formulate a lottery choice game with a list of 10 paired lotteries. Lotteries on each line of the list are represented as *Option A* and *Option B*. *Option A* is a "safe" lottery in which the difference between its high/low payoffs is smaller than that of the "risky" *Option B*. Down the list, the high/low payoff values of *Option A* and *Option B* remain constant, while probabilities for the high (low) payoff to be selected will increase (decrease) such that an expected utility maximizer is induced to switch from *Option A* to *Option B*. Switching further down the line is indicative of higher risk aversion.

3.2.1.3. Single Choice List

In contrast to the Multiple Price List (MPL), in the Eckel and Grossman (2002, 2008) Single Choice List (SCL) asks subjects to choose one lottery among a set of five. In this elicitation method, the probability associated with high/low payoffs of each lottery remains constant at 50%. While the most risk averse subject will choose the first lottery in which high/low payoffs are the same, thus guaranteeing a payoff with certainty, risk neutral (as well as risk seeking) subjects will take a 50/50 gamble between receiving a very high payoff and a low payoff of $0; this option also has the high expected payoff.

3.2.1.4. Certainty Equivalent Method

Abdellaoui et al. (2011) offer a variation of the lottery method by which subjects choose between playing a lottery or receiving a sure payment. In the original iteration, the Certainty Equivalent Method (CEM) has nine choices, i.e. one per row of a list of nine lines. *Option A* is a 50/50 lottery between a high and low payoff; probabilities and payoffs remain the same throughout the list. *Option B* is a sure payment that increases marginally down the line, from the value of low payoff in the first line, to the highest payoff in the last. Note that, in this setup, choices in lines 1 and 9 are stochastically dominated such that a rational player is always better off choosing the lottery option in Line 1 and sure payment in Line 9. Like in the MPL method, the switching point indicates the degree of risk aversion. Risk averse subjects generally prefer sure payment options in earlier lines, whereas risk-neutral and risk-seeking individuals will select more lottery options, up to the stochastically dominant sure payment of Line 9.

3.2.1.5. Investment Game

In the Investment Game developed by Gneezy and Potters (1997), the subject is given \$X and tasked to allocate this endowment between a risky project (\$x) and a riskless pot ($\$X - x$). If the project succeeds with probability p, the subject receives $\$(X - x + kx)$; if the project fails $(1 - p)$, she keeps the amount that had been set aside $\$(X - x)$. The specifications are set up such that $(k > 1)$ and $p \cdot k > 1$. Thus, the most risk-averse subjects choose to invest little to none of their endowment, while those who are risk neutral (or risk seeking) will invest most to all of their investment.

3.2.1.6. Connection to economic theory: identifying risk-seeking individuals

The design of EMs will limit the range of implied risk parameters that may be elicited. For instance, of the four tasks selected in this project, the Single Choice List (SCL) and Investment Game (INV) measure varying degrees of risk aversion, but do not have available choices of gambles that would discern risk seeking subjects from risk neutral ones. In these two particular tasks, the most risk-neutral subjects (and also those who are risk-seeking) would be expected to select the gamble that has the highest variance and also the highest expected payoff. In contrast, the Multiple Price List (MPL) and Certainty Equivalent Method (CEM) extend into the risk seeking domain; when faced with a choice between a risky gamble and safe gamble, the risk-seeking subject is more likely to select the risky gamble (e.g. low probability of high payoff, high probability of low payoff), despite its probability-weighted expected payoff being smaller than that of the safe gamble.

3.2.2. Risk self-evaluation in psychology and economics

Outside of economics, researchers in psychology, neuroscience and cognitive science have developed measurements of risk aversion that suitable to their respective fields. For instance, psychologists favour self-evaluation for risk preference elicitation, often rated on a Likert scale. Self-evaluation scales differ in sensitivity, ranging from 5 points (Blais and Weber, 2006; Nicholson et al., 2005; Weber et al., 2002) to 10 points (SOEP et al., 2007). Self-assessment is well-regarded in psychology (Frey et al., 2017; Pedroni et al., 2017), and its validity is often tested in economic literature in conjunction with experimental EMs (Crosetto and Filippin, 2016; Deck et al., 2013; Reynaud and Couture, 2012). In cognitive neuroscience, researchers will generally apply techniques like functional magnetic resonance imaging (fMRI) technology to measure the biophysical indicators of risk-seeking behaviour in the brain (Schonberg et al., 2011). The following is a review of the most prevalent non-experimental measures outside of economics, including self-evaluated risk-taking on a Likert scale (General risk-taking, Domain-Specific risk-taking) and biophysical measures of risk-taking.

3.2.2.1. Self-evaluation: General risk-taking

At its most elemental level, a self-evaluated risk method consists of a single question in which the subject is tasked with evaluating their self-perceived risk on a Likert scale. This is a method employed by the German Social Economic Panel (SOEP) questionnaire, a longitudinal survey of over 15,000 private households (SOEP, Wagner et al. 2007). Dohmen et al. (2011) also replicated this method; their results lead them to support the elicitation of a general risk attitude, contending that this approach mitigates subjective beliefs about the riskiness of a decision environment, which could bias stated risk preferences.

3.2.2.2. Self-evaluation: Domain-Specific Risk-Taking

Domain-specific stated risk preference methods ask subjects to rate their self-perceived risk preferences in different contexts. The Domain-Specific Risk-Taking (DOSPERT) Scale in Weber et al. (2002) asks subjects to state their outlook on 50 scenarios, made up of 10 scenarios in each of five domains: financial decisions, health/safety, recreational, ethical, and social decisions. These scenarios ranged from “Forging somebody’s signature.” (Ethical) to “Investing 5% of your annual income in a conservative stock.” (Financial) to “Engaging in unprotected sex.” (Health). Subjects rate their likelihood of engaging in these activities, as well as their perception of the risk inherent in these scenarios, and expected benefits of these activities. A regression analysis is seeking the relationship between expected benefits and perceived risk on the likelihood risky activity participation. Findings suggest gender and domain differences in stated risk taking are associated with perception of risky activities’ benefits and risk. Blais and Weber later revised this questionnaire, condensing it to 30 questions across the same five domains (2006).

Nicholson et al. (2005) propose a domain-specific risk assessment tool addressing everyday risk-taking in six domains: recreational, health, career, financial, safety, social. In contrast to WBB’s large selection of questions, Nicholson et al. method is concise: for each one of the six domains, subjects indicate, on a 5-point scale, their self-perceived risk tolerance for the domain, currently, as well as in their adult past. They discover that risk propensity has relationships with age and gender, career decisions.

3.2.3. Measuring risk preference consistency across economic elicitation methods

Since the foundational publications of Holt and Laury (2002) Multiple Price List method (MPL) and the Eckel and Grossman (2008) Single Choice List method (SCL) experimental economists have been interested in in developing new risk elicitation methods (Abdellaoui et al., 2011; Crosetto and Filippin, 2013; Figner et al., 2009; Lejuez et al., 2002; Pedroni et al., 2017). MPL and SCL have also been adapted for field experiments. For instance, Tanaka et al. (2010) extends Holt and Laury's MPL into a format suitable for Prospect Theory analysis, by designing multiple lists with varying probabilities and payoff values, some of which were in the domain of loss. As well, Reynaud and Couture (2012) adapts Eckel and Grossman's SCL for Prospect Theory analysis by adding additional gambles extending the choice set into the risk-loving domain.

Experimentalists have also been interested in evaluating the predictive reliability of EMs, individuals' risk preference consistency across multiple EMs, and for the presence of a "base-rate effect" inherent in EMs that would bias revealed preference estimates (Crosetto and Filippin, 2016; Friedman et al., 2018; Mata et al., 2018; Frey et al., 2017). While each EM has its unique theoretical and practical advantages (and disadvantages), the "risk elicitation puzzle", as termed by Pedroni et al. (2017), remains unsolved; researchers vacillate in supporting one method over the other, unable to declare any one elicitation method to be unequivocally superior.

Structural modelling: Expected Utility Theory & Prospect Theory

Lottery list risk elicitation researchers usually employ structural modelling approaches to analyze risk preferences. Often, models are examined under the classical Expected Utility Theory framework (EUT) framework (Von Neumann and Morgenstern, 1947), with the assumption of a Constant Relative Risk Aversion (CRRA) utility function (Dave et al., 2010; Eckel and Grossman, 2008; Holt and Laury, 2002; Holzmeister and Stefan, 2020). Other papers apply Prospect Theory

(Kahneman and Tversky, 1979; Tversky and Kahneman, 1992), in which researchers explore the possibility that subjects weigh probabilities unevenly. Prospect Theory (PT) can uncover behaviour that would not be captured by standard EUT (Pedroni et al., 2017; Reynaud and Couture, 2012; Tanaka et al., 2010). When applied in the field, PT has its advocates (Abdellaoui et al., 2013; Barberis, 2013; Pope and Schweitzer, 2011; Ruggeri et al., 2020), conditional supporters (Harrison et al., 2010; List, 2004) and its detractors (Kachelmeier and Shehata, 1992; Levy and Levy, 2021). In the lottery list EM literature, support for PT over EUT is also divided (Crosetto and Filippin, 2016; Holzmeister, 2016; Pedroni et al., 2017; Reynaud and Couture, 2012; Tanaka et al., 2010).

3.2.3.1. Expected Utility Theory

Dave et al. (2010) employ a structural estimation approach to test consistency of risk preference measurement between two EMs of different complexity: the complex Holt-Laury multiple price list and the relatively simpler Eckel-Grossman single choice list. Subjects are recruited from conventional pools and non-conventional pools (i.e. college students and non-college student adults respectively), and in addition to completing the two incentivized tasks, they provided demographic information and completed a numeracy test.

After coding task decision data into a series of binary choices between choice sets[17], Dave et al. apply the (Holt and Laury, 2002) Holt and Laury (2002) specification of a CRRA utility function, consisting of a risk parameter and a noise parameter to accommodate decision-making error. In the first set of models, MPL & SCL data are pooled to estimate parameters as functions of the set of demographic variables and a dummy variable for MPL. Estimates demonstrate that female and low math score subjects are more likely to make risk averse choices, and that, in

[17] This method will be explained in further detail in Section 3.4.1: Coding Choice Data, below.

comparison to the SCL task, MPL elicits more risk aversion with more noise. When this specification is applied separately to each task, females are significantly more likely to make risk averse choices in both MPL and SCL with less noise; subjects with low numeracy exhibit significantly more noise in the MPL task, signalling their miscomprehension.

Dave et al. observe that non-conventional subjects struggle with the MPL task, often requesting clarification and being more likely to make inconsistent choices (i.e. switching back and forth between the list of safe and risky choices). Subjects who made inconsistent MPL choices are also more likely to be in the low-numeracy cohort; when this cohort is omitted, the estimate of low math score subjects on the noise parameter lost significance. These results lead Dave et al. to suggest MPL be used for subjects with higher math skills, as this EM elicits risk preference with more nuance; SCL should be considered for non-conventional subjects and for subjects with low numeracy.

The between-task structural estimation method employed by Dave et al. (2010) is also used by Reynaud and Couture (2012) to compare MPL and SCL, though they find, contrarily, that SCL induces more risk aversion than MPL. Holzmeister and Stefan (2020) apply this structure to four tasks, task-by-task; their result on the MPL task effect corroborates the findings of Dave et al. (2010).

3.2.3.2. Predictive accuracy of CRRA estimates

Dave et al. (2010) further test the task effects of SCL and MPL, with a focus on the low numeracy cohort of subjects, by applying the *predictive accuracy* measure: the proportion of choices that can be predicted using CRRA parameters that had been previously estimated based on subjects' demographic characteristics. While MPL predicts 0.84 of choices for all subjects, there is a significantly lower degree of predictive accuracy for low-numeracy subjects (0.76) when compared to the non-low group (0.85); in contrast, the predictive accuracy of SCL (0.72) does not vary significantly between numeracy cohorts. These results reinforce the conclusion by Dave et al. that MPL is beneficial for in high numeracy subjects, but the simpler SCL is applicable to all subjects without needing to consider their numeracy.

3.2.3.3. Across-subject rank order

Pedroni et al. (2017) elicit risk behaviour through six incentivized EMs. After ordering subjects in each task based on counts of risky choices, the authors correlate rank orders between tasks to determine between-subject consistency. All 15 pairwise correlations between six tasks are positive, however, magnitudes and significances of these correlations vary, signalling there is minor/moderate consistency across EMs, across subjects.

3.2.3.4. Permutation statistics

To evaluate the distribution of risk preference consistency among subjects, Pedroni et al. (2017) first deem subjects as Risk Averse / Risk Seeking in each EM by comparing their decisions against a risk-neutral, expected utility-maximizing baseline. Subjects are categorized into six consistency classes on a histogram based on their revealed risk preferences across six EMs (e.g. 0 seeking / 6 averse, to, 5 seeking / 1 averse [18]). The authors apply permutation statistics to estimate base-rate distribution of consistency classes. Comparing the two histograms, Pedroni et al. find that there are significantly different distributions across four of the six consistency classes, leading them to conclude that consistency of risk-seeking / risk-averse behaviour is significantly influenced by the base-rate effect of EMs.

[18] One of the seven tasks in Pedroni et al. (2017) did not extend into the risk-seeking domain.

3.2.3.5. Prospect Theory

Pedroni et al. (2017) further examine three of the six methods by estimating their parameters using both EUT and PT frameworks. On average, 57% of subjects are best described by PT throughout the three tasks, however, a minority of subjects are best characterized throughout all three tasks by a single framework (17% by CPT). These findings suggest that subjects rarely anchor onto a single framework throughout multiple tasks. The authors surmise that EMs may not actually reveal a subject's innate risk preference, but rather, the subject actively constructs her preferences when interacting with the EM.

Reynaud and Couture (2012) also test for Prospect Theory behaviour in MPL by estimating a mixture model, following (Harrison and Rutström, 2009), to discover that their subject pool are evenly divided between being characterized by EUT and by PT.

3.2.3.6. Simulation with virtual subjects

Towards demonstrating the effect of EM mechanics on risk aversion, Crosetto and Filippin (2016) task participants with four methods: MPL, SCL, INV, and their own "Bomb" risk elicitation task (BRET) developed in (Crosetto and Filippin, 2013). In addition, to assess robustness, the authors ran a simulation exercise with 100,000 virtual subjects characterized by a CRRA function in which 10% of subjects set to exhibit random behaviour. The revealed parameters of actual experimental subjects' choices were plotted against those of virtual subjects on a cumulative density graph, in which distributions were found to be significantly different in certain sections. In particular, the authors note that real subjects are more risk averse in MPL and SCL and tend to avoid both risk averse and risk seeking extreme ends of SCL and BRET.

3.2.3.7. Comparing absolute scale of choices

Crosetto and Filippin (2016) also evaluate the gender effect, which has been a focal characteristic in the pioneering risk elicitation literature (Eckel and Grossman, 2008, 2002; Gneezy and Potters, 1997; Holt and Laury, 2002). While CRRA estimates on demographic characteristics confirm that females tend to be more risk averse in SCL and INV, the Wilcoxon *t*-test on male and female cohorts' mean absolute choices fails to confirm the presence of gender differences in MPL and BRET. Taken together with findings from the simulation exercise, these results lead Crosetto and Filippin to speculate that the presence of a riskless choice in SCL and INV will induce violation of the independence axiom of the Expected Utility Theory. The riskless option induces certainty effects, acts as a reference point against the uncertain outcomes, and increases the salience of regret (Loomes and Sugden, 1982).

3.2.3.8. Overlap of implied parameter intervals

Bruner (2009) tests for risk consistency between two variations of an MPL-style task, in which the author applies a structural EUT model, corresponding choices made in the two tasks with the implied CRRA parameter range. Risk preference is considered to be consistent when there is overlap between the intervals of the two tasks' implied parameters.

Holzmeister and Stefan (2020) employ this technique to determine within-subject risk preference consistency between four EMs: MPL, SCL, CEM and Crosetto and Filippin (2013) BRET. The authors critique previous literature for typically assessing across-method variation by the positive correlations of choices on the absolute scale. They contest that correlation of these absolute values is insufficient to evaluate preference consistency because, options in EMs may have vastly different implied parameter intervals. Instead, Holzmeister and Stefan create an individual-level consistency index that is the count of pairs in which the implied CRRA parameter intervals overlap (six pairs between four tasks). The authors acknowledge this measure serves only as a proxy of consistency, as overlapping intervals signal that risk aversion parameters are close, though not necessarily identical.

3.2.3.9. Risk perception and noisy behaviour

Holzmeister and Stefan (2020) also determine how perception of task difficulty influences choice behaviour. Immediately after completing each task, subjects are asked to rank their a) perceived riskiness of their decision and b) perceived complexity of the task, on a 7-point Likert scale. Results from maximum likelihood estimates of the CRRA model, including perception indices as variables, show that the parameterized risk parameter significantly relates to a), and also that the noise parameter varies with b). As evaluated perceptions relate to observed behaviour, Holzmeister and Stefan find that subjects are aware of riskiness and complexity of the task.

3.2.4. Comparing elicitation methods between economics and psychology

While risk behaviour is defined in economics literature by subjects' response to (often incentivized) elicitation methods, psychology clinicians are more likely to characterize "naturalistic risk-taking" as the propensity to engage in activities harmful to oneself (Schonberg et al., 2011). The gap between the two disciplines' definition of risk behaviour can be reconciled when subjects' revealed risk preferences are analyzed in conjunction with their self-evaluated risk levels[19].

[19] Nevertheless, this "reconciliation" itself is a bit overgeneralized. Kellen et al. (2017) critique empirical research into risk attitudes and self-reflection, due largely to "a *weak* definition of risk" as defined primarily by choice between lotteries.

3.2.4.1. Risk parameters and self-evaluated risk

In addition to evaluating MPL-SCL consistency with estimates of a CRRA model, Reynaud and Couture (2012) also employ the CRRA function to determine how self-evaluated risk preference correlates within revealed risk behaviour. Domain-specific risk behaviour is assessed using the revised DOSPERT (Blais and Weber, 2006), and these self-evaluation indices are correlated with revealed risk parameters. Correlations are significant with self-evaluated risk indices in Financial and Recreational domains, weakly significant with the Ethical domain and insignificant with Health and Social domains. Resultingly, Reynaud and Couture caution against interpreting subjects' lottery task revealed risk outside the relevant domains.

3.2.4.2. Self-assessed domain-specific and general risk

Dohmen et al. (2011) construct a single general risk question and validate this tool using domain-specific questions, self-evaluated risky behaviour (stock investment, sport activity, self-employment, smoking), and an incentivized lottery task and hypothetical lottery task. The authors discover that while self-evaluated risk in the financial domain is the best determinant for predicting hypothetical investment task, the response to a single general risk question is the best all-around determinant of risk.

3.2.4.3. Interval regression on self-evaluated general risk taking

Sauter et al. (2015) were the first researchers to compare elicited and self-evaluated risk preferences between two occupational groups (foresters and farmers). The authors task a variation of MPL (Laury et al., 2012) and the Reynaud and Couture (2012) specification of SCL, and task subjects to state self-evaluated risk using Dohmen et al. (2011).

Sauter et al. apply interval regressions to analyze profession and demographic effects on revealed CRRA parameters of MPL and SCL, finding that foresters exhibit more risk aversion than farmers in both tasks, while self-employed participants tend to be more risk-averse in the SCL. When the interval regression is applied on the self-evaluated risk preferences, all estimates were insignificant.

3.2.4.4. Scaled risk preferences

To further assess the elicited-self-evaluated risk preference relationship, Sauter et al. (2015) apply the Wilcoxon *t*-test to MPL and the self-evaluated risk index, both transformed to a 0 to 1 scale; results showing marginally significant and insignificant correlations for farmers and foresters, respectively. Contrary to the conclusion of Reynaud and Couture (2012) using self-evaluated domain-specific risk, Sauter et al. find that a self-evaluated risk index does not substitute for risk parameters elicited by MPL. Sauter et al. also speculate that their use of real monetary incentives may have induced an effect not captured by the hypothetical lotteries of Reynaud and Couture (2012).

3.2.4.5. Risk parameters and cognitive ability

Andersson et al. (2016) acknowledge research in which a negative correlation is determined between cognitive ability and risk aversion (Benjamin et al., 2013; Burks et al., 2009; Dohmen et al., 2010), and argue that this correlation could be spurious if choice inconsistency (e.g. multiple switching between risky/safe choices in a MPL) is not considered as part of behavioural noise. To prove this hypothesis, the authors present subjects with two MPL's with fixed probabilities and varying payoff values (Binswanger, 1980; Tanaka et al., 2010) as well as tests for cognitive ability and cognitive reflection (Frederick, 2005), while controlling for the Big Five personalities (Almlund et al., 2011). Correlation and OLS analyses demonstrate that subjects with higher levels of cognitive ability and reflection tend to make more risky choices in the first MPL task, but more safe choices in the second. However, in both MPL tasks, subjects who make consistent decisions tend to have a higher score of cognitive ability. The contrasting results lead the authors to conclude that while cognitive ability does not relate to revealed risk preference, individuals with higher cognitive ability will make decisions with less noise. The results withstand a series of robustness checks, including exclusion of subjects who spend very little time on tasks, inconsistent decisionmakers, as well as focusing on individuals who have a unique interior switch point in both tasks.

3.2.5. A review of empirical evidence of heterogeneity in risk aversion

Table 3.1 summarizes the literature that provides empirical evidence of heterogeneity in risk aversion, as measured through experimental and non-experimental elicitation methods. The literature proves that subjects in cohorts with risky backgrounds (in terms of profession, recreation, or other lifestyle circumstances) tend to differ from the general public in their risk preferences. Professional traders (Haigh and List, 2005), amateur race car drivers (Riddel and Kolstoe, 2013) and volunteer firefighters (Krčál et al., 2019) all tend to exhibit less risk aversion in revealed risk or self-reported risk. Subjects that have recently faced trauma, such as hurricane evacuees (Eckel et al., 2009), also demonstrate less risk aversion. However, subjects with precarious livelihoods, like low-income high school students (Eckel et al., 2012) and sharecroppers (Dillon and Scandizzo, 1978), are more likely to be more risk averse compared against wealthier individuals that form the control group.

Table 3.1 Empirical evidence of heterogeneity in risk aversion

Citation	Elicitation method(s)	Control subjects	Subjects of interest	Risk aversion *(Subject of itnerest relative to Control)* *
		Incentivized task		
Kroll and Davidovitz (2003)	SCL	Children *(urban)*	Children *(commune)*	=
Eckel et al. (2009)	SCL	Female *(general population)*	Females *(recent hurricane evacuees)*	–
Eckel et al. (2012)	SCL	High school students	*Group 1*†*:* HS students in smaller-sized classes	+
			Group 2: HS students with more low-income peers	+ +
Sauter et al. (2015)	MPL & SCL	Farmers	Foresters	+
Haigh and List (2005)	INV	University Students	Professional traders	–
Krčál et al. (2019)	BRET	University Students	*Group 1:* Novice firefighters	–
			Group 2: Experienced firefighters	– –
		Non-incentivized task		
Dillon and Scandizzo (1978)	CEM (realistic context)	Small farm owners	Share-croppers	+
Riddel and Kolstoe (2013)	Health risk question	University students	*Group 1:* Amateur race car drivers	–
			Group 2: Elite rock climbers	–
			Group 3: SCUBA divers	=
Brown et al. (2011)	Investment risk question	Salary employed	Self-employed	–
Balaz and Williams (2011)	Willingness-to-pay for lottery play	Female university students *(domestic)*	Female university students *(international)*	–
Ayaita and Stürmer (2020)	Self-assessment	Civil servants	Teachers	+

† Some papers compare two or more groups of subjects of interest to a cohort of control subjects
* Subjects of interest demonstrate:
=: the same level of risk aversion as the Control group
+ (–): more (less) risk aversion than the Control
+ + (– –): even more (less) risk aversion than the first group of Subjects of Interest

3.3. Experiment design

Recruitment

The experiment recruited two groups of participants. First, a total of 124 subjects were recruited through the University of Alberta's Department of Resource Economics and Environmental Sociology's Online Recruitment System for Economic Experiments (ORSEE). This group of participants, comprised mainly of undergraduate students, represent the conventional demographic of control subjects used in the risk elicitation literature, and therefore are referred to as the control group. Next, 62 Incident Commanders were recruited through an internal Alberta Wildfire email distribution. Incident Commanders are wildland firefighters who are permanent staff members of Alberta Wildfire, and have received specialized training in leading suppression operations[20]. Subjects in both groups were informed that their voluntary participation in this experiment would assist research into risk behaviour, and that they would receive $10 for participation, plus the chance to earn additional money.

Selection of Elicitation Methods

This experiment design follows that of Holzmeister and Stefan (2020) in which subjects' choices across four prominent revealed risk elicitation methods are evaluated to determine within-subject consistency. Holzmeister and Stefan opt for the Bomb Risk Elicitation Task (BRET), Certainty Equivalent Method (CEM), Multiple Price List (MPL), Single Choice List (SCL). In our experiment we have replaced BRET with the Investment Game (INV), along with the addition of a self-evaluated domain-specific risk assessment (Nicholson et al., 2005) to analyze consistency between elicited and self-evaluated risk behaviour. As conducted in Holzmeister and Stefan

[20] Alberta Wildfire employs seasonal firefighters every fire season (April 1 to October 31), as well as contractor and emergency firefighters, on an as-needed basis.

(2020), during the experiment our four EMs are referenced as colours, so that participants' behaviours are not influenced by the name of the task[21]. Tasks are presented to our participants in a randomized order as to mitigate order effects. In order to reduce "portfolio effects" (Harrison et al., 2008), subjects are informed in advance that outcomes would be revealed at the end of the experiment, and that the final payoff will be the outcome of a single, randomly-selected task, which is not revealed until the experiment is over. Participants were informed that each task had an equal chance of being selected for payment.

To reduce the potential for participants to express risk aversion differently across four tasks due to differences in expected outcomes, tasks are specified such that a risk neutral participant has roughly the same expected outcome value in each task. Table 3.2 compares expected outcomes between participants who exhibit different risk behaviours. To obtain the values in Table 3.2, I first simulate the choices made by hypothetical players with different levels of risk aversion: risk neutral, most risk averse and most risk seeking[22]. To compare the expected outcomes of simulated players to real-life players, I aggregate and average the choices made by subjects in the original literature, available through the data in the papers' supplementary material (Abdellaoui et al., 2011; Eckel and Grossman, 2002; Gillen et al., 2019; Holt and Laury, 2002). The expected outcome in each task is the product of high/low payoff values in the hypothetical player's choice multiplied by the respective probabilities. As each task has an equal chance of being selected for payment, the final payoff is the average of the four tasks' expected outcomes.

[21] Refer to screenshots of the experiment in Appendix I: Instructions for the Risk Elicitation Economics Experiment
[22] The risk neutral player will maximize the expected outcome in each task; the most risk averse player opts for a lottery in which the difference between high/low payoff values are minimized (maximized). Parameterization of incentivized elicitation methods is explained in further detail in Section 3.3.1: Revealed risk elicitation methods

Table 3.2 Expected payoffs of four preference incentivized risk elicitation methods, based on literature vs simulated behaviour

	Per literature	If Risk Neutral	If Most Risk Averse	If Most Risk Seeking
Multiple Price List	$22.87[1]	$24.28	$20.05	$21.63
Single Choice List	$20.73[2]	$24.00	$16.00	$24.00
Certainty Equivalent Method	$22.23[3]	$22.78	$21.11	$21.11
Investment Game	$23.75[4]	$25.00	$20.00	$25.00
Final payoff †	**$22.40**	**$24.02**	**$19.29**	**$22.94**

1: Holt and Laury (2002); 2: Eckel and Grossman (2002); 3: Abdellaoui et al. (2011); 4: Gillen et al. (2019).
† The hypothetical final payoff is the average of all four task expected payoffs.

Online laboratory

The experiment is coded by the research team using the *oTree* framework (Chen et al., 2016), and is hosted on the Heroku cloud application platform; of the four incentivized elicitation methods in our experiment, MPL, SCL and CEM are adapted from packages developed by Holzmeister (2017) and INV has been developed in-house. Observing University of Alberta and provincial public health guidelines, the experiment took place entirely online from March to April 2021 during the COVID-19 pandemic. Incident Commanders are invited to participate by responding to an invitation email distributed through an internal Alberta Wildfire mailing list; prospective Control participants are recruited from the Department of Resource Economics and Environmental Sociology's Online REES Experiment Recruitment System (ORSEE).

A participant who expressed interest receives a unique hyperlink to the experiment platform, and is instructed to complete the experiment within 48 hours. During their sessions, participants are not monitored by the research team, however they are instructed to undertake the experiment individually, without external distractions, and to avoid disclosing details of the experiment to other prospective participants. Seven participants were dropped: one Incident Commander participant and six Control participants began the experiment and did not proceed past initial Informed Consent page within 48 hours. In total, 179 participants, consisting of 118 Control and 61 ICs, completed the experiment.

Enforcing transitivity in elicitation methods

Transitivity of preferences is a fundamental principle of classical economics (Regenwetter et al., 2011), assuming that an individual who prefers x to y and y to z must prefer x to z. In multiple lottery tasks MPL and CEM, transitivity may be violated when participants are given the option to switch back and forth between lists. Violation of transitivity may be attributed to participants' poor understanding of the task (Dave et al., 2010), but as this is not a focus of this research, we avoid this potential issue by asking participants to choose a "switching point" from the safe option to risky option (MPL: *Option A* to *Option B*) or vice versa (CEM: lottery option to sure payment). Enforcing a single switching point is consistent with previous adaptations of elicitation methods (Holzmeister and Stefan, 2020; Jacobson and Petrie, 2009; Tanaka et al., 2010).

Sections 3.3.1 to 3.3.3 summarize instructions given to participants in three parts of the experiment: Revealed risk elicitation methods, Self-assessed risk elicitation method, and the Demographic survey. Screenshots of the full experiment are available in Appendix I: Instructions for the Risk Elicitation Economics Experiment).

In the next section, I also outline how the range of choices made in each task reveal participants' direction of risk behaviour (from risk aversion, to risk neutrality, to risk seeking).

3.3.1. Revealed risk elicitation methods

3.3.1.1. Multiple Price List (Task "Orange")

This task is based on the original Holt and Laury (2002) specification, with payoff values multiplied by a factor of 10. Participants represented with a list of ten paired lotteries, named

Option A and *Option B*. *Option A* has high/low payoff values of $20.00 and $16.00; *Option B*, $38.50 and $1.00. Payoff values for each lottery stay the same throughout the ten lines, while the probabilities associated with the high (low) payoff value marginally increases (decreases) by 10% on each successive next line. The participant is tasked to consider which lottery she prefers on each line. However, as transitivity is enforced, selecting *Option B* on any line will force *Option B* to be selected in all preceding lines, so that the participant is in fact making a singular decision in selecting a desired switch point. One of the ten lines will be selected at random. A lottery will be played for the *Option* selected in order to determine the MPL task payoff.

On each line, one lottery has a higher expected outcome than the other. From lines 1 to 4, *Option A* has a larger expected outcome than *Option B*; from Line 5 to Line 10, *Option B* has a larger expected outcome. The expected outcome (EO) of *Option A* and *B* lotteries in Line 5 are:

$$\begin{aligned} EO(A) &= 0.5 \cdot \$20.00 + 0.5 \cdot \$16.00 = \$18.00 \\ < \; EO(B) &= 0.5 \cdot \$38.50 + 0.5 \cdot \$1.00 = \$19.75 \end{aligned}$$

The participant's switch point from *Option A* to *Option B* reveals her risk behaviour: Risk neutral players will stay with *Option A* from lines 1 to 4, switching at Line 5 when the expected outcome of *Option B* is larger; risk seeking players will tend to switch to *Option B* before Line 5 as they want the chance to attain high $38.50 payoff, even when the expected outcome of *Option B* is comparatively lower than that of *Option B*; and risk averse players will tend to switch to *Option B* at some point after Line 5[23].

[23] See *Task "Orange" Decision* in Appendix I: Instructions for the Risk Elicitation Economics Experiment. In Line 10, each lottery has a 100% probability delivering the high payoff Players should select *Option B*, which gives her $38.50 with certainty, over *Option A*'s $20.00, however, we observe some players who stick with *Option A* throughout all lines, and discuss this form of "inconsistency" in Appendix E. Inconsistent players in MPL task

3.3.1.2. Single Choice List (Task "Violet")

SCL follows the original Eckel and Grossman (2002) specification, in which participants must select one preferred lottery from a list of five. Each lottery has two possible outcomes (high/low payoffs), and there is a 50/50 probability for high/low payoff[24]. In this task, the probabilities stay the same throughout all five lotteries, but the value of high/low payoffs change (*Lottery No. 1*: \$16/\$16; *2*: \$24/\$12; *3*: \$32/\$8; *4*: \$40/\$4; *5*: \$48/\$0). The selected lottery will be played to determine the SCL task payoff.

As the gap between high/low payoffs widens from *Lottery No. 1* to *5*, lotteries become inherently riskier and the expected outcomes increase from \$16 in *Lottery No. 1* to \$24 in *Lottery No. 5*:

$$EO(1) = 0.5 \cdot \$16 \ + 0.5 \cdot \$16 = \$16$$
$$EO(5) = 0.5 \cdot \$48 \ + 0.5 \cdot \$0 = \$24$$

The most risk averse participant will choose *Lottery No. 1* that guarantees \$16, while less risk averse players will choose one of the other lotteries where high and low payoffs are slightly different. Risk neutral players will opt for *Lottery No. 5*, as its expected outcome is higher than those of the other lotteries[25].

3.3.1.3. Certainty Equivalent Method (Task "Blue")

CEM follows the original Abdellaoui et al. (2011) specification, in which participants are presented with nine lines of paired choices: playing a *50/50 Coin Toss* lottery or receiving a *Sure Payment*. If she selects *Coin Toss*, there is a 50% chance of receiving \$30, and an equal chance of

---

[24] 50/50 probabilities are characterized in SCL and CEM as a "virtual coin tosses"

[25] In SCL, risk seeking behaviour cannot be distinguished from risk neutral behaviour, as the riskiest decision, *Lottery No. 5*, also has the largest expected outcome.

receiving $10; if *Sure Payment* is selected, there is a 100% probability of receiving the value listed on the line. The value of the *Sure Payment* increases from Line 1 ($10.00) to Line 9 ($30.00) by consistent increments of $2.50.

Transitivity is enforced in this task, so participants are asked to choose a switch point from *Coin Toss* to *Sure Payment*. The choices in lines 1 and 9 have be preselected, because *Coin Toss* and *Sure Payment* are, respectively, stochastically dominated choices[26]. One line will be randomly selected, and, depending on the choice made on this line, the CEM task payoff will be determined by a virtual "coin toss" or the *Sure Payment*.

A risk neutral player is expected to switch from lottery to sure payment at Line 5 or 6, in which the Expected Outcome of the sure payment is respectively equal to or greater than the EO of lottery option. A risk averse player will opt for *Sure Payment* earlier, even when $EO(Coin\ Toss) > EO(Sure\ Payment)$. A risk seeker will switch only in lines *7* to *9*.

3.3.1.4. Investment Game (Task "Green")

In this iteration of the Gneezy and Potters (1997) investment decision task, the participant is endowed with $20 and must choose the exact amount they wish to invest, in units of $0.01, into a *risky project*. We apply the Dreber et al. (2010) specification, in which the outcome of "success" returns to the participant 2.5 times the invested portion, and "failure" means she loses the invested portion; the participant always keeps the amount not invested, with certainty. "Success" and "failure" outcomes have an equal chance of success and failure.

[26] See *Task "Blue" Decision* in Appendix I: Instructions for the Risk Elicitation Economics Experiment

If the participant opts to invest nothing, she keeps \$20 with certainty. Investing the entire endowment of \$20 gives the participant a 50% chance of earning \$50 (2.5 times \$20), but also a 50% of ending the task with \$0. The expected outcome of investing the entire endowment is \$25:

$$EO(20) = 0.5 \cdot ((2.5 \times \$20) + \$0) + 0.5 \cdot (\$0 + \$0) = \$25$$

Investing half of the endowment yields and equal chance between \$35 and \$10 with an Expected Outcome of \$22.5:

$$EO(10) = 0.5 \cdot ((2.5 \times \$10) + \$10) + 0.5 \cdot (\$0 + \$10) = \$22.5$$

Risk averse subjects are expected to invest a small portion of the endowment, whereas the risk-neutral expected utility maximizer opts to invest everything because the expected outcome of investing the entire endowment is greater than the expected outcomes of investing anything less[27]:

$$EO(20) = 0.5 \times \$50 + 0.5 \times \$0 = \$25 > \ EO_i \ \forall \, i \ \neq 20$$

3.3.2. Self-assessed risk elicitation method

After having completed the four incentivized revealed risk tasks, participants are state their personal risk-taking across six risk domains (*Recreational, Health, Career, Financial, Safety, Social*). This survey follows Nicholson et al. (2005) in their original format and wording, which can be found in its entirety on *Survey: Self-Evaluation* in Appendix I: Instructions for the Risk Elicitation Economics Experiment

[27] In INV, risk seeking behaviour cannot be distinguished from risk neutral behaviour because the riskiest decision to invest the entire \$20 endowment also has the largest expected outcome.

3.3.3. Demographic survey

Towards the end of the experiment session, all participants are asked to voluntarily disclose their gender, age and education level. Additionally, we ask subjects if they have family dependents, to control for the effect of family variable on risk aversion (Chaulk et al., 2003; Jianakoplos and Bernasek, 2006). In addition to general demographic questions, Incident Commander participants also see career-specific questions that asked: the year in which they started their Alberta Wildfire service, Incident Commander certification level, their current role title, and the number of years spent in their current role. All questions provided a nondisclosure option. The text of this demographic survey can be found on *Survey* in Appendix I: Instructions for the Risk Elicitation Economics Experiment.

3.4. Structural Parameter Estimation

3.4.1. Coding Choice Data

To format data for structural estimation, we follow Dave et al. (2010) by converting choices in all revealed preference elicitation methods into a binary format. Paired lottery tasks like the Multiple Price List (MPL) and the Certainty Equivalent Method (CEM) lend themselves well to this format; for each line of the task, “0” represents selection of the “safe” choice, and “1”, the selection of the “risky” choice.

Consider MPL: *Option A* is the safe choice as the difference between \$20.00 and \$16.00 is small; *Option B* is the risky choice because of the large difference between \$38.50 and \$1.00. In MPL, the selection of *Option A* is coded “0” and selection of *Option B* is “1”. The choice of a risk neutral participant who opts for *Option A* from lines 1 to 4, and *Option B* from lines 5 to 10, will be coded as “0, 0, 0, 0, 1, 1, 1, 1, 1”.

Likewise, in formatting CEM choice data, the selection of the safe *Sure Payment* is coded "0" and selecting the risky *50/50 Coin Toss* is "1". The choice of a very risk seeking participant, who selects *Coin Toss* from Line 2 to 8 is represented as "1, 1, 1, 1, 1, 1, 1, 0"[28].

Additional transformation is required for Single Choice List (SCL) because the participant makes only one explicit choice. As suggested by Dave et al. (2010), we transform SCL into a format analogous to the MPL binary choice lottery, in which the five lottery options are transformed into four pairwise lotteries. In the transformed SCL, *Option B* is the "riskier" lottery because compared to *Option A*, it has a larger difference between high/low payoffs. This transformed binary format is represented as such:

Line	**Option A**	**Option B**
1	50% probability of $40 50% probability of $4	50% probability of $48 50% probability of $0
2	50% probability of $32 50% probability of $8	50% probability of $40 50% probability of $4
3	50% probability of $24 50% probability of $12	50% probability of $32 50% probability of $8
4	50% probability of $16 50% probability of $16	50% probability of $24 50% probability of $12

The participant's choice for one lottery is presumed to dominate her preference for the other four lotteries. For instance, the choice of the $40/$4 lottery is coded "0, 1, 1, 1"; in Line 1, the *Option A* dominates the *Option B*, and in Line 2, *Option B* dominates *Option A*, so implicitly the participant would also prefer risky *Option B* over safe *Option A* through lines 3 and 4. The choice of $16/$16 by the most risk averse player is coded "0, 0, 0, 0", and the selection of $48/$0 by the least risk averse player is coded "1, 1, 1, 1".

---

[28] There are 9 lines in the CEM task, but choice is represented by 8 binary codes because the stochastically dominant choices in Line 1 and Line 9 have been preselected. See 3.3.1.3 *Certainty Equivalent Method (Task "Blue")* for a detailed explanation of the stochastically dominant choices.

Choice data from the Investment Game (INV) requires further transformation because the possible choices are nearly continuous (i.e. choosing to allocate an amount into the risky project, ranging from \$0.00, \$0.01, \$0.02 … to \$19.98, \$19.99, \$20.00). To format the choice data from INV such that is comparable to the data from the three other tasks (as mentioned above, MPL: 10 lines; SCL: 4 lines; CEM: 9 lines), we first round investment decision to the nearest \$2.00, the unit applied in Crosetto and Filippin (2016). This way, there are ten lines of binary choices:

Line	Option A	Option B
1	50% probability of \$47 50% probability of \$2	50% probability of \$50 50% probability of \$0
2	50% probability of \$44 50% probability of \$4	50% probability of \$47 50% probability of \$2
3	50% probability of \$41 50% probability of \$6	50% probability of \$44 50% probability of \$4
…	…	…
9	50% probability of \$23 50% probability of \$18	50% probability of \$26 50% probability of \$16
10	50% probability of \$20 50% probability of \$20	50% probability of \$23 50% probability of \$18

Similar to the transformed binary format in SCL, in INV the participant's decision to invest a certain amount in the "risky project" is presumed to dominate all other possible levels of investment choice. For instance, if the participant allocates \$16 into the risky project (keeping \$4 regardless of the outcome of the risky project), the lottery of *Option A* is preferred over *B* in lines 1 and 2, and *Option B* dominates *A* from lines 3 to 10. This choice is coded as "0, 0, 1, 1, 1, 1, 1, 1, 1, 1". The choice of the most risk averse player to invest none of the \$20 in the risky project is coded "0, 0, 0, 0, 0, 0, 0, 0, 0, 0", the least risk averse player who invests \$20 in the risky project is "1, 1, 1, 1, 1, 1, 1, 1, 1, 1".

After this coding procedure, an individual participant's four experiment choices are represented as 32 rows of binary code (MPL: 10 rows; SCL: 4 rows; CEM: 8 rows; INV: 10 rows).

3.4.2. CRRA utility function

We employ the Constant Relative Risk Aversion (CRRA) utility function, which as been widely used in the economic risk aversion literature (Crosetto and Filippin, 2016; Dave et al., 2010; Holt and Laury, 2002; Holzmeister and Stefan, 2020; Pedroni et al., 2017).

$$u(x) = \begin{cases} \dfrac{x^{1-\varphi}}{(1-\varphi)} & \text{if } \varphi \neq 1 \\ \ln(x) & \text{if } \varphi = 1 \end{cases}$$

Subjects are assumed to have utility $u(x)$ for cash incentives, given their risk parameter (φ). In this specification, the utility form of a perfectly risk neutral individual $(\varphi = 0)$ simplifies to $u(x) = \mathrm{x}$; risk-seeking individuals are characterized by $\varphi \rightarrow -\infty$, and risk-averse by $\varphi \rightarrow +\infty$. The CRRA utility function forms the basis of the structural estimation approach via the conditional log-likelihood function, which can be imposed upon the observed choices of selecting one option over the other.

3.4.3. Structural estimation

Once the data has been appropriately formatted, we specify the CRRA function. Expected Utility Theory (EUT) assumes that the participant, having internalized the probabilities and payoff values of all choices in a set, will make a decision that gives the best level of expected utility (EU).

$$EU(p, x, \varphi) = p_H \cdot u(x_H) + p_L \cdot u(x_L)$$

EU is a function of probabilities (p), high and low payoff values (x_H, x_L) as well as the utility that the participant derives from payoffs, informed by her risk parameter (φ). According to EUT, most participants would prefer the lottery of *Option B* over *Option A* so long as $EU_B > EU_A$ (although risk-seekers enticed by the potential high payoff of *Option B* may opt for *Option B* over *Option A* even when $EU_B < EU_A$).

However, it is unreasonable to expect that empirical behavior will perfectly fit the economic model. In taking the model to the data, it is important to consider how decisionmakers' *deviation* from expected theory, due to behaviour or calculation mistakes, will be reflected through a noise parameter. Following Crosetto and Filippin (2016), Hey and Orme (1994b), Holzmeister and Stefan (2020), we assume the Fechner error specification for the noise parameter, in which the error of a group of participants' evaluation of *Option A* and *B* is assumed to take a normal distribution. A latent index is yielded:

$$\nabla EU = EU_B - EU_A + \sigma\varepsilon \quad \text{with } \varepsilon \sim N(0, 1)$$

which includes a stochastic error term ($\sigma\varepsilon$) that is interpreted as the "noise" in the participant's decision-making process (Wilcox, 2008). Sigma (σ), also referred to as the noise parameter, is the standard deviation of this noise; as σ approaches 0, the participant's observed choice becomes increasingly likely to be an expression of her underlying preference.

This index, informed by a participant's latent preferences (including risk, φ), is linked to observed choices in the binary format using a cumulative standard normal distribution $\Phi(\nabla EU)$. Implicitly, the participant's choice on each line can be described using a probit link function:

$$P(Choice_B > Choice_A) = \Phi\left(\frac{EU_B - EU_A}{\sigma}\right)$$

The likelihood function is maximized with respect to risk (φ) and noise (σ) parameters. We employ a conditional log-likelihood function following the Holzmeister and Stefan (2020) specification:

$$\ln L(\varphi, \sigma|\vec{y}) = \sum_{i=1}^{n}([\ln \Phi(\nabla EU_i)]^{y_i} + [\ln \Phi(-\nabla EU_i)]^{1-y_i})$$

including risk parameter (φ), standard deviation of noise (σ), and the vector of n choices made in each individual task $(\vec{y})$, as coded in the binary format (Section 3.4.1. Coding Choice Data). In each Line i, y is "0" for relatively safer *Option A*; "1" for a choice of *Option B*. Following Holzmeister and Stefan (2020), standard errors are clustered on the subject level.

3.4.4. Applying Prospect Theory

Our Expected Utility Theory (EUT) model assumes that participants will gauge probabilities objectively. This assumption is relaxed in Prospect Theory (PT) model, in which we estimate potential probability bias expressed by participants' choice behaviour. While all four EMs may be analyzed with the EUT framework, only the Multiple Price List lends itself to PT analysis, as MPL has variation in probabilities between lines of paired lotteries. It is possible that participants bias probabilities during the decision-making process (e.g. overweighting the small probability of 10%, or underweighting a large 90% probability). The magnitude of this probability bias is estimated with the Tversky and Kahneman (1992) probability weighting specification for evaluation of gains:

$$w(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{\frac{1}{\gamma}}}$$

Probability bias is indicated by a value of γ (gamma) that deviates from 1; when γ approaches 1, the value of the weighted probability is close to the true probability, $w(p) \cong p$. Tversky and Kahneman (1992) support this functional form for its simplicity (only one parameter: γ), its ability to encompass weighting functions in both concave and convex regions, and for its reliable approximation of aggregate and individual data within the range of $0.5 < p < 0.95$.

Incorporating the weighting function, the participant's valuation for the expected utility of a lottery is expressed as:

$$EU(p, \gamma, x, \varphi) = w(p_H) \cdot u(x_H) + w(p_L) \cdot u(x_L)$$

With the addition of γ parameter in this equation, the aforementioned conditional log-likelihood function is configured to estimate three parameters: risk (φ), noise (σ) and probability weighting (γ). also accounts for noise $\ln L(\varphi, y, \sigma|\vec{y})$. Prospect theory encapsulates the Expected Utility form, in which $\gamma = 1$, and the individual's probability weighting is equal to probability, $w(p) = p$. Figure 3.1 is a visualization of the relationship between probabilities and weighting probabilities when $\gamma = 1$, and at other hypothetical levels.

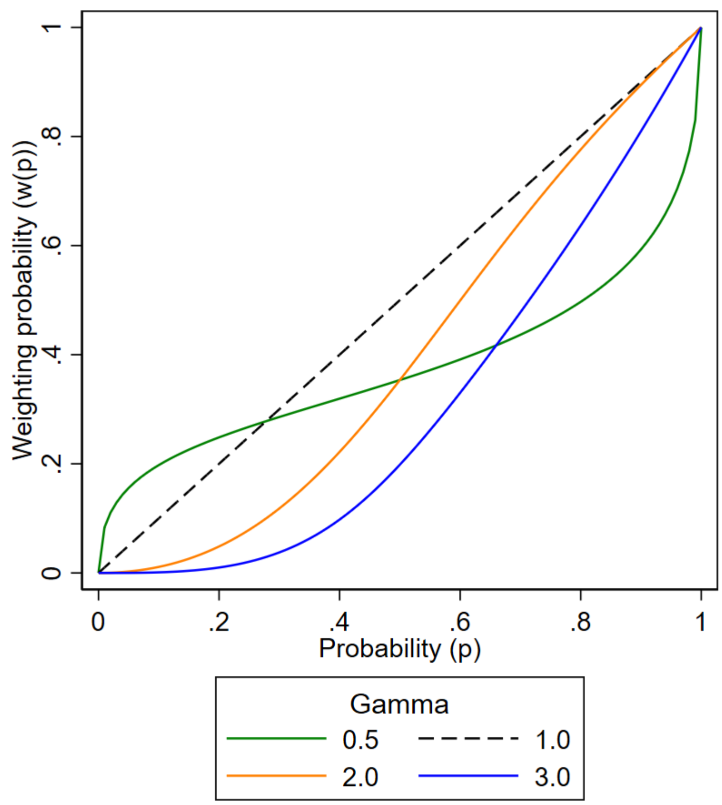

Figure 3.1 Visualization of probability weights at different levels of γ (gamma)

3.4.5. Accounting for heterogeneity between participants

Measures of revealed risk aversion in experimental games often correlate with characteristics such as gender (Eckel and Grossman, 2008; Filippin and Crosetto, 2016; Holt and Laury, 2002), age (Hanna and Lindamood, 2004), and education (Jung, 2015), as well as their vocation (see a summary of reviewed literature on heterogeneous risk preferences in Table 3.1 of Section 3.4.5. In order to account for observed heterogeneity between our 179 participants, risk and noise parameters are specified as linear functions of characteristic variables that are elicited from a voluntary survey at the end of the experiment (Incident commander, gender, university education, family, age).

Heterogeneity among Incident Commander participants is further investigated by matching the participant with their service records from Alberta Wildfire operations data. We investigate how risk aversion is affected by the number of days working in an operational capacity, including the number of days in which overtime was incurred, and days on an extended period beyond the standard 14-day deployment term.

Qualified Alberta Wildfire staff rotate through the weekly Duty Officer position, a role which carries heavy responsibility: The Forest Area Duty Officer is the first contact for all fires within a Forest Area. The Duty Officer is responsible for allocating the amount and type of initial suppression resources to be dispatched and advises their Wildfire Operations Officer on the import/export of resources to/from other jurisdictions. We measure the responsibility in fire operations by counting the number of days in which IC participants acted in capacity of Duty Officer.

3.5. Overview of empirical analysis

Through this experiment, we hope to explore two sets of research topics: the effect of Incident Commanders and their experiences on elicited risk aversion, and topics related to experiment design and analysis. The first set of research topics represent the focus of this study: to determine how the cohort of interest, Incident Commanders, differ from Control subjects, and amongst themselves, in expressing risk preferences. Through addressing the second set of topics, we come to a better understanding as to how the choice of analytical method affects interpretation of the results, and also examine the connection between elicited and self-evaluated risk preferences.

Analysis on Incident Commanders and Risk Aversion

Existing literature has demonstrated that subjects who are engaged in more risky professions are significantly less risk averse in their performance in elicitation methods, whether it be professional traders (Haigh and List, 2005), urban firefighters (Krčál et al., 2019), or race car drivers and rock climbers (Riddel and Kolstoe, 2013)[29].

The cohort of interest consists of 61 Alberta Wildfire firefighters with a wide range of experience in wildland firefighting operations. As Incident Commanders, these individuals play active roles in active wildfire suppression, from the direct supervision of wildfire crew members to high-level coordination of wildfire suppression programs. Firefighting is an occupational sector that is inherently risky, due to the nature of the work and its associated high rate of occupational disease (Alberta, 2018). Wildland firefighting, in particular, is perceived by the general public to be an extremely risky occupation (Desmond, 2009), a belief that is reinforced by news reports on wildland firefighters' strenuous work conditions (Dickson and Kulkarni, 2021), and high-profile

---

[29] For a review of literature on empirical evidence of heterogeneity in risk aversion, see Table 3.1 in Section 3.2.5.

fatalities (CTV News, 2021; Tucker, 2017). Given that wildland firefighters are immersed in a risky profession, it is possible that risk aversion levels differ between Control subjects and Incident Commanders.

Among the experimental participant cohort of Incident Commanders, individuals with more experience will have likely been exposed more to risk on the job. Risk elicitation literature demonstrates that individuals engaged in high-risk occupations, recreational activities, and those who have faced recent trauma have a tendency to be less risk averse in incentivized tasks (Eckel et al., 2009; Haigh and List, 2005; Krčál et al., 2019; Riddel and Kolstoe, 2013). Thus, we may find that Incident Commanders with more job experience are less risk averse. Job experience, in our definition, is measured as the number of days in which ICs were engaged in suppression operations, as well as the intensity at which ICs worked in the field (i.e. when and how often "Double Time" payment was earned).

To further investigate how job experience impacts risk aversion in the laboratory, we are interested whether the impact of job experience diminishes with time. Risk elicitation experiments with natural disaster survivors demonstrates that the impact of trauma in influencing risk behaviour tends to diminish as time passes (Cameron and Shah, 2015; Eckel et al., 2009). Inasmuch, it is possible to observe that the effect of the effect of Incident Commanders' past job experience, as defined above, on risk aversion diminishes with time. That is to say, the IC's job experience as measured from five years ago will have a smaller impact on her risk preferences, as compared to the experience of last year.

Analysis on Experiment Design and Analysis

Empirical research on risk elicitation methods have applied both Expected Utility Theory and Prospect Theory, and the research is inconclusive in determining which theory is superior in analyzing results elicited from these methods (see Section 3.2.3). Among our four tasks[30], MPL can be analyzed with both EUT and PT, and, due to the lack of consensus in the reviewed literature, it is possible to observe that Expected Utility Theory and Prospect Theory models provide comparable goodness-of-fit on risk aversion models.

Our elicitation methods vary in their format as well as in the range of risk parameters that they can elicit (MPL, CEM: risk-averse and -seeking behaviour; SCL, INV: only risk-averse behaviour). Due to these differences among elicitation methods, individual subjects often express different levels of risk, as measured by structural models (Holzmeister and Stefan, 2020; Pedroni et al., 2017). We may observe similar phenomena in our subject pool, by which Different elicitation methods will induce an individual to express different risk behaviours.

While we are primarily focused on risk measurement through elicitation methods, we are also interested in applying self-reporting methods. Given previous literature has documented the relationship between elicited and self-evaluated levels of risk (Dohmen et al., 2011; Lönnqvist et al., 2015), we may observe a significant positive correlation exists between elicited and self-evaluated measures of risk aversion. Regardless, a significant correlation between these two forms of risk measurement does not mean that experimentalists ought to forgo self-evaluated risk measures. Self-reporting can offer rich insights into an individual's risk profile, both in the present

[30] MPL: Multiple Price List; SCL: Single Choice List; CEM: Certainty Equivalent Method; INV: Investment Game

and the past, as well as differentiating how the individual perceives risk across various domains in life (Nicholson et al., 2005).

3.6. Experimental results

Table 3.3 reports the summary statistics for participants in Control and Incident Commander (IC) cohorts. The cohort of ICs are predominantly male (56/61). On average, our ICs are 42 years old; the youngest is 25 and the oldest, 60. With the exception of a single individual, all ICs have at least a polytechnic diploma or higher education level. 63% of ICs have family dependents. The Control group is made up largely of younger participants (mean age: 24) who have completed high school or a university undergrad program. 60% are female, and 87% majority of participants do not have family dependents.

In addition to their $10 participation payment, participants won further earnings of $0 to $50 based on the choices they made during elicitation tasks, the task that was randomly selected for payment, and the outcome determined by a random number generator. On average, participants completed the experiment with a total payoff of $31.81.

Table 3.3 Summary statistics for experiment participants

	Participant Cohort					
	Incident Commanders			Control		
	N	Mean	Std. Dev.	N	Mean	Std. Dev.
Survey Responses						
Female	61	0.08	0.28	118	0.60	0.49
Age	61	42.21	9.69	115	24.37	6.67
Family	59	0.63	0.49	116	0.13	0.34
University	60	0.32	0.47	115	0.57	0.50
ABWF experience (years)	61	19.46	9.89	.	.	.
IC Records						
No. Days Worked (2011-2020)						
Total	50	934.00	340.41	.	.	.
with Double Time	50	392.74	169.72	.	.	.
as Duty Officer	50	212.10	182.55	.	.	.
Extended period	50	69.88	67.03	.	.	.
Extended period w/ Double Time	50	41.54	35.68	.	.	.
No. Days Worked (2020)						
Total	61	73.52	42.50	.	.	.
with Double Time	61	23.74	19.53	.	.	.
as Duty Officer	61	15.98	18.77	.	.	.
Extended period	61	4.95	9.73	.	.	.
Extended period w/ Double Time	61	2.02	4.11	.	.	.

Double Time rate is paid on hours that extend beyond standard 7.25 hr on “Straight Time”, 2 hr on “Time and a half “ and 0.5 hr on “Straight Meal Time” during a day.
Extended days: the number of days beyond standard 14-day deployment.

3.6.1. Risk choice

A series of Wilcoxon rank-sum tests are applied to test the null of equality of distributions in risk averse task choices, between genders controlled for cohort (Table 3.4) and between cohorts controlled for gender (Table 3.5). Tests on the mean choices made by a cohort demonstrate that among Control participants, females are more risk averse through all tasks, and significantly more risk averse in INV and SCL *($p < 0.01$)*. This finding is consistent with Crosetto and Filippin (2017) and Filippin and Crosetto (2016), who suggest that it is the presence of a safe option in a task that induces risk averse behaviour in females. Among participants in the IC cohort, female risk aversion is significant in MPL *($p < 0.05$)*. Among male participants, risk aversion is not significantly different between Control and IC groups for any of the tasks. Among female participants, ICs are significantly more risk averse than Control only in MPL *($p < 0.05$)*.

Table 3.4 Mean choice across four tasks, compared between genders, controlled by cohort

Task	Control					Incident Commander				
	Male		Female		Wilcoxon p-value	Male		Female		Wilcoxon p-value
	N	Mean choice	N	Mean choice		N	Mean choice	N	Mean choice	
MPL	47	6.62	71	6.87	0.6381	56	6.89	5	9.20	0.0483
SCL	47	3.09	71	3.73	0.0058	56	3.23	5	3.80	0.4261
CEM	47	5.02	71	5.11	0.6019	56	4.61	5	4.40	0.8069
INV	47	8.31	71	10.85	0.0013	56	7.40	5	11.00	0.3741

MPL: Multiple Price List. SCL: Single Choice List. CEM: Certainty Equivalent Method. SCL: Single Choice List. Revealed risk aversion increases with the value of the choice.

Table 3.5 Mean choice across four tasks, compared between cohorts, controlled by gender

Task	Male					Female				
	Control		IC		Wilcoxon p-value	Control		IC		Wilcoxon p-value
	N	Mean choice	N	Mean choice		N	Mean choice	N	Mean choice	
MPL	47	6.62	56	6.89	0.7059	71	6.87	5	9.20	0.0273
SCL	47	3.09	56	3.23	0.5179	71	3.73	5	3.80	0.9470
CEM	47	5.02	56	4.61	0.1642	71	5.11	5	4.40	0.3795
INV	47	8.31	56	7.40	0.4278	71	10.85	5	11.00	0.9487

MPL: Multiple Price List. SCL: Single Choice List. CEM: Certainty Equivalent Method. SCL: Single Choice List. Revealed risk aversion increases with the value of the choice.

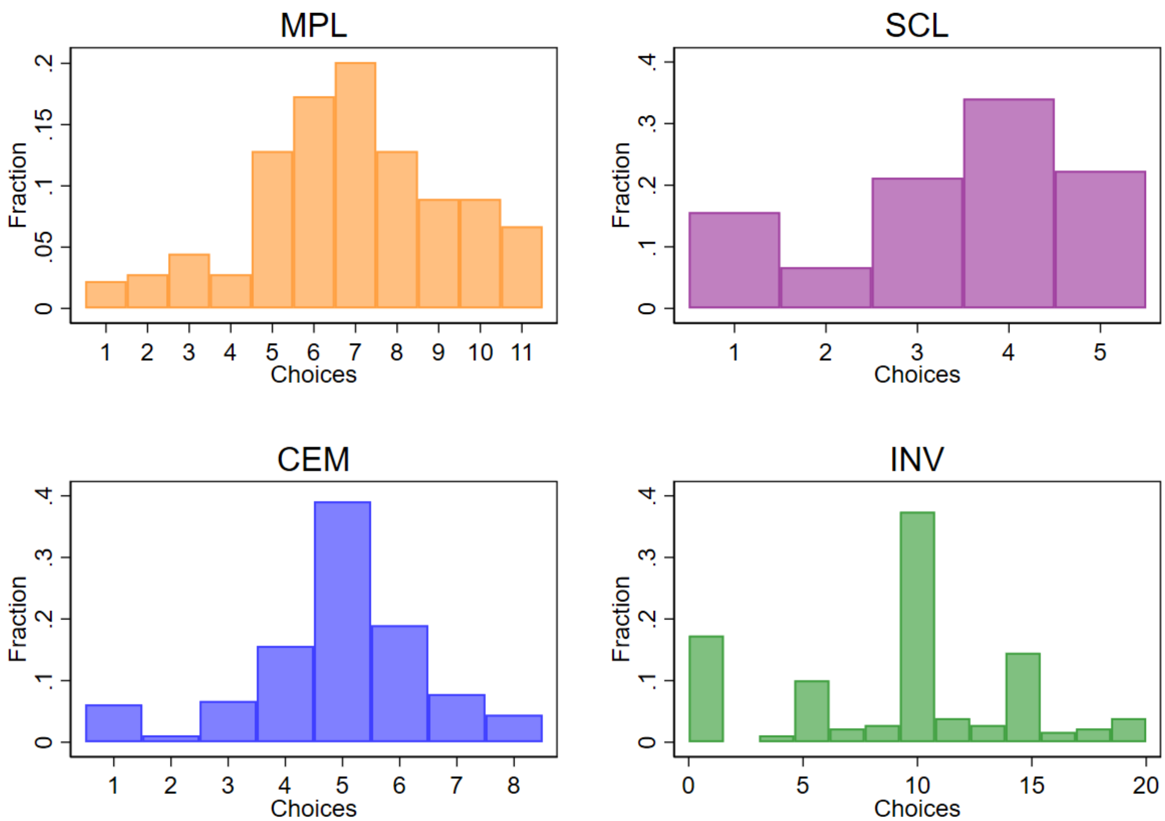

Figure 3.2 Participant choices in four elicitation methods
(Choices are in order of increasing risk aversion)

Figure 3.2 shows a distribution of participants' choices throughout four tasks. Notice the choice behaviour prevalent throughout four tasks; participants tend to centre their choices in MPL, CEM and INV, while skewing towards a higher level of risk aversion in SCL. Supplementary figures, in which the choice distribution is divided by gender and cohort can be found in Appendix D: Supplementary figures on experiment choices.

3.6.2. Structural models: Expected Utility Theory

We first analyze participants' choices using an Expected Utility maximum likelihood approach, which includes risk and noise parameters, as specified in Section 3.4.3. We estimate the set of model parameters (risk, φ; noise, σ) for each elicitation task.

In Table 3.6, risk and noise parameters (φ, σ) are estimated without accounting for heterogeneity across participants. In Table 3.7, the parameters are estimated on heterogeneous characteristics that participants reported in the demographic survey: Incident Commander, Female, University, Family (dummy variables) and Age (discrete variable). These models will be referred to, respectively, as homogeneous (Table 3.6) and heterogenous (Table 3.7).

Table 3.6 Estimates of risk and noise parameters (φ, σ) across tasks (homogeneous)

	MPL	SCL	CEM	INV
φ	0.519***	0.481***	0.468***	0.210***
	(0.049)	(0.034)	(0.119)	(0.011)
σ	2.237***	1.096***	0.992***	0.229***
	(0.293)	(0.097)	(0.363)	(0.014)
N	1790	716	1432	1790
N_clust	179	179	179	179

Robust standard errors clustered at the subject level in reported parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Table 3.6 reports that the average level of risk and noise are different across the four incentivized tasks. Participants express the highest level of both risk aversion and noise in MPL ($\hat{\varphi} = 0.519$; $\hat{\sigma} = 2.237$) and the lowest level of risk aversion and noise in INV ($\hat{\varphi} = 0.210$; $\hat{\sigma} = 0.229$).

Table 3.7 Estimates of risk and noise parameters (φ, σ) as functions of characteristics across tasks (heterogeneous on cohort, gender, education, family status and age)

	MPL	SCL	CEM	INV
φ				
IC	0.080	0.050	0.350**	0.024
	(0.190)	(0.093)	(0.151)	(0.039)
Female	0.122	0.224***	0.267***	0.068**
	(0.101)	(0.083)	(0.098)	(0.031)
University	0.163	0.184*	0.011	-0.010
	(0.111)	(0.099)	(0.153)	(0.025)
Family	0.026	-0.148**	-0.308	0.001
	(0.138)	(0.067)	(0.201)	(0.028)
Age	0.003	0.009	-0.008	-0.003**
	(0.009)	(0.007)	(0.009)	(0.001)
Constant	0.278	0.109	0.617**	0.273***
	(0.202)	(0.147)	(0.261)	(0.043)
σ				
IC	-1.086*	0.023	-0.989*	0.071
	(0.577)	(0.390)	(0.578)	(0.072)
Female	-0.764	0.344	-0.459**	-0.004
	(0.535)	(0.318)	(0.234)	(0.034)
University	-0.930	0.240	-0.135	0.053
	(0.644)	(0.422)	(0.255)	(0.048)
Family	0.604	-0.334	0.927	0.037
	(0.570)	(0.393)	(0.785)	(0.042)
Age	0.025	0.031	0.044	-0.001
	(0.033)	(0.025)	(0.032)	(0.002)
Constant	2.357***	0.033	-0.037	0.168***
	(0.795)	(0.507)	(0.643)	(0.043)
N	1710	684	1368	1710
N_clust	171	171	171	171

MPL: Multiple Price List. SCL: Single Choice List. CEM: Certainty Equivalent Method. SCL: Single Choice List.
Robust standard errors clustered at the subject level in reported parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Across the four tasks in the heterogeneous model (Table 3.7), Incident Commanders appear to be more risk averse than the control sample, and significantly so in CEM task ($p < 0.05$).

The noise parameter indicates how the participants' choice conforms to the predicted choices from EUT; a large noise parameter indicates that behaviour is farther removed from EUT[31]. We also observe marginal significance in ICs' being less noisy in MPL and CEM.

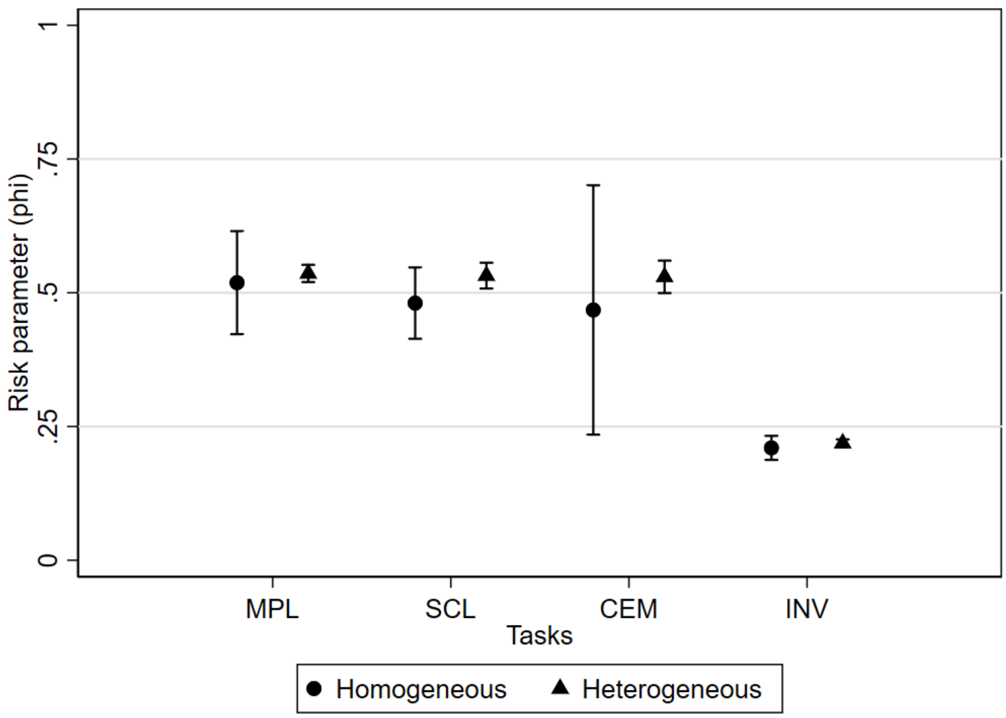

Figure 3.3 Estimates and confidence intervals (95%) of risk parameters in homogeneous and heterogeneous models by task

Figure 3.3 compares the estimated risk parameters and confidence intervals of the homogeneous model (Table 3.6) and those of the heterogeneous model (Table 3.7). In this figure,

[31] In their MPL & SCL experiment, Dave et al. (2010) discover females to exhibit significantly less noisy behaviour than male subjects in MPL, and no significant difference between genders in SCL. While females do tend to be less noisy in three tasks (MPL, CEM and INV), it is only in CEM that the gender effect on noise is significant, at the level of $p < 0.05$. Age and university degree attainment are not significant factors in shaping the noise parameter.

the plots of the heterogeneous model are the mean values of the $\hat{\varphi}_i$ per task, across all 179 individuals, based on the estimated risk parameter in Table 3.7. Confidence intervals are computed using the delta method. As observed in the figure, the confidence intervals on plots of the heterogeneous model are smaller than those on the homogeneous model, demonstrating that estimating risk and noise parameters on characteristic variables improves predictive precision in the structural estimation.

3.6.2.1. Comparing heterogeneous risk parameter estimates

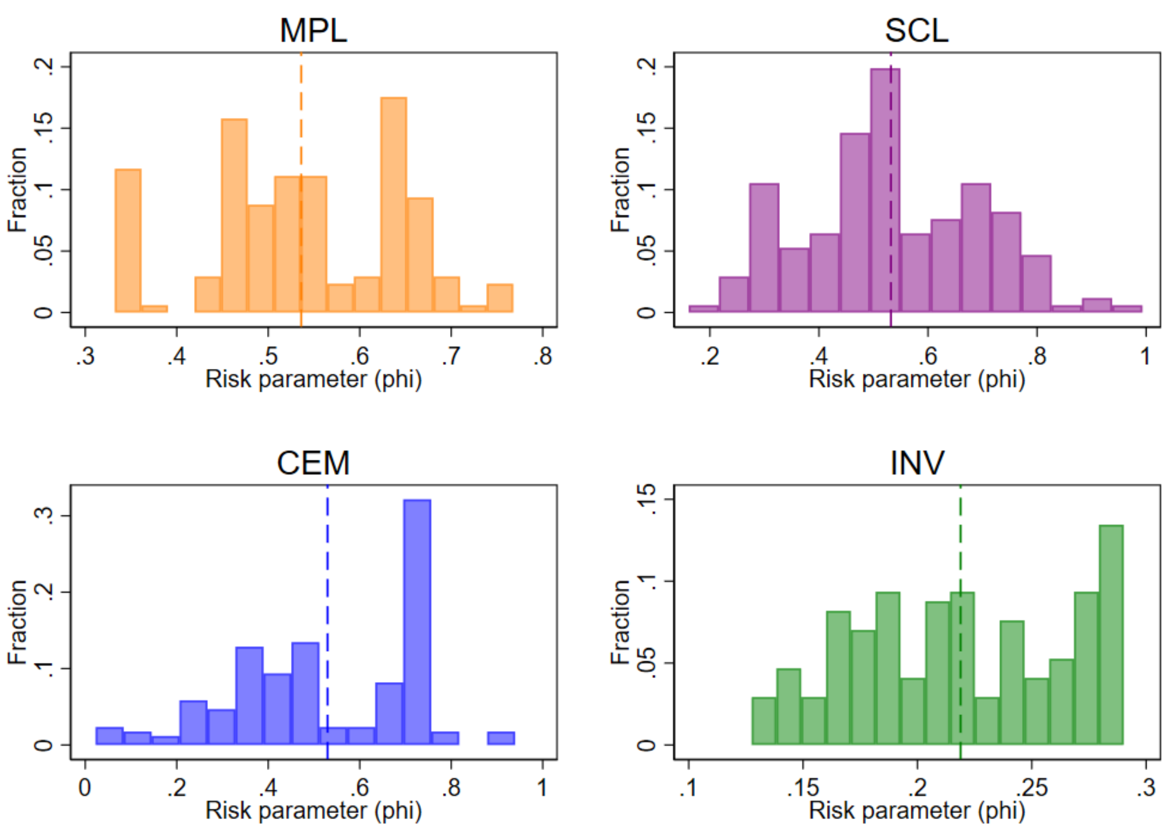

Figure 3.4 Distribution of estimated risk parameters ($\hat{\varphi}$) functioned on characteristics
(Dashed lines represent the mean $\hat{\varphi}$ across all participants)

Figure 3.4 shows the distributions of individual $\hat{\varphi}$ in each task, based on the estimates in Table 3.7. The distributions of estimated risk parameters appear to be centred in MPL and SCL, and slightly skewed to the right (higher levels of risk aversion) in CEM and INV.

3.6.2.2. Models with Incident Commander-specific traits

Towards discovering how Incident Commanders' job experience influence risk aversion, firstly we have linked IC participants with an Alberta Wildfire dataset that includes observations of daily, individual-level operational involvement. From this dataset, we are able to create variables that capture:

a) the number of total days in which the IC worked on wildfire suppression, and within those days,

b) the number of days in which Double Time was incurred (working over 9.75 hours),

c) the number of consecutive days extended beyond the standard 14-day deployment,

d) the number of consecutive days with Double Time extended beyond the standard 14-day deployment, and

e) the number of days the IC acted as Duty Officer, a role entailing higher responsibility.

Operational experience variables are in units of 100 days, so that results can be reported with three significant digits.

Firstly, we focus on observations of IC work experience in 2020, the year prior to the experiment. In order to achieve convergence in a model with a small sample size of 61 participants, the risk and noise parameters of the Expected Utility model are specified on a single explanatory variable at a time (Table 3.8 to Table 3.12). At the end of this sub-section, we will consider observations in previous years, to potentially discover a time-diminishing effect of work experience on risk aversion.

Table 3.8 Estimates of risk and noise parameters (φ, σ) as functions of Total days in operational deployment in 2020 (units of 100 days)

	MPL	SCL	CEM	INV
φ				
Total Days (100)	0.006	-0.249***	-0.272***	-0.059*
	(0.229)	(0.049)	(0.080)	(0.032)
Constant	0.587***	0.656***	0.226	0.222***
	(0.204)	(0.080)	(0.260)	(0.037)
σ				
Total Days (100)	-0.489	-0.690**	1.354	-0.096**
	(1.155)	(0.334)	(1.235)	(0.039)
Constant	2.445**	1.764***	3.283	0.315***
	(1.045)	(0.452)	(2.700)	(0.051)
N	610	244	488	610
N_clust	61	61	61	61

Robust standard errors clustered at the subject level in reported parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Table 3.9 Estimates of risk and noise parameters (φ, σ) as functions of Double Time in 2020 (units of 100 days)

	MPL	SCL	CEM	INV
φ				
Double Time (100 days)	-0.177	-0.476***	-0.607**	0.019
	(0.463)	(0.124)	(0.252)	(0.085)
Constant	0.634***	0.584***	0.239	0.169***
	(0.161)	(0.087)	(0.268)	(0.028)
σ				
Double Time (100 days)	-0.137	-1.286	2.326	-0.170*
	(2.380)	(1.034)	(2.727)	(0.092)
Constant	2.109***	1.513***	2.926	0.288***
	(0.730)	(0.436)	(2.457)	(0.041)
N	610	244	488	610
N_clust	61	61	61	61

“Double Time” is paid at twice the standard rate, and incurred after 7.25 hours of Standard time and 2 hours of Time-and-a-half
Robust standard errors clustered at the subject level in reported parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Table 3.8 show that Incident Commanders who spent more days on wildfire suppression operations in the previous year are significantly less risk averse in SCL and CEM (*$p < 0.01$* for both models), and are marginally significantly less risk averse in INV *($p < 0.10$)*. Similarly, Table 3.9 shows that ICs who had worked more days with Double Time in the previous year are also significantly less risk averse in both SCL *($p < 0.01$)* and CEM *($p < 0.05$)*.

Table 3.10 Estimates of risk and noise parameters (φ, σ) as functions of Extended period in 2020 (units of 100 days)

	MPL	SCL	CEM	INV
φ				
Extended period (100 days)	-0.586	-0.434***	-1.680	-0.115
	(0.591)	(0.140)	(1.552)	(0.086)
Constant	0.622***	0.474***	0.097	0.180***
	(0.110)	(0.075)	(0.270)	(0.019)
σ				
Extended period (100 days)	-0.090	-1.891	16.634	-0.277***
	(2.537)	(1.781)	(31.612)	(0.103)
Constant	2.064***	1.315***	3.766	0.258***
	(0.494)	(0.357)	(3.219)	(0.028)
ll	-266.213	-147.930	-194.921	-323.376
N	610	244	488	610
N_clust	61	61	61	61

The “extended period” are consecutive days beyond the standard 14-day deployment.
Robust standard errors clustered at the subject level in reported parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Table 3.11 Estimates of risk and noise parameters (φ, σ) as functions of Extended period with Double Time in 2020 (units of 100 days)

	MPL	SCL	CEM	INV
φ				
Extended period with Double Time (100 days)	-1.999	-1.600	-4.268**	-0.363
	(1.663)	(2.381)	(1.774)	(0.223)
Constant	0.631***	0.484***	0.135	0.183***
	(0.106)	(0.077)	(0.262)	(0.020)
σ				
Extended period with Double Time (100 days)	9.123	-2.216	56.248	-0.629**
	(16.250)	(19.291)	(67.807)	(0.249)
Constant	1.910***	1.189**	3.227	0.258***
	(0.440)	(0.537)	(2.633)	(0.030)
ll	-266.274	-146.590	-195.181	-323.719
N	610	244	488	610
N_clust	61	61	61	61

The “extended period” are consecutive days beyond the standard 14-day deployment.
“Double Time” is paid at twice the standard rate, and incurred after 7.25 hours of Standard time and 2 hours of Time-and-a-half
Robust standard errors clustered at the subject level in reported parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Table 3.12 Estimates of risk and noise parameters (φ, σ) as functions of Days as Duty Officer in 2020 (units of 100 days)

	MPL	SCL	CEM	INV
φ				
Duty Officer (100 days)	0.347	-0.263	-0.479**	-0.119
	(0.395)	(1.686)	(0.239)	(0.081)
Constant	0.536***	0.506	0.181	0.195***
	(0.119)	(0.521)	(0.272)	(0.024)
σ				
Duty Officer (100 days)	-2.003	-1.071	3.647	-0.006
	(1.517)	(6.481)	(3.940)	(0.133)
Constant	2.413***	1.433	2.923	0.242***
	(0.607)	(2.008)	(2.441)	(0.036)
N	610	244	488	610
N_clust	61	61	61	61

Robust standard errors clustered at the subject level in reported parentheses.
* p<0.10, ** p<0.05, *** p<0.01

The point estimates of the risk parameter in Table 3.10 are negative in all models which indicates that ICs who spent more days in extended deployment are generally less risk averse. However, the estimate of the parameter φ is only statistically significantly in SCL *($p < 0.01$)*. As well, additional extended deployment days in which the IC incurred Double Time are also associated with lower risk aversion across all tasks, significantly in CEM *($p < 0.05$)*, as shown in Table 3.11. In Table 3.12, we see that when Incident Commanders spent more days in 2020 taking up the responsibility of the Duty Officer, they also tend to be less risk averse across SCL, CEM, INV, significantly for CEM *($p < 0.05$)*.

To explore the possibility that the effect of Incident Commanders' past experiences on risk aversion diminishes with time, we explore how the measures of IC experiences across different time periods affect revealed risk aversion. Results for each measure of IC experience (Total days, Double Time, Extended period, Extended period with Double Time, and Duty Officer) are reported in Table 3.13 to Table 3.17.

Table 3.13 Estimates of risk parameters (φ) as functions of Total days in operational deployment from (year) to 2020 (units of 100 days)

	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
MPL	-0.008	-0.014	-0.010	-0.012	-0.020	-0.035	-0.046	-0.033	-0.039	0.006
	(0.024)	(0.025)	(0.030)	(0.032)	(0.033)	(0.036)	(0.046)	(0.075)	(0.101)	(0.229)
ll	-217.576	-241.040	-258.345	-262.293	-267.624	-266.661	-266.392	-267.370	-267.147	-267.362
R2	0.1892	0.1017	0.0372	0.0225	0.0027	0.0062	0.0072	0.0036	0.0044	0.0036
N	500	550	590	600	610	610	610	610	610	610
SCL	-0.029***	-0.033***	-0.036***	-0.040***	-0.047***	-0.058***	-0.078***	-0.099***	-0.146***	-0.249***
	(0.007)	(0.007)	(0.008)	(0.010)	(0.008)	(0.010)	(0.013)	(0.018)	(0.025)	(0.049)
ll	-122.156	-133.093	-141.267	-143.239	-145.862	-145.119	-144.419	-146.109	-146.168	-146.542
R2	0.1848	0.1118	0.0572	0.0441	0.0266	0.0315	0.0362	0.0249	0.0245	0.0220
N	200	220	236	240	244	244	244	244	244	244
CEM	-0.042***	-0.053***	-0.064***	-0.074***	-0.079***	-0.094***	-0.111**	-0.133**	-0.151***	-0.272***
	(0.012)	(0.016)	(0.020)	(0.024)	(0.025)	(0.033)	(0.044)	(0.054)	(0.039)	(0.080)
ll	-163.065	-177.375	-188.157	-189.952	-191.842	-191.674	-191.547	-192.034	-191.549	-192.383
R2	0.1752	0.1028	0.0483	0.0392	0.0297	0.0305	0.0312	0.0287	0.0312	0.0269
N	400	440	472	480	488	488	488	488	488	488
INV	-0.004	-0.007	-0.007	-0.008	-0.008	-0.010	-0.015	-0.024*	-0.035*	-0.059*
	(0.005)	(0.005)	(0.006)	(0.006)	(0.007)	(0.008)	(0.010)	(0.013)	(0.019)	(0.032)
ll	-269.929	-293.090	-315.512	-318.476	-324.592	-323.387	-321.567	-320.073	-318.658	-319.446
R2	0.1783	0.1078	0.0395	0.0305	0.0119	0.0155	0.0211	0.0256	0.0299	0.0275
N	500	550	590	600	610	610	610	610	610	610
N_clust	50	55	59	60	61	61	61	61	61	61

Robust standard errors clustered at the subject level in reported parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Table 3.14 Estimates of risk parameters (φ) as functions of Double Time from (year) to 2020 (units of 100 days)

	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
MPL	-0.014	-0.021	-0.021	-0.030	-0.044	-0.064	-0.089	-0.120	-0.199	-0.177
	(0.052)	(0.050)	(0.056)	(0.061)	(0.072)	(0.089)	(0.098)	(0.142)	(0.221)	(0.463)
ll	-217.408	-240.489	-257.345	-261.177	-266.534	-265.803	-265.377	-265.512	-265.167	-267.375
R2	0.1898	0.1038	0.0410	0.0267	0.0067	0.0094	0.0110	0.0105	0.0118	0.0036
N	500	550	590	600	610	610	610	610	610	610
SCL	-0.028	-0.037	-0.043	-0.049	-0.069	-0.102***	-0.130***	-0.193***	-0.341***	-0.476***
	(0.040)	(0.042)	(0.046)	(0.050)	(0.048)	(0.034)	(0.041)	(0.050)	(0.069)	(0.124)
ll	-124.040	-135.038	-143.495	-145.407	-148.510	-147.066	-146.435	-145.447	-144.969	-145.336
R2	0.1722	0.0988	0.0424	0.0296	0.0089	0.0185	0.0228	0.0293	0.0325	0.0301
N	200	220	236	240	244	244	244	244	244	244
CEM	-0.079**	-0.099***	-0.114***	-0.117**	-0.127***	-0.142**	-0.175***	-0.217***	-0.302***	-0.607**
	(0.033)	(0.036)	(0.043)	(0.047)	(0.049)	(0.061)	(0.066)	(0.075)	(0.106)	(0.252)
ll	-162.306	-177.159	-186.946	-189.122	-191.591	-192.315	-191.327	-191.215	-190.992	-192.583
R2	0.1791	0.1039	0.0544	0.0434	0.0309	0.0273	0.0323	0.0328	0.0340	0.0259
N	400	440	472	480	488	488	488	488	488	488
INV	0.010	0.005	0.006	0.003	0.003	-0.002	-0.007	-0.014	-0.027	0.019
	(0.016)	(0.014)	(0.017)	(0.015)	(0.016)	(0.018)	(0.021)	(0.029)	(0.045)	(0.085)
ll	-269.625	-296.047	-317.959	-321.471	-327.128	-326.392	-325.900	-325.644	-324.768	-326.365
R2	0.1792	0.0987	0.0320	0.0214	0.0041	0.0064	0.0079	0.0086	0.0113	0.0065
N	500	550	590	600	610	610	610	610	610	610
N_clust	50	55	59	60	61	61	61	61	61	61

Robust standard errors clustered at the subject level in reported parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Table 3.15 Estimates of risk and noise parameters (φ) as functions of Extended period from (Year) to 2020 (units of 100 days)

	2012	2013	2014	2015	2016	2017	2018	2019	2020
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
MPL	-0.022	0.039	-0.010	-0.101	-0.201	-0.192	-0.226	-0.278	-0.586
	(0.188)	(0.293)	(0.372)	(0.237)	(0.171)	(0.207)	(0.376)	(0.701)	(0.591)
ll	-241.159	-257.977	-262.287	-267.768	-266.581	-266.688	-267.275	-267.724	-266.213
R2	0.1013	0.0386	0.0225	0.0021	0.0065	0.0061	0.0040	0.0023	0.0079
N	550	590	600	610	610	610	610	610	610
SCL	-0.090***	-0.097***	-0.113***	-0.099***	-0.172***	-0.185***	-0.332***	-0.325***	-0.434***
	(0.022)	(0.021)	(0.000)	(0.022)	(0.054)	(0.000)	(0.083)	(0.105)	(0.140)
ll	-132.248	-141.025	-142.710	-147.159	-147.303	-145.995	-146.538	-148.254	-147.930
R2	0.1174	0.0589	0.0476	0.0179	0.0170	0.0257	0.0221	0.0106	0.0128
N	220	236	240	244	244	244	244	244	244
CEM	-0.215***	-0.217**	-0.259**	-0.271***	-0.326***	-0.320***	-0.470***	-0.760***	-1.680
	(0.081)	(0.086)	(0.105)	(0.104)	(0.120)	(0.116)	(0.164)	(0.263)	(1.552)
ll	-181.028	-192.042	-193.658	-195.439	-195.373	-195.815	-195.967	-195.022	-194.921
R2	0.0844	0.0287	0.0205	0.0115	0.0118	0.0096	0.0088	0.0136	0.0141
N	440	472	480	488	488	488	488	488	488
INV	-0.039	-0.035	-0.041	-0.041	-0.046***	-0.069*	-0.107**	-0.097	-0.115
	(0.033)	(0.024)	(0.033)	(0.036)	(0.010)	(0.039)	(0.042)	(0.068)	(0.086)
ll	-286.398	-308.905	-310.684	-316.004	-315.725	-314.574	-316.087	-320.688	-323.376
R2	0.1281	0.0596	0.0542	0.0380	0.0388	0.0423	0.0377	0.0237	0.0156
N	550	590	600	610	610	610	610	610	610
N_clust	55	59	60	61	61	61	61	61	61

Note: 2011 is not included, as the ML model does not converge for *Extended period from 2011 to 2020*.
Robust standard errors clustered at the subject level in reported parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Table 3.16 Estimates of risk and parameters (φ) as functions of Extended period with Double Time from (year) to 2020 (units of 100 days)

	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
MPL	-0.256	-0.219	-0.176	-0.210	-0.304	-0.429	-0.469	-0.723	-1.268	-1.999
	(0.198)	(0.221)	(0.279)	(0.290)	(0.283)	(0.265)	(0.321)	(0.509)	(0.818)	(1.663)
ll	-215.359	-239.520	-257.474	-261.519	-266.448	-265.350	-264.942	-265.422	-265.053	-266.274
R2	0.1974	0.1074	0.0405	0.0254	0.0070	0.0111	0.0126	0.0109	0.0122	0.0077
N	500	550	590	600	610	610	610	610	610	610
SCL	-0.200	-0.216	-0.237	-0.249	-0.217	-0.284	-0.489***	-0.617	-0.567**	-1.600
	(0.159)	(0.138)	(0.155)	(0.188)	(0.231)	(0.203)	(0.150)	(0.493)	(0.250)	(2.380)
ll	-123.337	-134.241	-142.888	-144.940	-149.024	-148.266	-146.217	-148.270	-148.589	-146.590
R2	0.1769	0.1041	0.0464	0.0327	0.0055	0.0105	0.0242	0.0105	0.0084	0.0217
N	200	220	236	240	244	244	244	244	244	244
CEM	-0.575***	-0.534***	-0.478***	-0.565***	-0.639***	-0.594***	-0.577***	-0.827***	-1.301**	-4.268**
	(0.161)	(0.154)	(0.161)	(0.195)	(0.205)	(0.195)	(0.197)	(0.297)	(0.512)	(1.774)
ll	-162.583	-178.964	-190.439	-192.488	-194.267	-195.173	-195.739	-195.659	-195.017	-195.181
R2	0.1777	0.0948	0.0368	0.0264	0.0174	0.0128	0.0100	0.0104	0.0136	0.0128
N	400	440	472	480	488	488	488	488	488	488
INV	-0.067*	-0.073*	-0.078**	-0.090**	-0.104**	-0.119**	-0.173***	-0.257**	-0.262	-0.363
	(0.036)	(0.037)	(0.033)	(0.039)	(0.043)	(0.050)	(0.065)	(0.117)	(0.195)	(0.223)
ll	-262.283	-287.065	-308.922	-311.355	-315.338	-312.835	-309.498	-314.039	-318.276	-323.719
R2	0.2015	0.1261	0.0596	0.0521	0.0400	0.0476	0.0578	0.0440	0.0311	0.0145
N	500	550	590	600	610	610	610	610	610	610
N_clust	50	55	59	60	61	61	61	61	61	61

Robust standard errors clustered at the subject level in reported parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Table 3.17 Estimates of risk and parameters (φ) as functions of Days as Duty Officer (2020) (units of 100 days)

	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se	b/se
MPL	-0.045	-0.032	-0.027	-0.014	0.008	0.024	0.041	0.122	0.206	0.347
	(0.048)	(0.051)	(0.055)	(0.064)	(0.073)	(0.091)	(0.114)	(0.150)	(0.155)	(0.395)
ll	-216.116	-240.997	-258.272	-262.621	-268.254	-267.869	-267.563	-265.863	-263.933	-265.799
R2	0.1946	0.1019	0.0375	0.0213	0.0003	0.0017	0.0029	0.0092	0.0164	0.0095
N	500	550	590	600	610	610	610	610	610	610
SCL	-0.030	-0.043**	-0.051**	-0.047	-0.040	-0.031	-0.030	-0.026	-0.030	-0.263
	(0.023)	(0.021)	(0.020)	(0.029)	(0.044)	(0.068)	(0.090)	(0.127)	(0.201)	(1.686)
ll	-122.064	-133.123	-142.648	-145.395	-149.442	-149.694	-149.733	-149.812	-149.821	-149.674
R2	0.1854	0.1116	0.0480	0.0297	0.0027	0.0010	0.0007	0.0002	0.0002	0.0011
N	200	220	236	240	244	244	244	244	244	244
CEM	-0.041	-0.058	-0.058	-0.061	-0.072	-0.099	-0.130	-0.220	-0.217*	-0.479**
	(0.032)	(0.038)	(0.050)	(0.066)	(0.074)	(0.096)	(0.131)	(0.187)	(0.115)	(0.239)
ll	-165.481	-181.076	-192.875	-194.984	-196.869	-196.721	-196.669	-196.405	-196.484	-196.453
R2	0.1630	0.0841	0.0244	0.0138	0.0042	0.0050	0.0053	0.0066	0.0062	0.0063
N	400	440	472	480	488	488	488	488	488	488
INV	-0.010	-0.012	-0.015	-0.016	-0.020	-0.021	-0.025	-0.027	-0.046	-0.119
	(0.009)	(0.010)	(0.011)	(0.012)	(0.014)	(0.018)	(0.023)	(0.031)	(0.043)	(0.081)
ll	-268.741	-293.662	-314.837	-318.611	-324.146	-325.463	-325.461	-326.761	-325.834	-323.895
R2	0.1819	0.1060	0.0415	0.0301	0.0132	0.0092	0.0092	0.0052	0.0081	0.0140
N	500	550	590	600	610	610	610	610	610	610
N_clust	50	55	59	60	61	61	61	61	61	61

Robust standard errors clustered at the subject level in reported parentheses.
* p<0.10, ** p<0.05, *** p<0.01

In Table 3.13 to Table 3.17, observations that make up IC experience variables are extended from 2020 to previous years. For instance, in each table, model "2011" includes observations from years 2011, 2012, 2013, …, up to and including 2020[32]. In the "2020" model, the results are the same as those reported in Table 3.8 to Table 3.12.

For most tasks in Table 3.13 to Table 3.17, we observe that the magnitude of the negative impact of work experience on risk aversion is reduced as observations are extended into the past. (The exception is for MPL, in which no IC experience variable has a significant effect.) Thus, recent operational experience, such as through additional operation days worked, extended periods of deployment, and more Duty Officer responsibilities, seems to be more impactful on reducing an IC's risk aversion during the experiment. However, we also observe that the goodness-of-fit (McFadden's pseudo-R2) improves with specifications that extend further into the past. These findings are discussed in further detail in Section 3.7. Discussion.

3.6.3. Structural models: Prospect Theory

As outlined in Section 3.4.4, our Multiple Price List is the only elicitation method that lends itself to Prospect Theory analysis, because the lotteries of this task have multiple probabilities. In addition to risk and noise parameters (φ, σ) that are present in EUT model, the PT model also includes a parameter that captures probability weighting (γ).

As observed in Table 3.18, below, in the structural model without heterogeneity, the risk parameter estimates of both EUT and PT are similar ($\hat{\varphi}_{EUT} = 0.519$; $\hat{\varphi}_{PT} = 0.521$) and both significant at $p < 0.01$. However, the estimated probability weighting estimator in the PT model

[32] The number of participant-level observations change from 2011 to 2015, as some ICs only joined Alberta Wildfire during these years.

($\hat{\gamma} = 2.239$) absorbs much of the variation captured in the noise parameter in EUT ($\hat{\sigma}_{EUT} = 2.37$; $\hat{\sigma}_{PT} = 1.014$).

Table 3.18 Estimates of risk, noise, and probability weighting parameters $(\varphi, \sigma, \gamma)$, MPL experiments using data from all participants

	EUT	PT
φ	0.519***	0.521***
	(0.049)	(0.053)
σ	2.237***	1.014***
	(0.293)	(0.103)
γ		2.239***
		(0.293)
N	1790	1790
N_clust	179	179

Robust standard errors clustered at the subject level in reported parentheses.
* p<0.10, ** p<0.05, *** p<0.01

Table 3.19 reports a heterogenous structural model in which risk, noise and probability weighting parameters $(\varphi, \sigma, \gamma)$ are estimated on participant characteristics. Results of the PT estimates are compared to estimates from the EUT structural estimation model.

Figure 3.5 visually compares the mean risk parameter estimates of the homogeneous/heterogeneous models estimated using EUT and PT.

Table 3.19 Estimates of risk, noise, and probability weighting parameters $(\varphi, \sigma, \gamma)$ on characteristics, MPL (all participants)

	EUT	PT
φ		
IC	0.080	-0.113
	(0.190)	(0.247)
Female	0.122	0.220**
	(0.101)	(0.106)
Age	0.003	0.023
	(0.009)	(0.016)
Family	0.026	-0.097
	(0.138)	(0.198)
University	0.163	0.221**
	(0.111)	(0.111)
Constant	0.278	-0.215
	(0.202)	(0.305)
σ		
IC	-1.086*	-0.344
	(0.577)	(0.227)
Female	-0.764	0.374*
	(0.535)	(0.207)
Age	0.025	0.054**
	(0.033)	(0.027)
Family	0.604	-0.388
	(0.570)	(0.289)
University	-0.930	0.266
	(0.644)	(0.232)
Constant	2.357***	-0.535
	(0.795)	(0.487)
γ		
IC		-0.650
		(0.781)
Female		-0.595*
		(0.309)
Age		-0.004
		(0.046)
Family		0.747
		(0.580)
University		-0.530
		(0.415)
Constant		2.309**
		(0.969)
Pseudo-R2	.0949	.1066
N	1710	1710
N_clust	171	171

Robust standard errors clustered at the subject level in reported parentheses.
* p<0.10, ** p<0.05, *** p<0.01

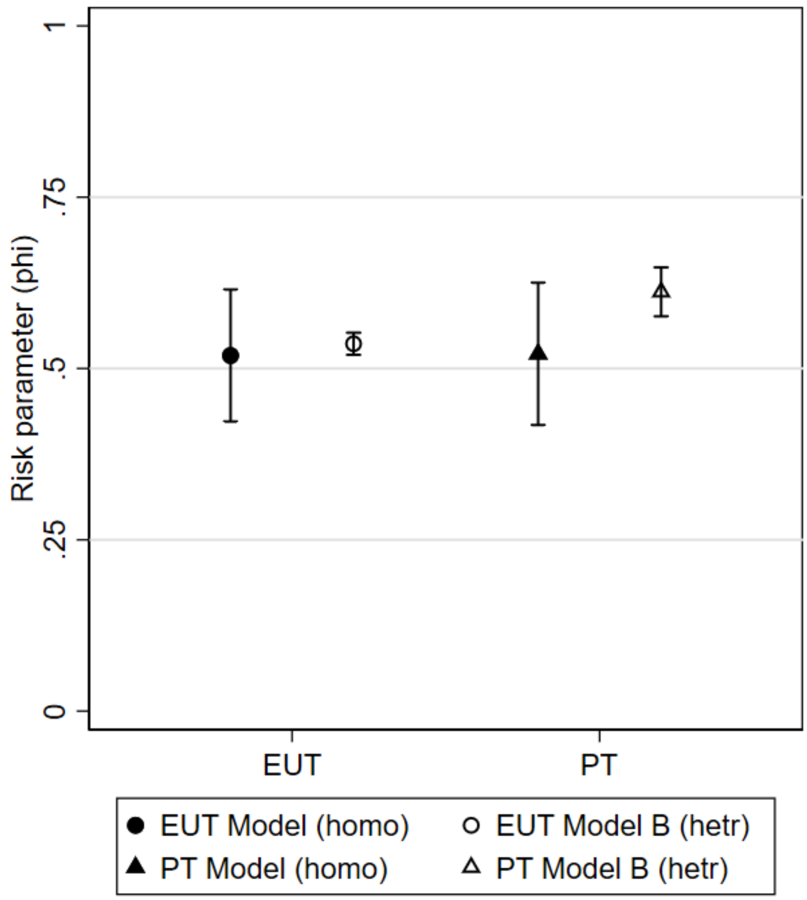

Figure 3.5 Estimates of risk parameters in EUT and PT models of MPL task, homogeneous and heterogeneous

In Table 3.19, the EUT structural estimation model on characteristics is compared to the PT model. When the probability weighting parameter (γ) of the PT model captures some of the observed variation previously captured by solely by risk or sigma parameters (φ, σ) in the EUT model we observe that: effects of Female and University on risk become significantly positive *($p < 0.05$)*, IC effect on noise is no longer significant, Female effect on noise becomes positive and marginally significant *($p < 0.10$)*, and Age effect on noise becomes significant *($p < 0.05$)*.

For the estimators for probability weighting in the PT model, only Female is marginally significant in its negative effect *($p < 0.10$)*.

Regarding differences between Incident Commanders and Control participants, ICs are not significantly different in risk aversion in either model; ICs are marginally significant in exhibiting less behavioural noise in EUT, however, with the introduction of the γ parameter, the IC estimator on the noise parameter is no longer significant, nor is the IC estimator significant on the probability weighting parameter.

Finally, to compare the goodness-of-fit of Expected Utility Theory and Prospect Theory in this model specification, we compare calculated McFadden's pseudo R-squared [33] . PT (0.1066) appears to be slightly better than EUT (0.0949) in capturing variation in the data.

Nevertheless, the difference in mean risk parameter estimates between homogeneous EUT and PT models are minute, while the mean risk estimated in heterogeneous PT model is slightly higher than that in the heterogeneous EUT model (Figure 3.5).

3.6.3.1. Counterfactual analysis of probability weighting parameters

Using estimated parameters from the PT model Table 3.19, we find that the median player has a probability weighting parameter equal to ($\gamma = 1.6$). Against an expected utility maximizer ($\gamma = 1.0$), the relationship between probability weighting against probability values for median player is displayed in Figure 3.6.

---

[33] McFadden's pseudo R-squared: $R^2 = 1 - \frac{\ln \hat{L}(M_{heteregeneous})}{\ln \hat{L}(M_{homogeneous})}$ in which $\ln \hat{L}$ is the estimated log likelihood and M is the model specification (McFadden, 1974).

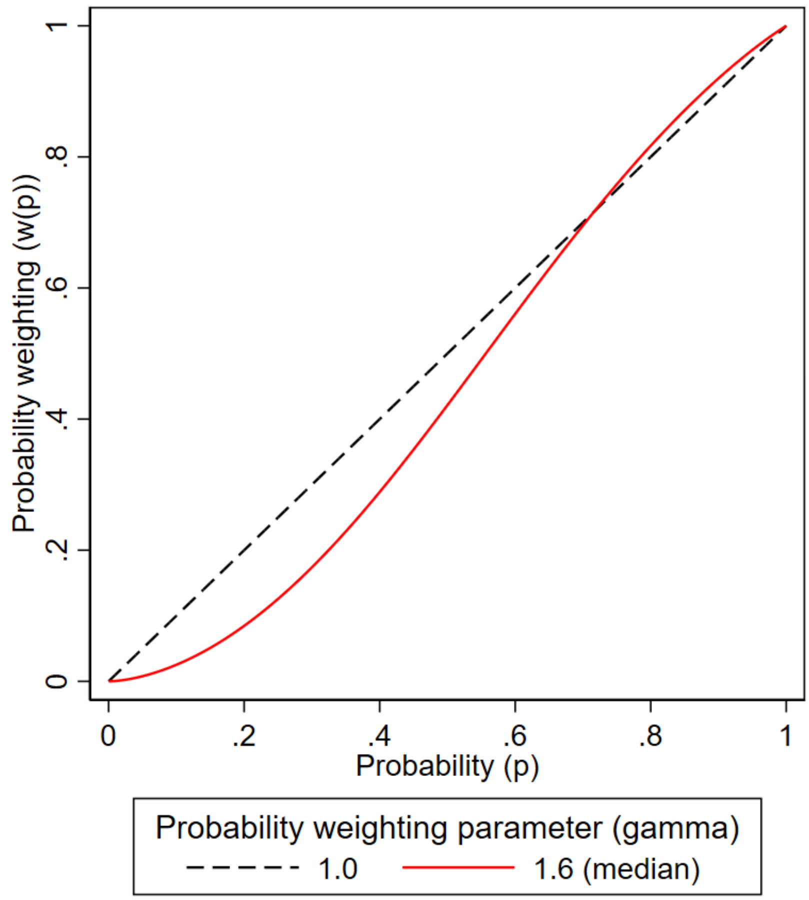

Figure 3.6 Probability weighting (γ) and Probability (p) for the median player in MPL

PT parameters from Table 3.19 can also be used to predict the characteristics of a counterfactual participant. Per the probability weighting estimators, such are the traits of a hypothetical individual who would have a probability weighting parameter (γ) close to the levels displayed in Figure 3.1 (Section 3.4.4): a 22-year-old female IC with university degree and no family dependents ($\gamma \sim 0.5$); a 50-year-old female IC with university degree and family dependents ($\gamma \sim 1.0$); a 60-year-old male Control with no university degree and no family dependents ($\gamma \sim 2.0$); and, a 30-year-old male Control with a university degree and family dependents ($\gamma \sim 3.0$).

3.6.4. Between-task consistency

While individuals' risk preferences are not likely to change over the course of the experiment, the estimates generated from their choices can be different from one task to another. Risk preference is determined to be "consistent" when there is overlap between the interval ranges of two tasks' implied parameters. The implied Constant Relative Risk Aversion (CRRA) parameters for all choices in the four elicitation methods are available in Table G.1. of Appendix G: CRRA risk parameters for experiment choices.

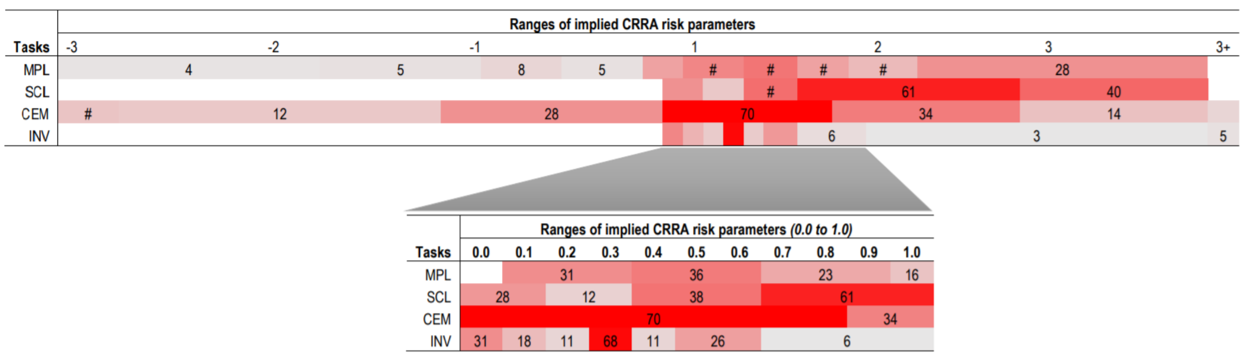

Figure 3.7 Count of participants within ranges of implied CRRA risk parameters.
MPL: Multiple Price List. SCL: Single Choice List.
CEM: Certainty Equivalent Method. INV: Investment Game.

As visually represented in Figure 3.7, 31 participants made a choice[34] in MPL that had an implied CRRA risk parameter that spans from $0.1 \leq \varphi < 0.4$. Of those 31 individuals, those who selected one of the two options[35] in SCL that ranged from $0.1 \leq \varphi < 0.4$ will be considered to be consistent between MPL-SCL.

[34] MPL: 31 participants chose to switch in Line 6 from *Option A* (60% probability of $20.00; 40% probability of $16.00) to *Option B* (60% probability of $38.50; 40% probability of $1.00). For a visual representation of this task, See *Task "Orange" Decision* in Appendix I: Instructions for the Risk Elicitation Economics Experiment.
[35] SCL: 28 participants chose the $48/0 lottery; 12 participants chose the $40/$4 lottery.

Table 3.20 reports pairwise task consistency to be strong between SCL-CEM (62%) MPL-SCL (58% of participants), MPL-CEM (51%).

Table 3.20 Proportion of participants with implied risk parameter consistency between tasks

	MPL	SCL	CEM	INV
MPL	1.0000			
SCL	0.5810	1.0000		
CEM	0.5140	0.6201	1.0000	
INV	0.3073	0.2067	0.4860	1.0000

MPL: Multiple Price List; SCL: Single Choice List; CEM: Certainty Equivalent Method INV: Investment Game

Following Holzmeister and Stefan (2020), we establish an index of between-task consistency. Each participant has an index ranging from 0/6 to 6/6 that indicates the number of task pairs in which the implied CRRA risk parameter intervals overlap. A distribution of participants across the between-task consistency index is displayed in Figure 3.8, below.

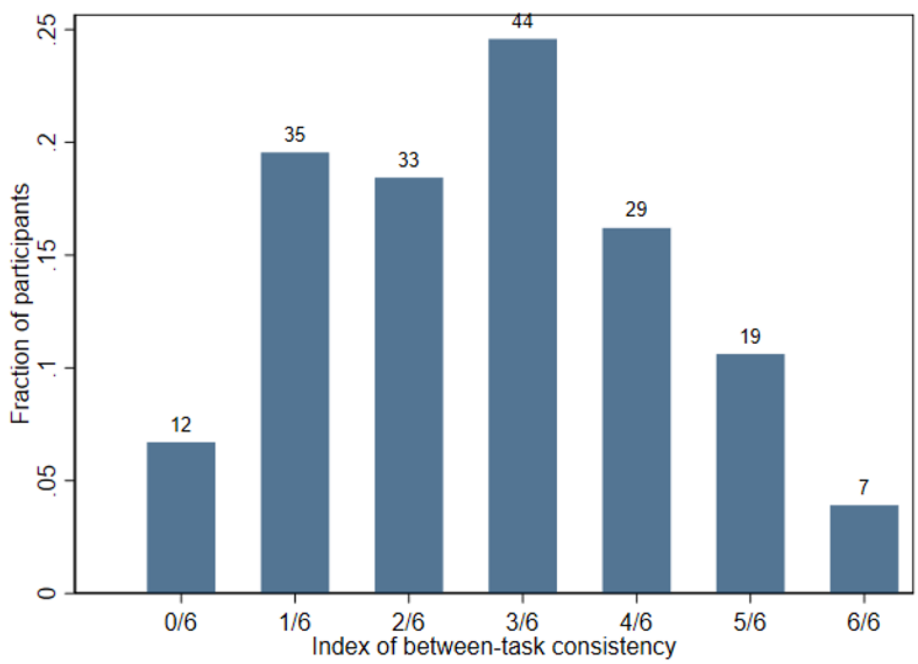

Figure 3.8 Distribution of participants across an index of consistency between pairs of tasks
(Count of participants per index is labelled above the graphs)

Figure 3.8 shows that on the bulk of participants had 1/6 to 4/6 task pairs in which their levels revealed risk were consistent between tasks. Among 179 participants, there are 12 whose implied CRRA ranges did not overlap between any task-wise pair (0/6), and seven whose implied CRRA ranges overlapped across all four elicitation methods (6/6).

3.6.5. Self-evaluated risk

3.6.5.1. Comparing between cohorts, and past/present risks

After having completed all risk elicitation methods (randomized order of MPL, SCL, CEM, INV), participants were asked to evaluate their own level of risk-taking. The format of the survey follows Nicholson et al. (2005), in which individuals are asked to rate, from a scale of 1 to 5, how risk-taking applies to themselves throughout six domains, both now and in their adult past. (For full text of this survey, see Appendix I: Instructions for the Risk Elicitation Economics Experiment.)

The self-evaluations of Control and Incident Commander participants are visually represented in Figure 3.9, and statistical tests on differences between risks across timeframes and cohorts are reported in Table 3.21 and Table 3.22.

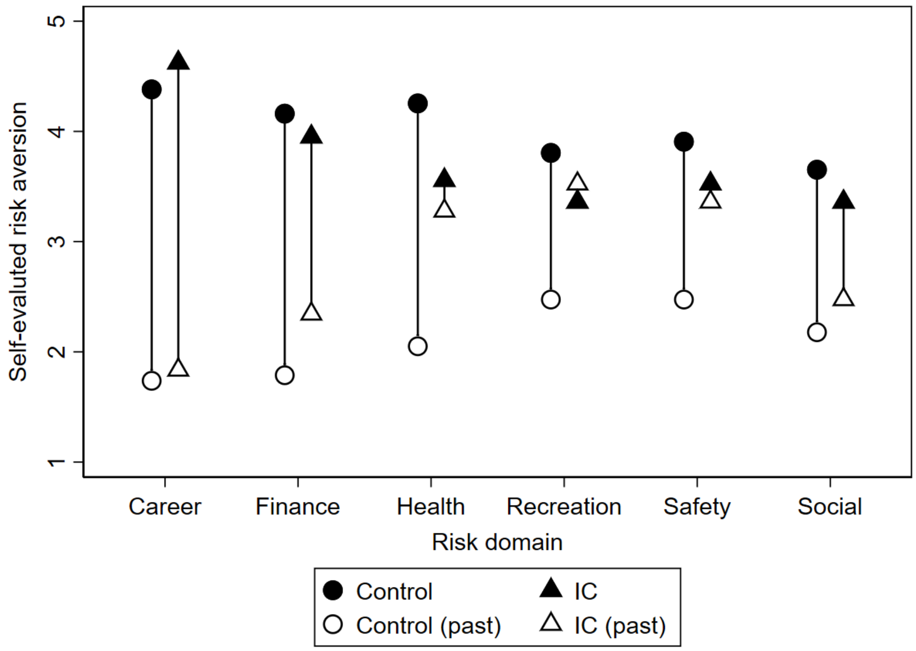

Figure 3.9 Responses to self-evaluated domain-specific risks
(1: low risk aversion; 5: high risk aversion)

Table 3.21 Comparing self-evaluated risk aversion between cohorts, timeframe controlled

Risk Domain	Now					Past				
	Control		IC		Wilcoxon	Control		IC		Wilcoxon
	N	Mean	N	Mean	p-value	N	Mean	N	Mean	p-value
Career	118	4.38	61	4.62	0.1393	118	1.74	61	1.84	0.2061
Financial	118	4.16	61	3.95	0.0213	118	1.79	61	2.34	0.0003
Health	118	4.25	61	3.56	0.0000	118	2.05	61	3.28	0.0000
Recreational	118	3.81	61	3.36	0.0051	118	2.47	61	3.52	0.0000
Safety	118	3.91	61	3.52	0.0134	118	2.47	61	3.36	0.0000
Social	118	3.65	61	3.36	0.0357	118	2.18	61	2.48	0.0740

Table 3.22 Comparing self-evaluated risk aversion between timeframes, controlled by cohort

Risk Domain	Control					Incident Commanders				
	Now		Past		Wilcoxon	Now		Past		Wilcoxon
	N	Mean	N	Mean	p-value	N	Mean	N	Mean	p-value
Career	118	4.38	118	1.74	0.0000	61	4.62	61	1.84	0.0000
Financial	118	4.16	118	1.79	0.0000	61	3.95	61	2.34	0.0000
Health	118	4.25	118	2.05	0.0000	61	3.56	61	3.28	0.1502
Recreational	118	3.81	118	2.47	0.0000	61	3.36	61	3.52	0.3636
Safety	118	3.91	118	2.47	0.0000	61	3.52	61	3.36	0.3344
Social	118	3.65	118	2.18	0.0000	61	3.36	61	2.48	0.0001

In general, participants report they are more risk averse today than in their adult past. The exception to this trend is ICs' risk taking in *Health*, *Recreational* and *Safety* domains, in which the differences between mean present-day and past risks are not significant. Although the average IC is older than the average Control (42 and 24-years-old), ICs tend to express a smaller difference in between current and past risk-taking in most domains, with the exception of *Career*. As such, ICs are simultaneously significantly less risk averse than Control in 5 of 6 domains in the present-day timeframe, and also significantly more risk averse in 4 of 6 domains in the past.

Considering that gender may affect self-evaluated risk preferences, we also compare between cohorts within gender, as well between genders within cohorts in Appendix H: Self-reported risk by cohort, gender, and timeframe. Tables in the appendix provide evidence that variation in self-evaluated risk is explained better by cohort rather than by gender.

3.6.5.2. Correlation between experiment choices and self-evaluated risk

To examine the correlation between participants' experiment-elicited risk and self-evaluated risk, both sets of data must be transformed so that they can be examined together. Firstly, we use principal component analysis (PCA) to extract the first principal component (PC1) from the set of four variables representing the choices in the four experimental tasks.

Participants' self-evaluated risk preferences are characterized by 12 variables. These variables are responses to a five-point Likert scale questionnaire on risk-taking across six domains and two timeframes (present, past)[36]. To capture the six variables of present-timeframe self-valuated risk as a single variable, PCA is repeated to extract the PC1. PCA is applied once again to capture past-timeframe self-evaluated risk as a single variable.

In Figure 3.10, below, each point represents an individual participant's PC1 of experimental choices and present-timeframe self-evaluated risk. Participants who make more risk averse choices in the incentivized tasks also tend to assess themselves as more risk averse. This positive relationship is highly significant *($p < 0.01$)* as reported in the OLS regression in Table 3.23.

[36] The self-valuation data is transformed to follow ascending risk aversion order. That is, in a Likert scale response to "Please could you tell us if any of the following have ever applied to you …", *1* becomes *very often* (least risk averse), and *5, never* (most risk averse).

For a screenshot of this questionnaire, see *Survey: Self-Evaluation* in Appendix I: Instructions for the Risk Elicitation Economics Experiment.

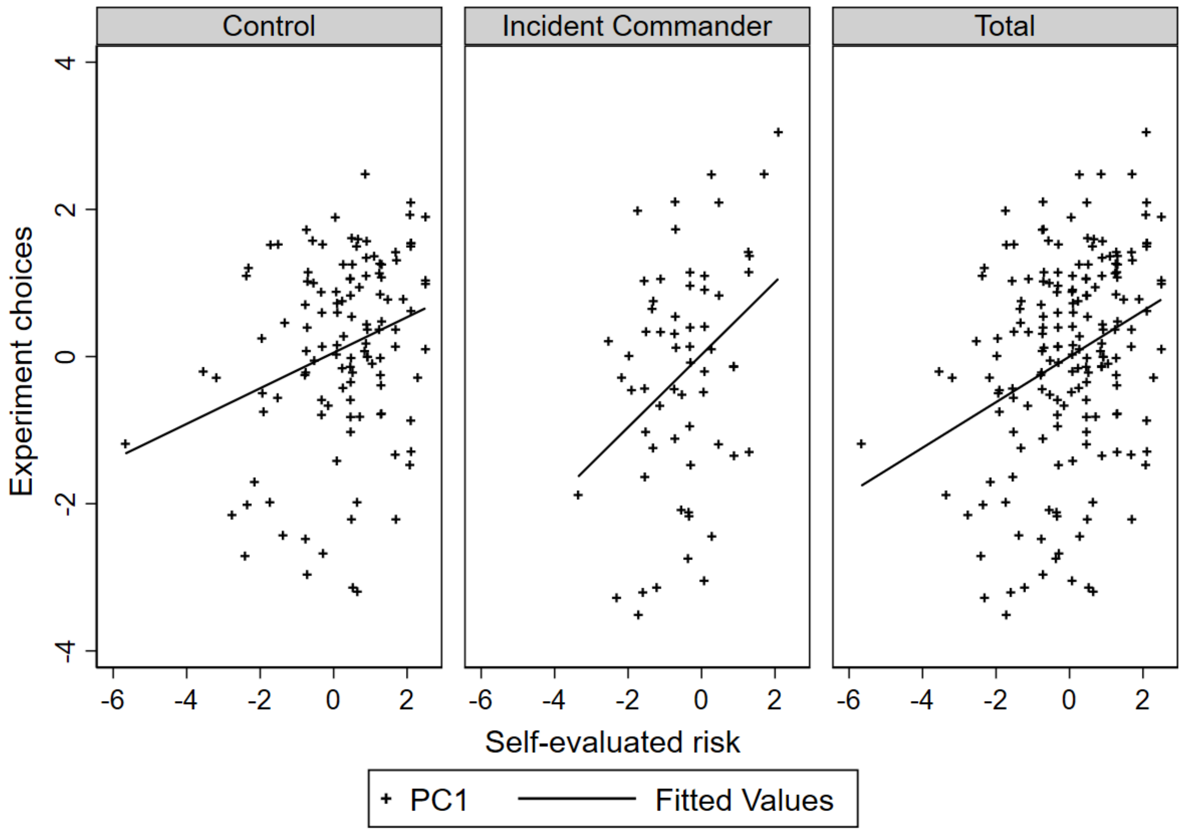

Figure 3.10 Correlation between First Principal Components of Risk elicitation experiment choices and Domain-specific self-evaluated risk (present timeframe).
(All choices are scaled in order of increasing risk aversion)

The relationship between risk aversion in Experiment choices and Self-evaluation is more pronounced among ICs (0.493) than Control (0.242). Regardless, in both cohorts and across all participants, the correlation is highly significant *($p < 0.01$)*, as reported in Table 3.23 below.

Table 3.23 OLS regression on PC1 of Experiment choices on PC1 of Self-evaluated risk (present)

	Control	Incident Commander	All participants
Self-evaluated risk (PC1)	0.242***	0.493***	0.310***
	(0.078)	(0.173)	(0.072)
Constant	0.052	0.026	-0.000
	(0.115)	(0.209)	(0.099)
Observations	118	61	179

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

In contrast, as seen below in Table 3.24, participants' self-evaluated risk-taking in the past timeframe appears to correlate negatively with elicited risk from experiment choices. This negative relationship is significant for all pooled participants and for the Control group.

Table 3.24 OLS regression on PC1 of Experiment choices on PC1 of Self-evaluated risk (past)

	Control	Incident Commander	All participants
Self-evaluated past risk (PC1)	-0.259***	-0.199	-0.242***
	(0.079)	(0.152)	(0.065)
Constant	-0.006	-0.042	-0.000
	(0.118)	(0.243)	(0.100)
Observations	118	61	179

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Principal Component Analysis captures domain-specific risk self-evaluations as one variable. However, we are also interested in seeing which individual risk domain(s) tend to correlate best with experimental-elicited risk. To investigate this, we regress PC1 of experiment choices on the Likert scale units of individual risk domains. Table 3.25 reports six regressions of Experiment choice PC1 present-timeframe self-evaluated risk by domain, across all participants.

Table 3.25 OLS regression of Experiment choice PC1 on Self-evaluated risk by domain

	Career	*Financial*	*Health*	*Recreation*	*Safety*	*Social*
Self-eval. risk	-0.020	0.485***	0.144	0.218**	0.330***	0.144
	(0.109)	(0.104)	(0.098)	(0.086)	(0.096)	(0.088)
Constant	0.090	-1.983***	-0.578	-0.798**	-1.244***	-0.511
	(0.497)	(0.436)	(0.406)	(0.331)	(0.375)	(0.331)
Observations	179	179	179	179	179	179

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Among all the self-evaluated risk domains, *Financial* is most strongly and highly significantly associated with number of safe choices in experiments (0.485, $p < 0.01$); *Safety* is also highly significant *($p < 0.01$)*, and *Recreation* is significant *($p < 0.05$)*. However, when participants are separated by gender and cohort, it is evident that the relationship between certain domain-specific self-evaluated risks and experiment choice varies among subgroups (Table 3.26). While all split samples have a positive correlation between self-evaluated *Financial* risks and experiment choices, this relationship is not significant for Males in the Control group. As well, the relationship between self-evaluated *Safety* risks and experiment choices only maintains significance for Male participants in Control group *($p < 0.05$)* and Incident Commander group *($p < 0.01$)*, and *Recreation* is only marginally significant among Female Control participants *($p < 0.10$)*.

Table 3.26 OLS regression on PC1 of experiment choices on self-evaluated risks, by domain (separated by subgroups of participants)

	Career	Financial	Health	Recreation	Safety	Social
Panel A: Control, Male						
Self-eval. risk	-0.000	0.217	0.134	0.071	0.450**	0.227
	(0.181)	(0.177)	(0.190)	(0.180)	(0.215)	(0.152)
Constant	-0.231	-1.063	-0.786	-0.490	-2.060**	-1.052*
	(0.808)	(0.709)	(0.815)	(0.690)	(0.896)	(0.586)
N	47	47	47	47	47	47
Panel B: Control, Female						
Self-eval. risk	-0.138	0.505***	0.090	0.209*	0.129	0.120
	(0.142)	(0.148)	(0.156)	(0.108)	(0.120)	(0.118)
Constant	0.957	-1.866***	-0.045	-0.472	-0.144	-0.096
	(0.644)	(0.661)	(0.686)	(0.441)	(0.475)	(0.456)
N	71	71	71	71	71	71
Panel C: Incident Commander, Male						
Self-eval. risk	0.145	0.575**	0.095	0.141	0.605***	-0.086
	(0.258)	(0.254)	(0.196)	(0.190)	(0.185)	(0.195)
Constant	-0.972	-2.567**	-0.645	-0.768	-2.425***	-0.020
	(1.202)	(1.017)	(0.725)	(0.655)	(0.675)	(0.682)
N	56	56	56	56	56	56
Panel D: Incident Commander, Female						
Self-eval. risk	0.000 †	2.004**	-0.195	1.127	-0.516	1.477
	(.)	(0.611)	(0.713)	(1.037)	(1.274)	(0.878)
Constant	0.735	-7.681*	1.435	-4.225	2.695	-4.581
	(0.848)	(2.608)	(2.743)	(4.636)	(4.935)	(3.239)
N	5	5	5	5	5	5

† No variation among subgroup in this domain-specific risk.
Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

3.7. Discussion

This experiment was motivated by a research question on whether, and how, Alberta Wildfire Incident Commanders differ from a Control group in their level of risk aversion. As well, we were interested in comparing the application of two prevalent behavioural analysis theories, and considering how our choice of elicitation methods influences expressions of risk behaviour.

Discussion on Analysis on Incident Commanders and Risk Aversion

Risk literature has proven that subjects who engage in risk professions (Krčál et al., 2019), recreation (Riddel and Kolstoe, 2013), or have recently experienced trauma (Eckel et al., 2009) tend to be less risk averse. In this project, we were interested in finding empirical evidence for whether Incident Commanders (ICs) have risk preferences that are different than the control population. We find that, after controlling for heterogeneity in gender, age, education and family status, ICs are not particularly different from Control participants.

While evidence of risk aversion is not clear between cohorts, we discover that participants within the IC cohort are differentiated by the level of experience in operational deployment. ICs who with more operations experience in the past year tend to be less risk averse across nearly every elicitation method (Table 3.8 – Table 3.12). This finding demonstrates that there is a clear inverse correlation between an Incident Commander's engagement in wildland firefighting in the year previous to the experiment, and risk aversion elicited from incentivized tasks. When we extend observations of ICs' operations experience from 2020 to 2011 (Table 3.13 – Table 3.17), we find that additional experience in past years are not as impactful on risk behaviour as experience in recent years. Recent experience is more impactful than past experience on risk aversion revealed in the laboratory. This finding on the time-decaying effect of heightened-stress experience conforms with previous literature (Eckel et al., 2009).

Discussion on Analysis on Experiment Design and Analysis

All four tasks were analyzed using the most commonly used framework in experimental literature: the Expected Utility Model (Crosetto and Filippin, 2016; Dave et al., 2010; Eckel and Grossman, 2008; Holt and Laury, 2002). Prospect Theory, utilized in Pedroni et al. (2017) and Tanaka et al. (2010), was applied to the Multiple Price List (MPL), to account for probability weighting. We discover the PT model captures variation in MPL choices slightly better (pseudo-R^2 = 0.1066) than EUT (pseudo-R^2 = 0.0949).

When testing for between-task risk consistency, we observe a large proportion of our participants (93%) can be considered consistent between at least one pair of tasks (Table 3.20). Individuals exhibit consistency predominantly between tasks in which the implied CRRA parameters of available choices are comparable; among our tasks, these are CEM and MPL.

The reason why subjects seem to behave differently between elicitation methods during an experiment is a subject of much contention amongst experimental economists (Crosetto and Filippin, 2016; Dave et al., 2010; Holzmeister and Stefan, 2020; Pedroni et al., 2017). Assertions made in some works are observed as well in the results of our experiment. For instance, Dave et al. (2010) praise MPL for its predictive accuracy, but also caution against using the method due to its complexity. Indeed, we do find that Control participants are marginally significant in exhibiting mildly noisy behaviour both in MPL and in the other lottery task, CEM (Table 3.7). Filippin and Crosetto (2016) argue that gender differences are more likely to be observed when a task, a) includes a riskless option in the set, b) has fixed 50/50 probabilities. This phenomenon is observed, too, in our experiment, as females display significantly higher risk aversion in the three tasks with equal outcome probabilities (SCL, INV, CEM), of which two have a riskless option (SCL, INV).

In addition to task design, the order in which a participant is presented a task can also influence her behaviour (Harrison et al., 2008, 2005), a limitation that is further discussed below.

In comparing experiment revealed and self-evaluated risk through Principal Component Analysis, we find a strong correlation between experiment risk and self-evaluated risk. Participants who make risky choices in the four elicitation tasks also tend to assess themselves as more risk-taking, particularly in domains of *Financial* and *Safety* risks. This finding leads us to believe that there is a clear relationship between risk as elicited through incentivized tasks, and risk that is reported by an individual, after careful self-reflection. We also consider the possibility that since the self-evaluation task followed the four incentivized risk elicitation tasks, reflection on personal risk tolerance during the lottery tasks may be influencing participants' self-evaluation.

Limitations

While many of our findings are statistically significant and robust, we wish to address some of the limitations in our experiment.

Firstly, we recognize that Control and Incident Commander groups are comprised of particularly diverse individuals, in terms of gender, age, education, and family status (Table 3.3). Heterogeneity between participants was controlled during structural estimation using characteristic variables, but nonetheless, we recognize that unobserved variation may persist between considerably different groups of individuals.

Randomizing the presentation order of incentivized tasks helps avoid order effects across tasks (Carlsson et al., 2012; Holzmeister and Stefan, 2020). However, it remains possible that participants' behaviour in the second, third or fourth task will be different than, and perhaps influenced by the choice made in the first task (Harrison et al., 2008). An alternative analysis approach towards addressing order choice would be to analyze tasks separately by *n*th-order of presentation. However, in this study, implementation of such an approach is challenging due to our modest sample size, particularly of Incident Commanders ($n = 61$).

Income effect is another concern, particularly for Incident Commanders' participation in the experiment. Through their responses to the invitation email, ICs demonstrated enthusiasm for contributing towards wildfire research. For ICs well into their professional lives, the modest compensation offered by the experiment may not influence their behaviour in incentivized tasks in the way it may for our selected Control group. The risk behaviour of control subjects, mainly university undergraduate students, may are more influenced by the incentive of earning $10 - $60 within half an hour.

In addition, as our experiment period (January – April 2021) took place during the COVID-19 pandemic, public health restrictions obligated us to carry out this experiment online, rather than in the experimental laboratory in the Department of REES. In order to accommodate as many participants as feasible, individuals were instructed to complete the experiment at their own pace within 48 hours. Participants were instructed at the start of the experiment to undertake the experiment in a confidential and serious manner (*Information Page* and *Consent Form,* Appendix I: Instructions for the Risk Elicitation Economics Experiment). However, in the absence of experimenter monitoring, we were unable to assess participant attentiveness.

3.8. Conclusion

Results from risk elicitation methods show that, after controlling for key factors like age and gender, a cohort of 61 Alberta Wildfire Incident Commanders do not exhibit risk aversion that is particularly different from that of the general public. However, when asked about their personal level of risk-taking, Incident Commanders would rank themselves to be comparatively less risk averse than those of control participants.

Within the IC cohort, there is a noticeable correlation between the level of risk an IC is willing to take in an experiment, and the amount of experience she has on the field, in terms of days worked as well as her frequency of taking the Duty Officer role. It is challenging to ascertain the direction of this relationship, i.e. whether ICs more comfortable with taking risk are more likely to work more and in positions of greater responsibility, or whether the experience gained by ICs make them exhibit less risk aversion. However, given the relationship between experience variables and revealed risk, we suggest that additional research can be performed to better understand this connection.

Considering that Alberta Wildfire ICs make decisions collaboratively and within an operational framework, we suggest that future research with ICs should capture the effect of group dynamics on decision-making. Additional research into individual and collective decision-making, including the role of risk preferences in influencing decisions, can address the current gap in literature which seeks to understand how wildfire suppression expenditures vary between incident management teams (Canton-Thompson et al., 2008; Hand et al., 2017). As well, insights into Incident Commanders' risk preferences could also be improved through qualitative studies (e.g. structured interviews), in which participants have the opportunity to express certain attitudes that may not be expressed solely through quantitative data.

Chapter 4. Conclusion

While suppression expenditures are largely driven by environmental variables, actions taken by Alberta Wildfire can make marginal, although significant impacts on reducing the cost of wildfire suppression. These include actions that Alberta Wildfire are already taking, such as fuel management in vulnerable areas, improving the rapidity of wildfire detection through technological renewal, and the decision to prioritize allocate resources to the necessary fires. Our empirical research demonstrates the effects of such impacts, such an 0.2% increase in costs when a wildfire report is delayed by an hour, as well as significant reductions up to 141% in certain large fires when resource allocation is strategically delayed.

Nevertheless, environmental factors, including environmental factors like high temperature, wind and fuel types are principally responsible for driving wildfire costs, by 79% or more. As climate change continues to exacerbate the fire weather conditions (Flannigan et al., 2000; Robinne et al., 2016; Tymstra et al., 2021; Wotton et al., 2017), Alberta Wildfire can expect to see environmental variables continue to drive the bulk of suppression expenditures. Focusing on making changes where improvements can be made, such as in improving wildfire detection, will allow the wildfire management agency to make meaningful, albeit modest reductions in expenditures.

Experiments in risk preference elicitation allow us to shed light on how wildland firefighters express risk aversion in four incentivized economic games. We had initially expected our sample of 61 Alberta Wildfire Incident Commanders (ICs) to be less risk-averse, or even risk-seeking, due to the nature of their particularly profession. Instead, we find that ICs do not differ too much from control subjects, and are actually significantly more risk averse in one of the four tasks. Consistent with the reviewed literature (Crosetto and Filippin, 2016; Eckel and Grossman, 2008; Pedroni et al., 2017), a subject's gender is the most influential characteristic on risky choices. Among the IC cohort, we observe a significant correlation between reduced risk aversion, as elicited through our incentivized tasks, and job experience. ICs with additional job experience, measured as the number of days in the operational field or days taking the Duty Officer role responsible for resource allocation, will generally make less risk averse choices in incentivized experiments. In addition, ICs' job experience in the recent past has more influence on this experience-risk relationship than experience of more distant past.

Incident commanders' risk perceptions is one piece of wildfire expenditure puzzle. An understanding of how ICs perceive risks can help wildfire managers better comprehend how and why certain resource allocation choices are made. Risk elicitation through standardized economic games, while familiar to the standard university experiment pool (i.e. students/recent graduates), may be unfamiliar to working professionals like our IC cohort. As such, future research with ICs may be developed with more realistic settings, such as the elicitation of suppression strategy preferences when presented with hypothetical wildfire risks, as carried out by US Forest Service researchers (Calkin et al., 2013; Hand et al., 2015). Additionally, insights into Incident Commanders' risk preferences could also be ameliorated by adding interviews qualitative studies, allowing participants to express certain attitudes otherwise not captured through empirical quantitative methods. Further understanding into ICs' behaviours and decision-making can help Alberta Wildfire, and other wildfire management agencies, better prepare their staff for their operational roles, in an everchanging wildland fire landscape.

Bibliography

Abdellaoui, M., Bleichrodt, H., Kammoun, H., 2013. Do financial professionals behave according to prospect theory? An experimental study. Theory Dec 74, 411–429. https://doi.org/10.1007/s11238-011-9282-3

Abdellaoui, M., Bleichrodt, H., L'Haridon, O., 2008. A tractable method to measure utility and loss aversion under prospect theory. J. Risk Uncertain. 36, 245–266. https://doi.org/10.1007/s11166-008-9039-8

Abdellaoui, M., Driouchi, A., L'Haridon, O., 2011. Risk aversion elicitation: Reconciling tractability and bias minimization. Theory Decis. 71, 63–80. https://doi.org/10.1007/s11238-009-9192-9

Adab, H., Kanniah, K.D., Solaimani, K., 2013. Modeling forest fire risk in the northeast of Iran using remote sensing and GIS techniques. Nat. Hazards 65, 1723–1743. https://doi.org/10.1007/s11069-012-0450-8

Adetona, O., Reinhardt, T.E., Domitrovich, J., Broyles, G., Adetona, A.M., Kleinman, M.T., Ottmar, R.D., Naeher, L.P., 2016. Review of the health effects of wildland fire smoke on wildland firefighters and the public. https://doi.org/10.3109/08958378.2016.1145771 28, 95–139. https://doi.org/10.3109/08958378.2016.1145771

Alam, R., Islam, S., Mosely, E., Thomas, S., Dowdell, V., Doel, D., 2017. Rapid Impact Assessment of Fort McMurray Wildfire.

Alberta, 2018. Occupational Disease Fatalities Accepted by the Workers' Compensation Board: Year to date, numbers as of September 30, 2018.

Alberta Wildfire, 2019. Wildfire Detection | AAF - Agriculture and Forestry [WWW Document]. URL https://wildfire.alberta.ca/operations/wildfire-detection/default.aspx (accessed 11.15.21).

Alberta Wildfire, 2017. Operations | AAF - Agriculture and Forestry [WWW Document]. URL https://wildfire.alberta.ca/operations/default.aspx (accessed 11.14.21).

Alberta Wildfire, 2015. Alberta Wildland Firefighter - Unit Crew Program (UNIT) - YouTube [WWW Document]. URL https://www.youtube.com/watch?v=yoGe1uXgJaM (accessed 1.18.22).

Almlund, M., Duckworth, A.L., Heckman, J., Kautz, T., 2011. Personality Psychology and

Economics. Handb. Econ. Educ. 4, 1–181. https://doi.org/10.1016/B978-0-444-53444-6.00001-8

Anderson, L.R., Mellor, J.M., 2008. Predicting health behaviors with an experimental measure of risk preference. J. Health Econ. 27, 1260–1274.

Andersson, O., Holm, H.J., Tyran, J.-R., Wengström, E., 2016. Risk Aversion Relates to Cognitive Ability: Preferences Or Noise? J. Eur. Econ. Assoc. 14, 1129–1154. https://doi.org/10.1111/JEEA.12179

Arienti, M.C., Cumming, S.G., Boutin, S., 2006. Empirical models of forest fire initial attack success probabilities: the effects of fuels, anthropogenic linear features, fire weather, and management. https://doi.org/10.1139/X06-188

Arnold, P.D., Zai, G., Richter, M.A., 2004. Genetics of anxiety disorders. Curr. Psychiatry Reports 2004 64 6, 243–254. https://doi.org/10.1007/S11920-004-0073-1

Arrow, K., 1965. Aspects of the Theory of Risk Bearing. Yrjö Jahnssonin Säätiö, Helsinki.

Ayaita, A., Stürmer, K., 2020. Risk aversion and the teaching profession: an analysis including different forms of risk aversion, different control groups, selection and socialization effects. Educ. Econ. 28, 4–25. https://doi.org/10.1080/09645292.2019.1675592

Balaz, V., Williams, A.M., 2011. Risk attitudes and migration experience. J. Risk Res. 14, 583–596. https://doi.org/10.1080/13669877.2010.547256

Barber, B.M., Odean, T., 2001. Boys will be Boys: Gender, Overconfidence, and Common Stock Investment. Q. J. Econ. 116, 261–292. https://doi.org/10.1162/003355301556400

Barberis, N.C., 2013. Thirty Years of Prospect Theory in Economics: A Review and Assessment. J. Econ. Perspect. 27, 173–96. https://doi.org/10.1257/JEP.27.1.173

Bayham, J., Yoder, J.K., 2020. Resource allocation under fire. Land Econ. 96, 92–110. https://doi.org/10.3368/le.96.1.92

Becker, G.M., DeGroot, M.H., Marschak, J., 1964. Measuring utility by a single-response sequential method. Behav. Sci. 9, 226–232. https://doi.org/10.1002/bs.3830090304

Benjamin, D.J., Brown, S.A., Shapiro, J.M., 2013. Who is 'Behavioral'? Cognitive Ability and Anomalous Preferences. J. Eur. Econ. Assoc. 11, 1231–1255. https://doi.org/10.1111/JEEA.12055

Bernoulli, D., 1738. Specimen Theoriae Novae de Mensura Sortis. Coment. Acad. Sci. Imp. Petropolitanae translated.

Beverly, J.L., Flannigan, M.D., Stocks, B.J., Bothwell, P., 2011. The association between Northern Hemisphere climate patterns and interannual variability in Canadian wildfire activity. Can. J. For. Res. 41, 2193–2201. https://doi.org/10.1139/X11-131

Binswanger, H.H.P., 1980. Attitudes toward Risk : Experimental Measurement In Rural India. Am. J. Agric. Econ. 62, 395–407. https://doi.org/10.2307/1240194

Blais, A.-R., Weber, E.U., 2006. A Domain-Specific Risk-Taking (DOSPERT) scale for adult populations. Judgm. Decis. Mak. 1, 33–47.

Bombardini, M., Trebbi, F., 2012. Risk Aversion and Expected Utility Theory: An Experiment with Large and Small Stakes. J. Eur. Econ. Assoc. 10, 1348–1399. https://doi.org/10.1111/J.1542-4774.2012.01086.X

Boulianne, S., Minaker, J., Haney, T.J., 2018. Does compassion go viral? Social media, caring, and the Fort McMurray wildfire. Inf. Commun. Soc. 21, 697–711. https://doi.org/10.1080/1369118X.2018.1428651

Brown, J.K., 1983. The "unnatural fuel buildup" issue., in: JE Lotan et al (Ed.), Proceedings - Symposium and Workshop on Wilderness Fire.USDA Forest Service Intermountain Forest and Range Experiment Station, General Technical Report INT-128. pp. 127–128.

Brown, S., Dietrich, M., Ortiz-Nuñez, A., Taylor, K., 2011. Self-employment and attitudes towards risk: Timing and unobserved heterogeneity. J. Econ. Psychol. 32, 425–433. https://doi.org/10.1016/j.joep.2011.02.015

Bruner, D.M., 2009. Changing the probability versus changing the reward. Exp. Econ. 12, 367–385.

Burks, S. V., Carpenter, J.P., Goette, L., Rustichini, A., 2009. Cognitive skills affect economic preferences, strategic behavior, and job attachment. Proc. Natl. Acad. Sci. 106, 7745–7750. https://doi.org/10.1073/PNAS.0812360106

Burlig, F., Knittel, C., Rapson, D., Wolfram, C., 2020. Machine Learning from Schools about Energy Efficiency. https://doi.org/10.7910/DVN/NDAYCR

Cai, X., Wang, X., Jain, P., Flannigan, M.D., 2019. Evaluation of Gridded Precipitation Data and Interpolation Methods for Forest Fire Danger Rating in Alberta, Canada. J. Geophys. Res. Atmos. 124, 3–17. https://doi.org/10.1029/2018JD028754

Calkin, D., Hyde, K., Gebert, K., Jones, G., 2005. Comparing resource values at risk from wildfires with forest service fire suppression expenditures: Examples from 2003 western Montana wildfire season, USDA Forest Service - Research Note RMRS-RN. https://doi.org/10.2737/RMRS-RN-24

Calkin, D.E., Venn, T., Wibbenmeyer, M., Thompson, M.P., 2013. Estimating US federal wildland fire managers' preferences toward competing strategic suppression objectives. Int. J. Wildl. Fire 22, 212–222. https://doi.org/10.1071/WF11075

Cameron, L., Shah, M., 2015. Risk-taking behaviour in the wake of natural disasters. J. Hum. Resour. 50, 484–515.

Canadian Wildland Fire Strategy Assistant Deputy Ministers Task Group, 2005. Canadian wildland fire strategy: A vision for an innovative and integrated approach to managing the risks. A report to the Canadian Council of Forest Ministers, prepared by the Canadian Wildland Fire Strategy Assistant Deputy Ministers Task Group, Canadian Forest Service.

Canton-Thompson, J., Gebert, K.M., Thompson, B., Jones, G., Calkin, D., Donovan, G., 2008. External human factors in incident management team decisionmaking and their effect on large fire suppression expenditures. J. For. 106, 416–424.

https://doi.org/10.1093/jof/106.8.416

Carlsson, F., Mørkbak, M.R., Olsen, S.B., 2012. The first time is the hardest: A test of ordering effects in choice experiments. J. Choice Model. 5, 19–37. https://doi.org/10.1016/S1755-5345(13)70051-4

Casamatta, G., Giannoni, S., Brunstein, D., Jouve, J., 2022. Host type and pricing on Airbnb: Seasonality and perceived market power. Tour. Manag. 88, 104433. https://doi.org/10.1016/J.TOURMAN.2021.104433

CBC News, 2021. Alberta to test new wildfire-fighting technology this season for about $4.3M. CBC News.

Charness, G., Eckel, C., Gneezy, U., Kajackaite, A., 2018. Complexity in risk elicitation may affect the conclusions: A demonstration using gender differences. J. Risk Uncertain. 56, 1–17. https://doi.org/10.1007/s11166-018-9274-6

Charness, G., Gneezy, U., 2012. Strong Evidence for Gender Differences in Risk Taking. J. Econ. Behav. Organ. 83, 50–58. https://doi.org/10.1016/j.jebo.2011.06.007

Charness, G., Gneezy, U., Imas, A., 2013. Experimental methods: Eliciting risk preferences. J. Econ. Behav. Organ. 87, 43–51. https://doi.org/10.1016/j.jebo.2012.12.023

Chaulk, B., Johnson, P.J., Bulcroft, R., 2003. Effects of Marriage and Children on Financial Risk Tolerance: A Synthesis of Family Development and Prospect Theory. J. Fam. Econ. Issues 2003 243 24, 257–279. https://doi.org/10.1023/A:1025495221519

Cheek, J., 2021. The lonely life of a wildfire lookout in northern Alberta - Macleans.ca. Macleans.

Chen, D.L., Schonger, M., Wickens, C., 2016. oTree-An open-source platform for laboratory, online, and field experiments. J. Behav. Exp. Financ. 9, 88–97. https://doi.org/10.1016/j.jbef.2015.12.001

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., Robins, J., 2018. Double/debiased machine learning for treatment and structural parameters. Econom. J. 21, C1–C68. https://doi.org/10.1111/ECTJ.12097

Christopoulos, G.I., Tobler, P.N., Bossaerts, P., Dolan, R.J., Schultz, W., 2009. Neural Correlates of Value, Risk, and Risk Aversion Contributing to Decision Making under Risk. J. Neurosci. 29, 12574–12583. https://doi.org/10.1523/JNEUROSCI.2614-09.2009

CIFFC, 2021. Homepage | CIFFC [WWW Document]. URL https://www.ciffc.ca/ (accessed 11.4.21).

Clark, A.M., Rashford, B.S., McLeod, D.M., Lieske, S.N., Coupal, R.H., Albeke, S.E., 2016. The impact of residential development pattern on wildland fire suppression expenditures. Land Econ. 92, 656–678. https://doi.org/10.3368/le.92.4.656

Colgan, G., 2021. Kananaskis Country sees record visitors for 2020. RMO Today.

Coogan, S.C.P., Daniels, L.D., Boychuk, D., Burton, P.J., Flannigan, M.D., Gauthier, S., Kafka, V., Park, J.S., Mike Wotton, B., 2021. Fifty years of wildland fire science in Canada. Can. J. For. Res. 51, 283–302. https://doi.org/10.1139/cjfr-2020-0314

Crockett, M.J., Fehr, E., 2014. Pharmacology of Economic and Social Decision Making. Neuroeconomics Decis. Mak. Brain Second Ed. 259–279. https://doi.org/10.1016/B978-0-12-416008-8.00014-0

Crosetto, P., Filippin, A., 2017. Safe Options Induce Gender Differences in Risk Attitudes. IZA Discuss. Pap.

Crosetto, P., Filippin, A., 2016. A theoretical and experimental appraisal of four risk elicitation methods. Exp. Econ. 19, 613–641.

Crosetto, P., Filippin, A., 2013. The "bomb" risk elicitation task. J. Risk Uncertain. 47, 31–65. https://doi.org/10.1007/s11166-013-9170-z

CTV News, 2021. "It's just tragic news": Pilot fighting wildfire killed in helicopter crash west of Edmonton | CTV News.

Dave, C., Eckel, C.C., Johnson, C.A., Rojas, C., 2010. Eliciting risk preferences: When is simple better? J. Risk Uncertain. 41, 219–243. https://doi.org/10.1007/s11166-010-9103-z

Deck, C., Lee, J., Rayes, J.A., Rosen, C.C., 2013. A failed attempt to explain within subject variation in risk taking behavior using domain specific risk attitudes. J. Econ. Behav. Organ. 87, 1–24. https://doi.org/10.1016/j.jebo.2012.11.010

Denburg, N.L., Cole, C.A., Hernandez, M., Yamada, T.H., Tranel, D., Bechara, A., Wallace, R.B., 2007. The Orbitofrontal Cortex, Real-World Decision Making, and Normal Aging. Ann. N. Y. Acad. Sci. 1121, 480. https://doi.org/10.1196/ANNALS.1401.031

Desmond, M., 2009. On the FireLine: Living and Dying with Wildland Firefighters. University of Chicago Press, Chicago.

Dickson, C., Kulkarni, A., 2021. B.C. fire crews risked their lives to help residents who ignored evacuation orders: minister. CBC News.

Dillon, J.L., Scandizzo, P.L., 1978. Risk Attitudes of Subsistence Farmers in Northeast Brazil: A Sampling Approach. Am. J. Agric. Econ. 60, 425–435. https://doi.org/10.2307/1239939

Dohmen, T., Falk, A., Huffman, D., Sunde, U., 2010. Are risk aversion and impatience related to cognitive ability? Am. Econ. Rev. 100, 1238–1260. https://doi.org/10.1257/aer.100.3.1238

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., Wagner, G.G., T Dohmen, D Huffman, J Schupp, A Falk, U Sunde, G.W., Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., Wagner, G.G., 2011. Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences. J. Eur. Econ. Assoc. 9, 522–550. https://doi.org/10.1111/j.1542-4774.2011.01015.x

Donovan, G.H., Brown, T.C., 2005. An alternative incentive structure for wildfire management on national forest land. For. Sci. 51, 387–395. https://doi.org/10.1093/forestscience/51.5.387

Donovan, G.H., Noordijk, P., Radeloff, V., 2004. Estimating the Impact of Proximity of Houses on Wildfire Suppression Costs in Oregon and Washington. Proc. Second Int. Symp. Fire Econ. Planning, Policy A Glob. View 697–702.

Donovan, G.H., Prestemon, J.P., Gebert, K., 2011. The effect of newspaper coverage and political pressure on wildfire suppression costs. Soc. Nat. Resour. 24, 785–798. https://doi.org/10.1080/08941921003649482

Dreber, A., Rand, D., Garcia, J., Wernerfelt, N., Lum, J., 2010. Dopamine and risk preferences in different domains, Harvard Kennedy School Faculty Research Working Paper Series.

Drolet, J.L., McDonald-Harker, C., Lalani, N., McGreer, S., Brown, M.R.G., Silverstone, P.H., 2021. Early Childhood Development in the Aftermath of the 2016 Wildfires in Alberta, Canada. J. Educ. Emergencies 7, 16. https://doi.org/10.33682/u3ar-wwzm

Eckel, C.C., El-Gamal, M.A., Wilson, R.K., 2009. Risk loving after the storm: A Bayesian-Network study of Hurricane Katrina evacuees. J. Econ. Behav. Organ. 69, 110–124. https://doi.org/10.1016/j.jebo.2007.08.012

Eckel, C.C., Grossman, P.J., 2008. Forecasting risk attitudes: An experimental study using actual and forecast gamble choices. J. Econ. Behav. Organ. 68, 1–17. https://doi.org/10.1016/j.jebo.2008.04.006

Eckel, C.C., Grossman, P.J., 2002. Sex differences and statistical stereotyping in attitudes toward financial risk. Evol. Hum. Behav. 23, 281–295. https://doi.org/10.1016/S1090-5138(02)00097-1

Eckel, C.C., Grossman, P.J., Johnson, C.A., de Oliveira, A.C.M., Rojas, C., Wilson, R.K., 2012. School environment and risk preferences: Experimental evidence. J. Risk Uncertain. 45, 265–292. https://doi.org/10.1007/s11166-012-9156-2

Ergibi, M., Hesseln, H., 2020. Awareness and adoption of FireSmart Canada: Barriers and incentives. For. Policy Econ. 119, 102271. https://doi.org/10.1016/j.forpol.2020.102271

F Holzmeister, A.P., 2016. oTree: The “bomb” risk elicitation task. J. Behav. Exp. Financ. 10, 105–108. https://doi.org/10.1016/j.jbef.2016.03.004

Figner, B., Mackinlay, R.J., Wilkening, F., Weber, E.U., 2009. Affective and Deliberative Processes in Risky Choice: Age Differences in Risk Taking in the Columbia Card Task. J. Exp. Psychol. Learn. Mem. Cogn. 35, 709–730. https://doi.org/10.1037/A0014983

Filippin, A., Crosetto, P., 2016. A reconsideration of gender differences in risk attitudes. Manage. Sci. 62, 3138–3160. https://doi.org/10.1287/mnsc.2015.2294

FireSmart Canada, 2021. What is FireSmart? Empowering the public and increasing neighbourhood resilience. [WWW Document]. URL https://www.firesmartcanada.ca/what-is-firesmart/ (accessed 11.4.21).

Flannigan, M.D., Logan, K.A., Amiro, B.D., Skinner, W.R., Stocks, B.J., 2005. Future area burned in Canada. Clim. Change 72, 1–16. https://doi.org/10.1007/s10584-005-5935-y

Flannigan, M.D., Stocks, B.J., Wotton, B.M., 2000. Climate change and forest fires. Sci. Total Environ. 262, 221–229. https://doi.org/10.1016/S0048-9697(00)00524-6

Flory, J.D., Manuck, S.B., Matthews, K.A., Muldoon, M.F., 2004. Serotonergic function in the central nervous system is associated with daily ratings of positive mood. Psychiatry Res. 129, 11–19. https://doi.org/10.1016/J.PSYCHRES.2004.06.010

Frederick, S., 2005. Cognitive reflection and decision making. J. Econ. Perspect. 19, 25–42. https://doi.org/10.1257/089533005775196732

Frey, R., Pedroni, A., Mata, R., Rieskamp, J., Hertwig, R., 2017. Risk preference shares the psychometric structure of major psychological traits. Sci. Adv. 3, 1–14. https://doi.org/10.1126/sciadv.1701381

Friedman, D., Habib, S., Duncan, J., Crockett, S., 2018. Varieties of risk elicitation (No. 744). Santa Cruz, CA.

Frydman, C., Camerer, C.F., 2016. The Psychology and Neuroscience of Financial Decision Making. Trends Cogn. Sci. 20, 661–675. https://doi.org/10.1016/J.TICS.2016.07.003

Gaure, S., 2010. OLS with Multiple High Dimensional Category Dummies OLS with Multiple High Dimensional Category Dummies, Memorandum, No. 2010,14. Oslo.

Gebert, K.M., Black, A.E., 2012. Effect of suppression strategies on federal wildland fire expenditures. J. For. 110, 65–73. https://doi.org/10.5849/jof.10-068

Gebert, K.M., Calkin, D.E., Yoder, J., 2007. Estimating suppression expenditures for individual large wildland fires. West. J. Appl. For. 22, 188–196. https://doi.org/10.1093/wjaf/22.3.188

Gillen, B., Snowberg, E., Lariv, L., Yariv, L., 2019. Experimenting with measurement error: Techniques with applications to the caltech cohort study. J. Polit. Econ., NBER Working Paper No. w21517 127, 1826–1863. https://doi.org/10.1086/701681

Glaeser, E.L., Laibson, D.I., Scheinkman, J.A., Soutter, C.L., 2000. Measuring Trust. Q. J. Econ. 115, 811–846. https://doi.org/10.1162/003355300554926

Gneezy, U., Potters, J., 1997. An experiment on risk taking and evaluation periods. Q. J. Econ. 112, 631–645. https://doi.org/10.1162/003355397555217

Goeree, J.K., Holt, C.A., Laury, S.K., 2002. Private costs and public benefits: unraveling the effects of altruism and noisy behavior. J. Public Econ. 83, 255–276. https://doi.org/10.1016/S0047-2727(00)00160-2

Gonzalez-Caban, A., 1984. Costs of firefighting mopup activities.

Government of Alberta, 2021. Population Projections: Alberta and Census Deivisions, 2021–2046.

Government of Alberta, 2018. Workplace incident fatalities accepted by the Workers' Compensation Board in [year] - Open Government [WWW Document]. URL https://open.alberta.ca/publications/6539952 (accessed 1.18.22).

Groesch, D.A., Johnson, J.E., Smith, D.W., 1992. Early vegetative response to wildfire in a table mountain-pitch pine forest. Int. J. Wildl. Fire 2, 177–184. https://doi.org/10.1071/WF9920177

Guimarães, P., Portugal, P., 2010. A simple feasible procedure to fit models with high-dimensional fixed effects. Stata J. 10, 628–649.

Guiso, L., Sapienza, P., Zingales, L., 2018. Time varying risk aversion. J. financ. econ. 128, 403–421. https://doi.org/10.1016/j.jfineco.2018.02.007

Haigh, M.S., List, J.A., 2005. Do professional traders exhibit myopic loss aversion? an experimental analysis. J. Finance 60, 523–534. https://doi.org/10.1111/j.1540-6261.2005.00737.x

Hand, M., Katuwal, H., Calkin, D.E., Thompson, M.P., 2017. The influence of incident management teams on the deployment of wildfire suppression resources. Int. J. Wildl. Fire 26, 615–629. https://doi.org/10.1071/WF16126

Hand, M.S., Gebert, K.M., Liang, J., Calkin, D.E., Thompson, M.P., Zhou, M., 2014. Economics of Wildfire Management - The Development and Application of Suppression Expenditure Models. Springer.

Hand, M.S., Thompson, M.P., Calkin, D.E., 2016. Examining heterogeneity and wildfire management expenditures using spatially and temporally descriptive data. J. For. Econ. 22, 80–102. https://doi.org/10.1016/j.jfe.2016.01.001

Hand, M.S., Wibbenmeyer, M.J., Calkin, D.E., Thompson, M.P., 2015. Risk Preferences, Probability Weighting, and Strategy Tradeoffs in Wildfire Management. Risk Anal. 35, 1876–1891. https://doi.org/10.1111/risa.12457

Hanes, C.C., Wang, X., Jain, P., Parisien, M.A., Little, J.M., Flannigan, M.D., 2019. Fire-regime changes in canada over the last half century. Can. J. For. Res. 49, 256–269. https://doi.org/10.1139/cjfr-2018-0293

Hanna, S.D., Lindamood, S., 2004. An Improved Measure of Risk Aversion. Financ. Couns. Plan. 15, 27–38.

Harrison, G.G.W., Rutström, E.E., Ruström, E., 2008. Risk aversion in the laboratory. Res. Exp. Econ. 12, 41–196. https://doi.org/10.1016/S0193-2306(08)00003-3

Harrison, G.W., Humphrey, S.J., Verschoor, A., 2010. Choice under Uncertainty: Evidence from Ethiopia, India and Uganda. Econ. J. 120, 80–104. https://doi.org/10.1111/J.1468-0297.2009.02303.X

Harrison, G.W., Johnson, E., Mcinnes, M.M., Elisabet Rutström, E., 2005. Risk aversion and incentive effects: Comment. Am. Econ. Rev. 95, 897–901. https://doi.org/10.1257/0002828054201378

Harrison, G.W., List, J.A., 2004. Field experiments. J. Econ. Lit. 42, 1009–1055. https://doi.org/10.1257/0022051043004577

Harrison, G.W., Rutström, E.E., 2009. Expected utility theory and prospect theory: One wedding and a decent funeral. Exp. Econ. 12, 133–158. https://doi.org/10.1007/s10683-008-9203-7

Hey, J.D., Orme, C., 1994a. Investigating generalizations of expected utility theory using experimental data. Econometrica 62, 1291–1326. https://doi.org/10.2307/2951750

Hey, J.D., Orme, C., 1994b. Investigating Generalizations of Expected Utility Theory Using Experimental Data Author(s): John D. Hey and Chris Orme Source: Econometrica 62, 1291–1326.

Hirsch, K.G., Podur, J.J., Janser, R.F., Mcalpine, R.S., Martell, D.L., 2004. Productivity of Ontario initial-attack fire crews: results of an expert-judgement elicitation study. https://doi.org/10.1139/X03-237

Holt, C.A., Laury, S.K., 2002. Risk Aversion and Incentive Effects. Am. Econ. Rev. 92, 1644–1655. https://doi.org/10.1257/000282802762024700

Holtz, D., Zhao, M., Benzell, S.G., Cao, C.Y., Rahimian, M.A., Yang, J., Allen, J., Collis, A., Moehring, A., Sowrirajan, T., Ghosh, D., Zhang, Y., Dhillon, P.S., Nicolaides, C., Eckles, D., Aral, S., 2020. Interdependence and the cost of uncoordinated responses to COVID-19. Proc. Natl. Acad. Sci. U. S. A. 117, 19837–19843. https://doi.org/10.1073/PNAS.2009522117/-/DCSUPPLEMENTAL

Holzmeister, F., 2017. oTree: Ready-made apps for risk preference elicitation methods. J. Behav. Exp. Financ. 16, 33–38. https://doi.org/10.1016/j.jbef.2017.08.003

Holzmeister, F., Stefan, M., 2020. The risk elicitation puzzle revisited: Across-methods (in)consistency? Exp. Econ. 24, 593–616. https://doi.org/10.1007/s10683-020-09674-8

Hope, E.S., McKenney, D.W., Pedlar, J.H., Stocks, B.J., Gauthier, S., 2016. Wildfire Suppression Costs for Canada under a Changing Climate. PLoS One 11, e0157425. https://doi.org/10.1371/JOURNAL.PONE.0157425

Houtman, R.M., Montgomery, C.A., Gagnon, A.R., Calkin, D.E., Dietterich, T.G., McGregor, S., Crowley, M., 2013. Allowing a wildfire to burn: Estimating the effect on future fire suppression costs. Int. J. Wildl. Fire 22, 871–882. https://doi.org/10.1071/WF12157

Jacobson, S., Petrie, R., 2009. Learning from mistakes: What do inconsistent choices over risk tell us? J. Risk Uncertain. 38, 143–158. https://doi.org/10.1007/s11166-009-9063-3

Jianakoplos, N.A., Bernasek, A., 2006. Financial Risk Taking by Age and Birth Cohort. South. Econ. J. 72, 981. https://doi.org/10.2307/20111864

Johnson, E.A., Miyanishi, K. (Eds.), 2001. Forest Fires: Behavior and Ecological Effects. Academic Press.

Johnson, E.A., Miyanishi, K., Bridge, S.R.J., 2001. Wildfire regime in the boreal forest and the idea of suppression and fuel buildup. Conserv. Biol. 15, 1554–1557. https://doi.org/10.1046/j.1523-1739.2001.01005.x

Jung, S., 2015. Does education affect risk aversion? Evidence from the British education reform. Appl. Econ. 47, 2924–2938. https://doi.org/10.1080/00036846.2015.1011313

Kachelmeier, S.J., Shehata, M., 1992. Examining Risk Preferences Under High Monetary Incentives: Experimental Evidence from the People’s Republic of China 82, 1120–1141.

Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. Econometrica 47, 263–292. https://doi.org/10.2307/1914185

Kellen, D., Mata, R., Davis-Stober, C.P., 2017. Individual classification of strong risk attitudes: An application across lottery types and age groups. Psychon. Bull. Rev. 24, 1341–1349. https://doi.org/10.3758/S13423-016-1212-5/TABLES/4

Kimmerer, R.W., Lake, F.K., 2001. Maintaining the Mosaic: The role of indigenous burning in land management. J. For. 99, 36–41. https://doi.org/10.1093/jof/99.11.36

KPMG, 2017. May 2016 Wood Buffalo Wildfire, Post-Incident Assessment Report.

Krčál, O., Staněk, R., Slanicay, M., 2019. Made for the job or by the job? A lab-in-the-field experiment with firefighters. Res. Econ. 73, 271–276. https://doi.org/10.1016/j.rie.2019.08.001

Kroll, Y., Davidovitz, L., 2003. Inequality aversion versus risk aversion. Economica 70, 19–29. https://doi.org/10.1111/1468-0335.t01-1-00269

Kuhnen, C.M., Knutson, B., 2005. The Neural Basis of Financial Risk Taking. Neuron 47, 763–770. https://doi.org/10.1016/J.NEURON.2005.08.008

Labossière, L.M.M., McGee, T.K., 2017. Innovative wildfire mitigation by municipal governments: Two case studies in Western Canada. Int. J. Disaster Risk Reduct. 22, 204–210. https://doi.org/10.1016/j.ijdrr.2017.03.009

Lake, F.K., Christianson, A.C., 2019. Indigenous Fire Stewardship, in: Encyclopedia of Wildfires and Wildland-Urban Interface (WUI) Fires. pp. 714–722. https://doi.org/10.1007/978-3-319-52090-2_225

Lamoureux, M., Bellefontaine, M., 2016. Preliminary cost of Fort McMurray fire estimated at $615 million | CBC News [WWW Document]. CBC News. URL https://www.cbc.ca/news/canada/edmonton/preliminary-cost-of-fort-mcmurray-fire-estimated-at-615-million-1.3640935 (accessed 11.4.21).

Laury, S.K., McInnes, M.M., Swarthout, T.J., 2012. Avoiding the curves: Direct elicitation of time preferences. J. Risk Uncertain. 2012 443 44, 181–217. https://doi.org/10.1007/S11166-012-9144-6

Lawson, B.D., Armitage, O.B., 2008. Weather Guide for the Canadian Forest Fire Danger Rating System. Edmonton.

Lejuez, C.W., Richards, J.B., Read, J.P., Kahler, C.W., Ramsey, S.E., Stuart, G.L., Strong, D.R., Brown, R.A., 2002. Evaluation of a behavioral measure of risk taking: The balloon analogue risk task (BART). J. Exp. Psychol. Appl. 8, 75–84. https://doi.org/10.1037/1076-898X.8.2.75

Levy, D.J., Glimcher, P.W., 2011. Comparing Apples and Oranges: Using Reward-Specific and Reward-General Subjective Value Representation in the Brain. J. Neurosci. 31, 14693–14707. https://doi.org/10.1523/JNEUROSCI.2218-11.2011

Levy, H., Levy, M., 2021. Prospect theory, constant relative risk aversion, and the investment horizon. PLoS One 16, e0248904. https://doi.org/10.1371/JOURNAL.PONE.0248904

Liang, J., Calkin, D.E., Gebert, K.M., Venn, T.J., Silverstein, R.P., Liang, J., Calkin, D.E., Gebert, K.M., Venn, T.J., Silverstein, R.P., 2008. Factors influencing large wildland fire suppression expenditures. Int. J. Wildl. Fire 17, 650–659. https://doi.org/10.1071/WF07010

List, J.A., 2004. Neoclassical Theory Versus Prospect Theory: Evidence from the Marketplace. Econometrica 72, 615–625. https://doi.org/10.1111/J.1468-0262.2004.00502.X

Lönnqvist, J.E., Verkasalo, M., Walkowitz, G., Wichardt, P.C., 2015. Measuring individual risk attitudes in the lab: Task or ask? An empirical comparison. J. Econ. Behav. Organ. 119, 254–266. https://doi.org/10.1016/j.jebo.2015.08.003

Loomes, G., Sugden, R., 1982. Regret theory: An alternative theory of rational choice under uncertainty. Econ. J. 92, 805–824. https://doi.org/10.2307/2232669

MacVicar, A., 2019. Alberta Wildfire cuts rappel crews, detection tower staffing and air unit following provincial budget | Globalnews.ca. Glob. News.

Maguire, L.A., Albright, E.A., 2005. Can behavioral decision theory explain risk-averse fire management decisions? For. Ecol. Manage. 211, 47–58. https://doi.org/10.1016/J.FORECO.2005.01.027

Markowitz, H., 1952. Portfolio Selection. J. Finance 7, 77. https://doi.org/10.2307/2975974

Mata, R., Frey, R., Richter, D., Schupp, J., Hertwig, R., 2018. Risk preference: A view from psychology. J. Econ. Perspect. 32, 155–172. https://doi.org/10.1257/jep.32.2.155

McFadden, D., 1974. Conditional logit analysis of qualitative choice behavior, in: Zarembka, P. (Ed.), Frontiers in Econometrics. Academic Press, pp. 105–142.

McFarlane, B.L., McGee, T.K., Faulkner, H., 2011. Complexity of homeowner wildfire risk mitigation: An integration of hazard theories. Int. J. Wildl. Fire 20, 921–931. https://doi.org/10.1071/WF10096

McGee, T., McFarlane, B., Tymstra, C., 2015. Wildfire: A Canadian Perspective, in: Wildfire Hazards, Risks, and Disasters. pp. 35–58. https://doi.org/10.1016/B978-0-12-410434-1.00003-8

MNP LLP, 2020. Spring 2019 Wildfire Review Final Report - November 2020. Edmonton.

MNP LLP, 2017. A review of the 2016 Horse River wildfire: Alberta Agriculture and Forestry Preparedness and Response.

MNP LLP, 2016. Review of Alberta Agriculture and Forestry's Wildfire Management Program and the 2015 Fire Season. Volume 2: Detailed Report. Edmonton.

Mohr, P.N.C., Nagel, I.E., 2010. Variability in Brain Activity as an Individual Difference Measure in Neuroscience? J. Neurosci. 30, 7755–7757. https://doi.org/10.1523/JNEUROSCI.1560-10.2010

National Wildfire Coordinating Group, 2019. NWCG Standards for Fire Weather Stations.

Nicholson, N., Soane, E., Fenton-O'Creevy, M., Willman, P., 2005. Personality and domain-specific risk taking. J. Risk Res. 8, 157–176. https://doi.org/10.1080/1366987032000123856

Odean, T., 1999. Do Investors Trade Too Much? Am. Econ. Rev. 89, 1279–1298. https://doi.org/10.1257/AER.89.5.1279

Odean, T., 1998. Volume, Volatility, Price, and Profit When All Traders Are Above Average. J. Finance 53, 1887–1934. https://doi.org/10.1111/0022-1082.00078

Owen, B., 2021. Thousands could die and costs could reach billions if wildfire response doesn't change, scientists say | CBC News [WWW Document]. CBC News. URL https://www.cbc.ca/news/canada/british-columbia/wildfire-losses-1.6122673 (accessed 11.4.21).

Paulus, M.P., Rogalsky, C., Simmons, A., Feinstein, J.S., Stein, M.B., 2003. Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism. Neuroimage 19, 1439–1448. https://doi.org/10.1016/S1053-8119(03)00251-9

Pausas, J.G., Keeley, J.E., 2019. Wildfires as an ecosystem service. Front. Ecol. Environ. 17, 289–295. https://doi.org/10.1002/fee.2044

Pedroni, A., Frey, R., Bruhin, A., Dutilh, G., Hertwig, R., Rieskamp, J., A Pedroni, R Frey, A Bruhin, G Dutilh, R Hertwig, J.R., 2017. The risk elicitation puzzle. Nat. Hum. Behav. 1, 803–809. https://doi.org/10.1038/s41562-017-0219-x

Peters, E., Hess, T.M., Västfjäll, D., Auman, C., 2016. Adult Age Differences in Dual Information Processes: Implications for the Role of Affective and Deliberative Processes in Older Adults' Decision Making: https://doi.org/10.1111/j.1745-6916.2007.00025.x 2, 1–23. https://doi.org/10.1111/J.1745-6916.2007.00025.X

Peterson, R.L., 2007. Affect and Financial Decision-Making: How Neuroscience Can Inform Market Participants. http://dx.doi.org/10.1080/15427560701377448 8, 70–78. https://doi.org/10.1080/15427560701377448

PetroLMI, ENFORM, 2015. Rotation, Not Relocating: Alberta's Oil and Gas Rotational Workforce.

Plucinski, M.P., Mccaw, W.L., Gould, J.S., Wotton, B.M., 2014. Predicting the number of daily human-caused bushfires to assist suppression planning in south-west Western Australia. Int. J. Wildl. Fire 23, 520–531. https://doi.org/10.1071/WF13090

Pope, D.G., Schweitzer, M.E., 2011. Is Tiger Woods Loss Averse? Persistent Bias in the Face of Experience, Competition, and High Stakes. Am. Econ. Rev. 101, 129–57. https://doi.org/10.1257/AER.101.1.129

Popyk, T., 2021. Cost of fighting B.C. wildfires already above $95 million for 2021 — before peak fire season begins | CBC News [WWW Document]. CBC News. URL https://www.cbc.ca/news/canada/british-columbia/bc-wildfire-service-costs-2021-1.6092516 (accessed 11.4.21).

Post, T., van den Assem, M.J., Baltussen, G., Thaler, R.H., 2008. Deal or No Deal? Decision Making under Risk in a Large-Payoff Game Show. Am. Econ. Rev. 98, 38–71. https://doi.org/10.1257/AER.98.1.38

Pyne, S.J., 2008. Awful Splendour: A Fire History of Canada. UBC Press, Vancouver.

Regenwetter, M., Dana, J., Davis-Stober, C.P., 2011. Transitivity of Preferences. Psychol. Rev. 118, 42–56. https://doi.org/10.1037/a0021150

Reynaud, A., Couture, S., 2012. Stability of risk preference measures: Results from a field experiment on French farmers. Theory Decis. 73, 203–221. https://doi.org/10.1007/s11238-012-9296-5

Rhemtulla, J.M., Hall, R.J., Higgs, E.S., Macdonald, S.E., 2011. Eighty years of change: vegetation in the montane ecoregion of Jasper National Park, Alberta, Canada. https://doi.org/10.1139/x02-112 32, 2010–2021. https://doi.org/10.1139/X02-112

Riddel, M., Kolstoe, S., 2013. Heterogeneity in life-duration preferences: Are risky recreationists really more risk loving? J. Risk Uncertain. 46, 191–213. https://doi.org/10.1007/s11166-013-9161-0

Riley, K.L., Thompson, M.P., Scott, J.H., Gilbertson-Day, J.W., 2018. A model-based framework to evaluate alternative wildfire suppression strategies. Resources 7, 1–26. https://doi.org/10.3390/resources7010004

Robinne, F.-N., Parisien, M.-A., Flannigan, M., Robinne, F.-N., Parisien, M.-A., Flannigan, M., 2016. Anthropogenic influence on wildfire activity in Alberta, Canada. Int. J. Wildl. Fire 25, 1131–1143. https://doi.org/10.1071/WF16058

Ruggeri, K., Alí, S., Berge, M.L., Bertoldo, G., Bjørndal, L.D., Cortijos-Bernabeu, A., Davison, C., Demić, E., Esteban-Serna, C., Friedemann, M., Gibson, S.P., Jarke, H., Karakasheva, R., Khorrami, P.R., Kveder, J., Andersen, T.L., Lofthus, I.S., McGill, L., Nieto, A.E., Pérez, J., Quail, S.K., Rutherford, C., Tavera, F.L., Tomat, N., Reyn, C. Van, Većkalov, B., Wang, K., Yosifova, A., Papa, F., Rubaltelli, E., Linden, S. van der, Folke, T., 2020. Replicating patterns of prospect theory for decision under risk. Nat. Hum. Behav. 2020 46 4, 622–633. https://doi.org/10.1038/s41562-020-0886-x

Samanez-Larkin, G.R., Kuhnen, C.M., Yoo, D.J., Knutson, B., 2010. Variability in Nucleus Accumbens Activity Mediates Age-Related Suboptimal Financial Risk Taking. J. Neurosci. 30, 1426–1434. https://doi.org/10.1523/JNEUROSCI.4902-09.2010

Sauter, P., Hermann, D., Mußhoff, O., 2015. Risk attitudes of foresters, farmers and students: An experimental multimethod comparison. Diskussionbeitrag 1514.

Schonberg, T., Fox, C.R., Poldrack, R.A., 2011. Mind the gap: bridging economic and naturalistic risk-taking with cognitive neuroscience. Trends Cogn. Sci. 15, 11–19. https://doi.org/10.1016/J.TICS.2010.10.002

Schroeder, D., 2010. Fire behaviour in thinned jack pine : two case studies of FireSmart treatments in Canada’s Northwest Territories. FPInnovations Advant. Rep. 12, 1–11.

Snyder, G.W., 1999. Strategic Holistic Integrated Planning for the Future: Fire Protection in the Urban/ Rural/Wildland Interface (URWIN). Proc. Symp. Fire Econ. Planning, Policy December, 159–170.

SOEP, Wagner, G.G., Frick, J.R., Schupp, J., 2007. The German Socio-Economic Panel Study (SOEP): Scope, Evolution and Enhancements, SOEPpapers on Multidisciplinary Panel Data Research, No. 1. Berlin.

Steele, T.W., Stier, J.C., 1998. An Economic Evaluation of Public and Organized Wildfire Detection in Wisconsin. Int. J. Wildl. Fire 8, 205–215.

Stocks, B.J., Fosberg, M.A., Lynham, T.J., Mearns, L., Wotton, B.M., Yang, Q., Jin, J.Z., Lawrence, K., Hartley, G.R., Mason, J.A., McKenney, D.W., 1998. Climate change and forest fire potential in Russian and Canadian boreal forests. Clim. Change 38, 1–13.

https://doi.org/10.1023/A:1005306001055

Stocks, B.J., Martell, D.L., 2016. Forest fire management expenditures in Canada: 1970-2013. For. Chron. 92, 298–306. https://doi.org/10.5558/tfc2016-056

Tanaka, T., Camerer, C.F., Nguyen, Q., 2010. Risk and time preferences: Linking experimental and household survey data from Vietnam. Am. Econ. Rev. 100, 3–25. https://doi.org/10.1007/978-4-431-55402-8_1

Thagard, P., 2005. Mind: Introduction to cognitive science. MIT Press, Cambridge, Mass.

Trepel, C., Fox, C.R., Poldrack, R.A., 2005. Prospect theory on the brain? Toward a cognitive neuroscience of decision under risk. Cogn. Brain Res. 23, 34–50. https://doi.org/10.1016/J.COGBRAINRES.2005.01.016

Tucker, E., 2017. Firefighter who died battling Alberta wildfire remembered as 'great father' | Globalnews.ca. Glob. News.

Tversky, A., Kahneman, D., 1992. Advances in prospect theory: Cumulative representation of uncertainty. J. Risk Uncertain. 5, 297–323. https://doi.org/10.1007/BF00122574

Tymstra, C., Jain, P., Flannigan, M.D., Tymstra, C., Jain, P., Flannigan, M.D., 2021. Characterisation of initial fire weather conditions for large spring wildfires in Alberta, Canada. Int. J. Wildl. Fire 30, 823–835. https://doi.org/10.1071/WF21045

Tymstra, C., Stocks, B.J., Cai, X., Flannigan, M.D., 2020. Wildfire management in Canada: Review, challenges and opportunities. Prog. Disaster Sci. 5, 100045. https://doi.org/10.1016/j.pdisas.2019.100045

Vasilakos, Christos, Kostas, A.E., Ae, K., Hatzopoulos, John, Vasilakos, C, Hatzopoulos, J, Matsinos, Á.I., Kalabokidis, K., 2009. Identifying wildland fire ignition factors through sensitivity analysis of a neural network 50, 125–143. https://doi.org/10.1007/s11069-008-9326-3

Venn, T.J., Calkin, D.E., 2011. Accommodating non-market values in evaluation of wildfire management in the United States: Challenges and opportunities. Int. J. Wildl. Fire 20, 327–339. https://doi.org/10.1071/WF09095

Verstraeten, B.S.E., Elgbeili, G., Hyde, A., King, S., Olson, D.M., 2021. Maternal Mental Health after a Wildfire: Effects of Social Support in the Fort McMurray Wood Buffalo Study. Can. J. Psychiatry 66, 710–718. https://doi.org/10.1177/0706743720970859

Von Neumann, J., Morgenstern, O., 1947. Theory of games and economic behavior, 2nd ed. Princeton University Press.

Weber, E.U., Blais, A.R., Betz, N.E., 2002. A domain-specific risk-attitude scale: measuring risk perceptions and risk behaviors. J. Behav. Decis. Mak. 15, 263–290. https://doi.org/10.1002/bdm.414

Wibbenmeyer, M.J., Hand, M.S., Calkin, D.E., Venn, T.J., Thompson, M.P., 2013. Risk preferences in strategic wildfire decision making: A choice experiment with U.S. wildfire managers. Risk Anal. 33, 1021–1037. https://doi.org/10.1111/j.1539-6924.2012.01894.x

Wilcox, N.T., 2008. Stochastic models for binary discrete choice under risk: A critical primer and econometric comparison. Res. Exp. Econ. 12, 197–292. https://doi.org/10.1016/S0193-2306(08)00004-5

Wilson, R.S., Winter, P.L., Maguire, L.A., Ascher, T., 2011. Managing wildfire events: Risk-based decision making among a group of federal fire managers. Risk Anal. 31, 805–818. https://doi.org/10.1111/j.1539-6924.2010.01534.x

Wotton, B.M., Flannigan, M.D., Marshall, G.A., 2017. Potential climate change impacts on fire intensity and key wildfire suppression thresholds in Canada. Environ. Res. Lett. 12, 095003. https://doi.org/10.1088/1748-9326/aa7e6e

Yoder, J., Gebert, K., 2012. An econometric model for ex ante prediction of wildfire suppression costs. J. For. Econ. 18, 76–89. https://doi.org/10.1016/j.jfe.2011.10.003

Appendix A: Maps of the FPA and FAs

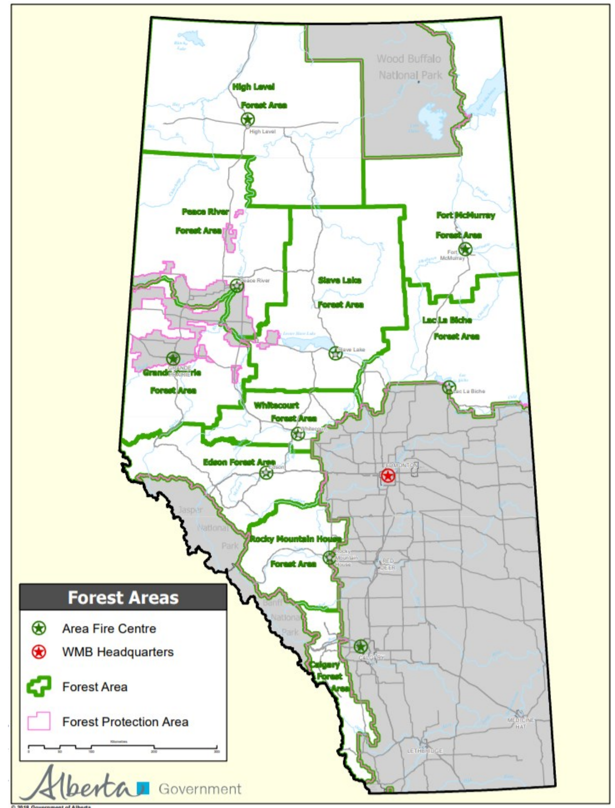

Figure A.1 Forest Protection Area (FPA) in Alberta
Alberta Wildfire is responsible for the suppression of all fires within the FPA.

Image source: https://wildfire.alberta.ca/resources/maps-data/documents/ForestProtectionAreaMap-May03-2017.pdf

Figure A.2 Forest Areas (FAs) in Alberta
FAs are administrative regions of Alberta Wildfire, and are centrally coordinated from headquarters in Edmonton.

Image source: https://wildfire.alberta.ca/resources/maps-data/documents/ForestAreas-Oct09-2018.pdf

Appendix B: Auxiliary figures

In this appendix, Figure B.1 examines the relationship between *Reporting delay* (hours between fire ignition and reporting) and *Report-Extinguished delay* (hours between reporting and extinguishment). Longer response times generally lead to longer fire durations. The magnitude and significance of this relationship is reported in Table C.2 of Appendix C: Auxiliary regressions.

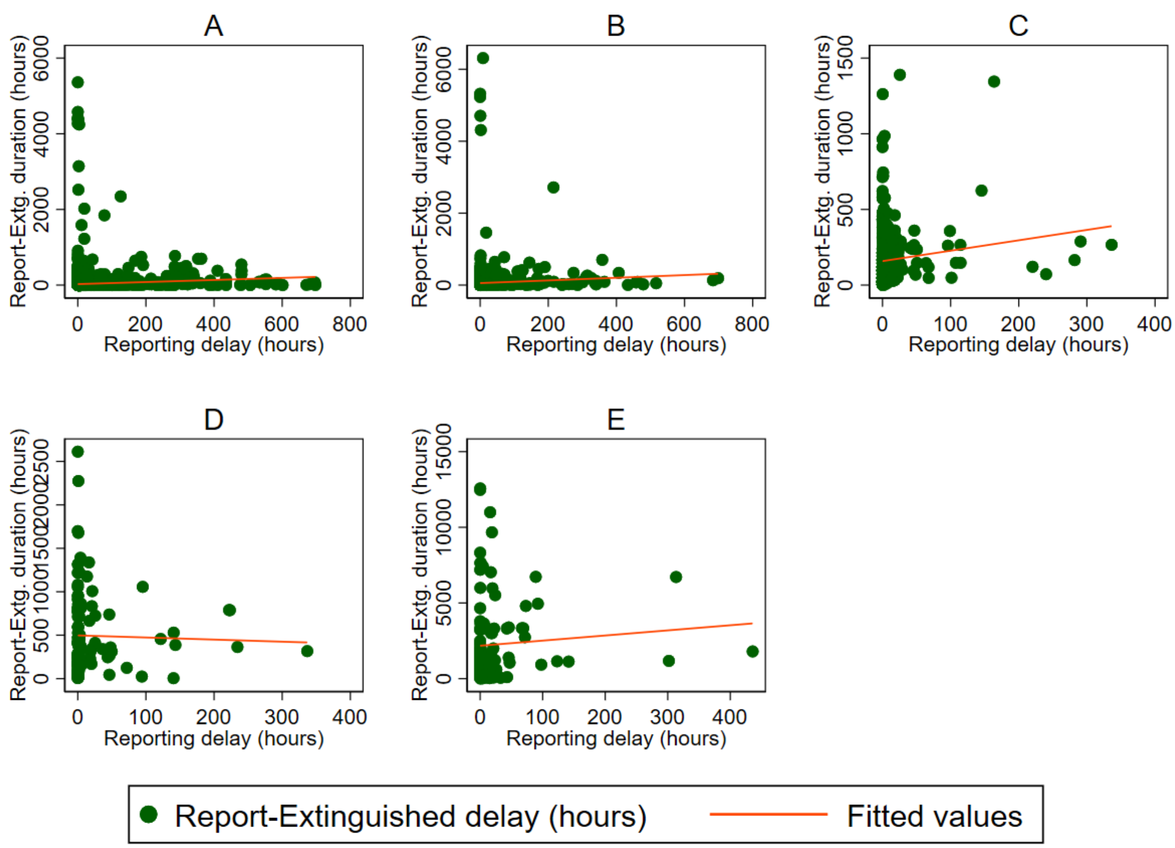

Figure B.1 *Report-extinguished phase duration* and *Reporting delay* by size class *(excluding observations with ≥ 30-day* Reporting delay*)*

Appendix C: Auxiliary regressions

Regression models in this appendix examine auxiliary specifications that have been mentioned in the main text of Chapter 2.

Table C.1 reports the effect of *Reporting delay* on *Log Cost*, displayed in Figure 2.4. The relationship between *Reporting delay* and *Log Cost* is significant size classes A and B.

Table C.1 OLS regression on the effect of *Reporting delay* on *Log Cost* by size class

	A	B	C	D	E
Start-Report delay (hours)	0.003***	0.003***	0.004	-0.004	0.001
	(0.001)	(0.001)	(0.003)	(0.003)	(0.004)
N	3,849	1,878	372	116	137
R-squared	0.006	0.006	0.007	0.019	0.001

Excluding observations with ≥ 30-day Reporting delay
Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table C.2 reports the effect of *Reporting delay* on *Report-extinguished phase duration*, displayed in Figure B.1. The relationship between *Reporting delay* on *Report-extinguished phase duration* is significant for fires in size classes A, B and C.

Table C.2 OLS regression on the effect of *Reporting delay* on *Report-extinguished phase duration* by size class

	A	B	C	D	E
Start-Report delay (hours)	0.273***	0.370***	0.688**	-0.242	3.385
	(0.056)	(0.129)	(0.267)	(0.799)	(4.021)
N	4,529	1,917	376	120	137
R-squared	0.005	0.004	0.017	0.001	0.005

Excluding observations with ≥ 30-day Reporting delay.
Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table C.3 reports the expenditure model, including observations of wildfires that start on military land. The inclusion of these 30 observations do not change results substantially.

Table C.3 OLS regression models, separated by wildfire size class (includes *Military land*)

	A	B	C	D	E
Fire environment					
Temperature (°C)	0.030**	0.034**	0.058**	0.074	0.380***
	(0.010)	(0.012)	(0.023)	(0.120)	(0.063)
Wind speed (km/h)	0.001	0.017*	0.004	0.055**	0.000
	(0.006)	(0.008)	(0.009)	(0.018)	(0.026)
Rain (mm)	-0.024*	-0.005	-0.053**	0.009	-0.138***
	(0.012)	(0.014)	(0.017)	(0.094)	(0.025)
Relative Humidity (%)	-0.007	-0.006	0.034***	0.045	0.063
	(0.008)	(0.005)	(0.008)	(0.038)	(0.061)
Fuel: Timberslash	1.275***	1.184***	1.225***	0.623**	-0.669
	(0.156)	(0.071)	(0.347)	(0.247)	(1.354)
Crown fire	-0.171	0.486***	0.101	0.229	0.248
	(0.277)	(0.133)	(0.168)	(0.123)	(0.376)
South Aspect (true south)	0.023	0.169*	0.088	-0.079	1.110*
	(0.130)	(0.077)	(0.246)	(0.364)	(0.534)
High elevation	0.084	0.207	0.539	0.000	0.000
	(0.212)	(0.158)	(0.362)	(.)	(.)
Elevation difference (m)	0.000	0.148***	0.015	-0.005	0.007
	(.)	(0.026)	(0.008)	(0.003)	(0.004)
Lake/River within 3km	-0.136	-0.059	0.114	0.197	0.184
	(0.092)	(0.077)	(0.232)	(0.406)	(1.087)
Operation					
Resource availability	-0.000	-0.002	-0.004	-0.002	-0.024**
	(0.002)	(0.002)	(0.005)	(0.007)	(0.009)
Delay (hours)	0.002***	0.002***	0.001	-0.011	0.006
	(0.000)	(0.000)	(0.002)	(0.007)	(0.005)
Strategic delay	-0.215*	-0.444***	-1.473**	-1.358***	-0.914
	(0.099)	(0.104)	(0.570)	(0.157)	(0.722)
Provincial land	0.277	0.397*	0.755*	0.055	0.000
	(0.200)	(0.196)	(0.360)	(0.163)	(.)
Indigenous land	0.315*	-0.001	0.097	-0.590	0.000
	(0.172)	(0.172)	(0.447)	(0.491)	(.)
Military land	0.414*	0.925***	1.987***	1.226	0.000
	(0.220)	(0.276)	(0.570)	(0.721)	(.)
Values-at-risk (within 3 km)					
Community	-0.449***	-0.577***	0.327	1.776***	1.671
	(0.086)	(0.172)	(0.489)	(0.137)	(1.071)
Park	0.173	0.110	1.165*	0.000	0.874
	(0.193)	(0.241)	(0.630)	(.)	(1.567)
Power generation	0.393	-0.193	1.190***	3.155***	0.000
	(0.287)	(0.419)	(0.326)	(0.426)	(.)
Road	-0.573***	-0.319***	-0.399**	-0.116	0.227
	(0.076)	(0.089)	(0.145)	(0.225)	(0.875)
N	2,973	1,658	326	87	81
R-squared	0.431	0.522	0.537	0.616	0.564

All models include corporate region, year, and month of the year fixed effects. Standard errors clustered at the corporate region are in parenthesis. * p<0.10, ** p<0.05, *** p<0.01

Table C.4 and Table C.5 compare estimate results between Ordinary Least Squares (OLS) regression modelling and double/debiased machine learning (DML). For models on both *Log cost* and *Reporting delay* dependent variables, Estimated coefficients between methods are similar, although DML estimates are highly significant throughout size classes A, B, C. Due to small sample sizes *(N<90)*, we do not estimate DML models for fires D and E.

Table C.4 OLS and DML: Estimates of *Reporting delay* on *Log cost*

A	A	B	C
OLS			
$\hat{\beta}$	0.00245***	0.00276***	0.00241
	(0.00054)	(0.00090)	(0.00230)
N	2,965	1,645	324
DML			
$\hat{\beta}$	0.00256***	0.00243***	0.00284***
	(5.2423e-06)	(9.4944e-06)	(3.1415e-05)
N	2,965	1,645	324

Size classes: A: 0 to 0.1 ha; B: >0.1 ha to 4 ha; C: > 4 ha to 40 ha
DML estimates not available for size class D and E due to small sample sizes.
Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Table C.5 OLS and DML: Estimates of *Reporting delay* on *Fire Duration*

	A	B	C
OLS			
$\hat{\beta}$	0.11271***	0.11549	0.53221*
	(0.04042)	(0.10045)	(0.30404)
N	3,247	1,668	327
DML			
$\hat{\beta}$	0.11390***	0.11707***	0.71173***
	(2.8150e-04)	(7.3472e-04)	(8.3587e-03)
N	3,247	1,668	327

Size classes: A: 0 to 0.1 ha; B: >0.1 ha to 4 ha; C: > 4 ha to 40 ha
DML estimates not available for size class D and E due to small sample sizes.
Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Appendix D: Supplementary figures on experiment choices

Figures in this appendix show how the distribution of choices throughout the four incentivized tasks vary by cohort and gender (Figure D.1) and by gender and cohort (Figure D.2). The cohort and gender effect have been explored in further detail through structure modelling in Sections 3.6.2 and 3.6.3.

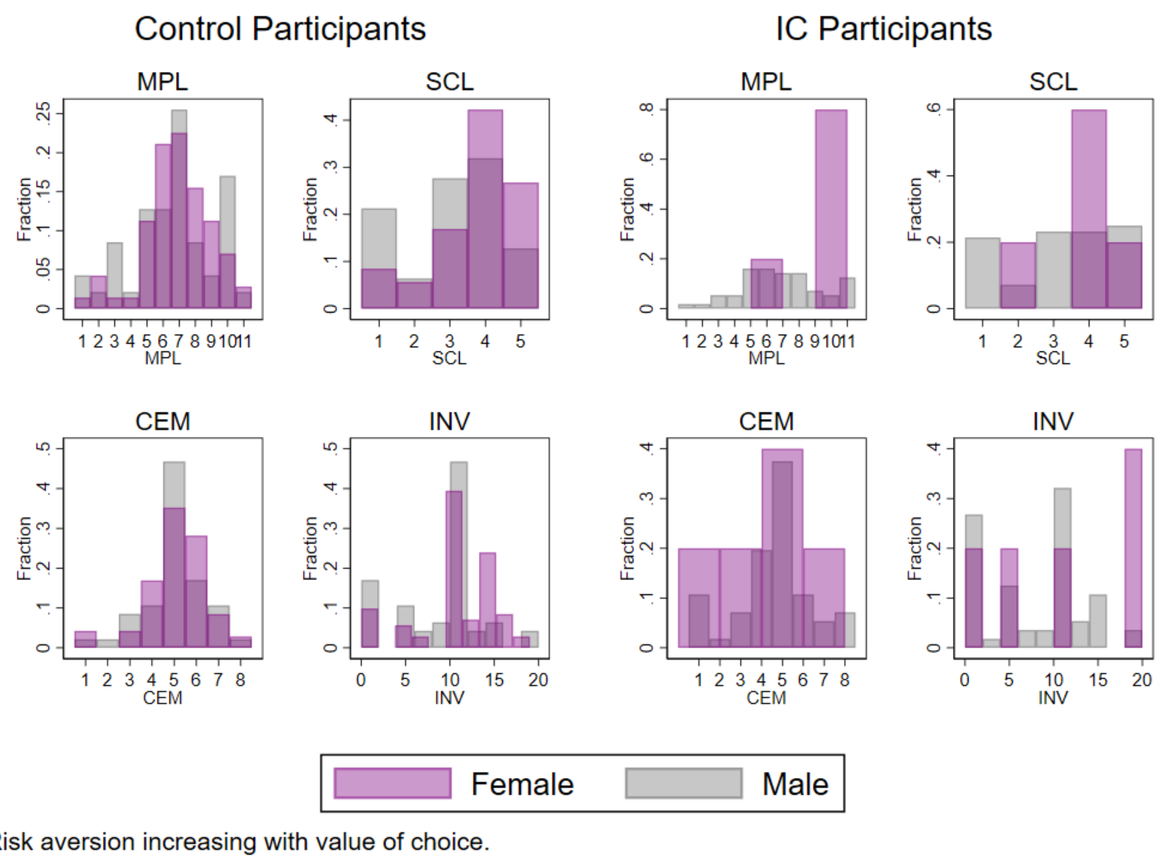

Figure D.1 Graphical representation of Table 3.4
Mean choice across four tasks, compared by cohort and gender.

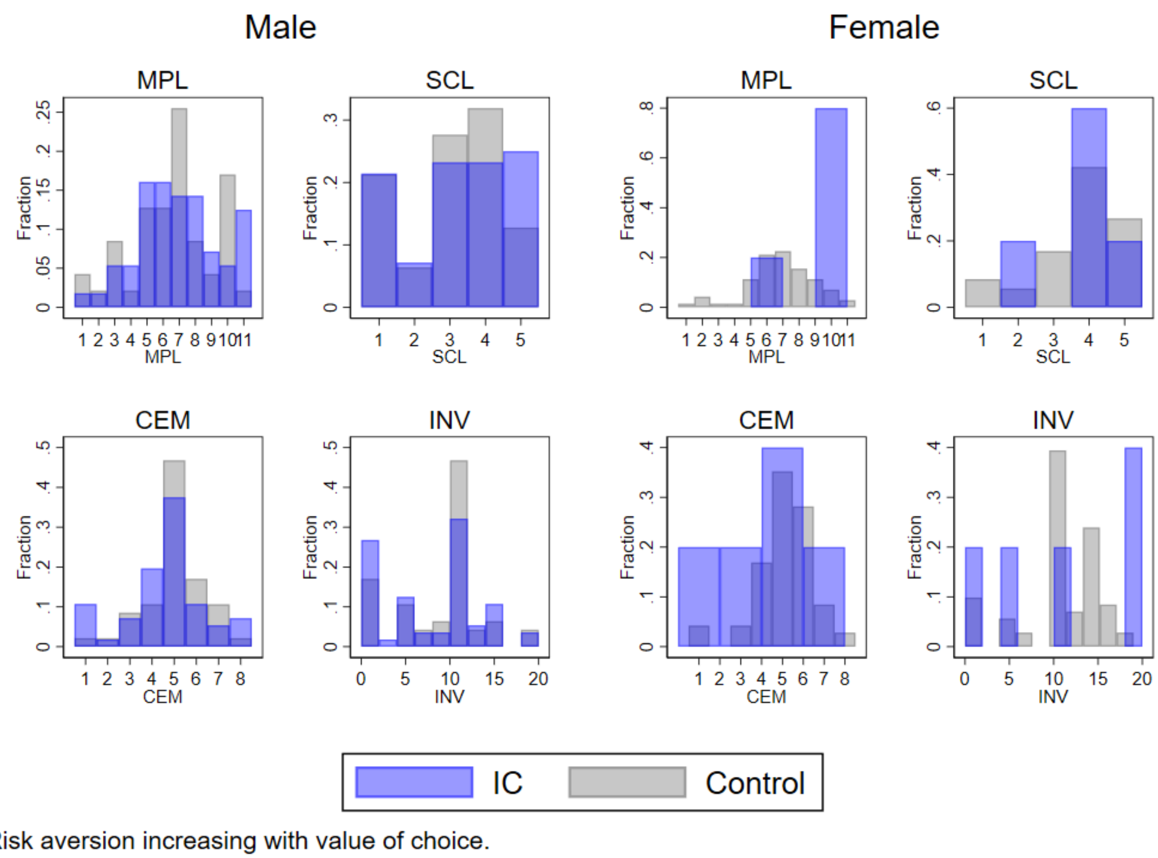

Figure D.2 Graphical representation of Table 3.5
Mean choice across four tasks, compared by gender and cohort.

Appendix E. Inconsistent players in MPL task

This appendix examines behaviour of players that are whose choices in the Multiple Price List task (MPL) are inconsistent with expected utility theory. Two forms of inconsistency may arise in the completion of a Multiple Price List task: multiple switches and the choice of a stochastically dominated option. As our specification of the MPL instructs the participant to select a switching point rather than individual choices between paired lotteries, we avoid the first form of inconsistency. However, some participants in our experiment are observed to violate the second form of consistency, through selection of the dominated option of the last line.

Line	**Option A**	**Option B**
10	100% probability of $20.00 0% probability of $16.00	100% probability of $38.50 0% probability of $1.00

In Line 10 (the last line), both Options guarantee the high payoff with certainty. The stochastically dominant choice is *Option B*, which has a larger value high payoff ($38.50) than that of *Option A* ($16.00). Nonetheless, some participants opt for *Option A* throughout the task. The issue of choice inconsistency has generally been addressed through one of two methods: counting the number of risky choices/safe choices, or, recognizing that inconsistent participants violate a rational player's decision framework, omitting these inconsistent participants from analysis (for a review, see: Filippin and Crosetto, 2016). Table E.1 below is summary on the prevalence of inconsistency among our participants.

Table E.1 Prevalence of MPL inconsistent decision (all participants)

MPL decision	Count by cohort			*Proportion by cohort*		
	All	Control	IC	*All*	*Control*	*IC*
Consistent	167	115	52	*0.93*	*0.97*	*0.85*
Inconsistent	12	3	9	*0.07*	*0.03*	*0.15*

We cannot ignore the possibility that a participant's selection of *Option A* in Line 10 could indicate: a) the participant struggles with understanding the MPL task, or worse, that b) she is unfocused in this task, and perhaps in other parts of the experiment. While we include all participants in analyses in the main text, in Table E.2 we explore which characteristics (Incident Commander, gender, university education, family, age) will make more likely to behave inconsistently, using a linear probability model.

Table E.2 Propensity of participants playing inconsistently in MPL (all participants)

	Inconsistent
IC	0.048
	(0.066)
Female	0.053
	(0.044)
University	-0.031
	(0.040)
Family	0.052
	(0.052)
Age	0.004
	(0.003)
Constant	-0.093
	(0.072)
N	171

Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Inconsistent players make up a larger proportion of the IC cohort (Table E.1), however, after controlling for the same characteristics in all previous analyses, we do not find evidence that participants of either cohort, gender, or with other observable characteristics are more likely to play inconsistently in MPL.

Appendix F: Investment Game: sensitivity of bins

This appendix examines the sensitivity of our results to the selection of bin sizes for choices in the Investment Game (INV) along a continuous scale. In INV, participants were given an endowment of $20, and tasked with choosing how much to invest in a lottery with high and low payoffs, and how much to keep as safe payments. Participants were able to choose to keep as safe payment any amount between $0.00 to $20.00, in units of cents. In our analysis of this task, we have rounded to choices to the nearest $2.00, so that the implied risk parameters of the task can be comparable to those of the other three elicitation methods. This makes for 11 bins, including a separate bin for choices to keep the entire endowment ($20.00) as safe payment. The decision to rounding the investment choice to the $2.00 level follows Crosetto and Filippin (2016). In order to test the sensitivity of estimates to rounding, below we will compare estimates of the heterogeneous Expected Utility Theory model with 11 bins (rounding to the nearest $2.00), against 21 bins ($1.00), 41 bins ($0.50), and 201 bins ($0.10).

Results are reported in Table F.1, and the values of estimates across bin sizes is compared visually in Figure F.1.

Table F.1 Estimates of risk and noise parameters (φ, σ) as a function of characteristics in INV, by bin size (all participants)

	11	21	41	201
φ				
IC	0.033	0.031	0.029	0.029
	(0.043)	(0.042)	(0.042)	(0.042)
Age	-0.003*	-0.003**	-0.003**	-0.003**
	(0.002)	(0.002)	(0.002)	(0.002)
Female	0.085**	0.076**	0.075**	0.073**
	(0.038)	(0.034)	(0.033)	(0.033)
Family	0.000	-0.001	-0.000	-0.000
	(0.030)	(0.030)	(0.030)	(0.030)
University	-0.009	-0.008	-0.009	-0.009
	(0.027)	(0.026)	(0.026)	(0.026)
Constant	0.290***	0.282***	0.279***	0.276***
	(0.051)	(0.046)	(0.046)	(0.045)
σ				
IC	0.066	0.035	0.017	0.004
	(0.120)	(0.052)	(0.027)	(0.006)
Age	-0.001	-0.000	-0.000	-0.000
	(0.003)	(0.001)	(0.001)	(0.000)
Female	0.015	0.001	0.001	0.000
	(0.043)	(0.020)	(0.010)	(0.002)
Family	0.036	0.019	0.010	0.002
	(0.059)	(0.027)	(0.014)	(0.003)
University	0.056	0.027	0.014	0.003
	(0.071)	(0.032)	(0.017)	(0.003)
Constant	0.173***	0.087***	0.044***	0.009***
	(0.061)	(0.027)	(0.014)	(0.003)
N	1710	3420	6840	34200
N_clust	171	171	171	171

Robust standard errors clustered at the subject level in reported parentheses.
* p<0.10, ** p<0.05, *** p<0.01

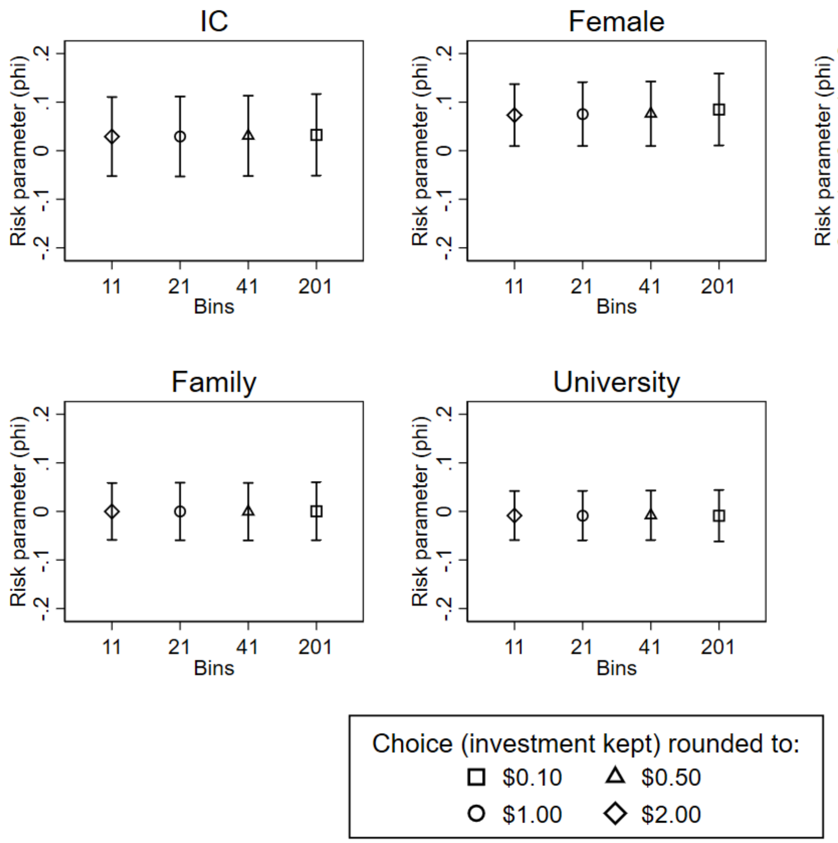

Figure F.1 Comparing estimates on risk parameter (φ) on characteristics in INV across bin sizes (all participants)

As observed in Table F.1, the only significant change is that at 21 or more bins (rounding to \$1.00 or lower), the negative effect of age on risk aversion increases in significance level ($p < 0.10$ to $p < 0.05$). Overall, Table F.1 and Figure F.1 demonstrate that, regardless whether INV choice data is rounded to \$2.00 or to \$0.10, the effects of characteristics on the estimated risk parameter (φ) remains relatively consistent.

Appendix G: CRRA risk parameters for experiment choices

In this appendix, Table G.1 supplements the discussion in Section 3.6.4 on implied Constant Relative Risk Aversion (CRRA) risk parameters and between-task consistency.

Each choice a participant makes implies a parameter of risk preference, which ranges from risk aversion $(\varphi > 0)$ to risk neutral $(\varphi = 0)$ to risk seeking $(\varphi < 0)$. Assuming that all participants' utility for experiment incentives are characterized by the CRRA utility function (as we have in Section 3.4.2), the choice reveals an "implied risk parameter" that lies between the lower and upper bounds.

Due to method design, some elicitation methods, like MPL and CEM, can distinguish risk behaviour of risk seeking from risk aversion. Other methods, like SCL and INV, are limited to the risk aversion domain, because the implied risk parameter range of the least risk averse choice in the task (i.e. SCL: choosing \$48/\$0 lottery; INV: investing nearly all of the \$20 into "risky project") extends from a lower bound of negative infinite (risk seeking domain), through 0 (risk neutral) to a positive upper bound (risk averse).

Table G.1 Elicitation method choices and their implied range of risk parameters

No. safe choices made in task:	Implied range of risk parameter (φ) lower				upper	No. safe choices made in task:	Implied range of risk parameter (φ) lower				upper
MPL											
1	-∞	≤	φ	<	-1.7128						
2	-1.7128	≤	φ	<	-0.9468	SCL					
3	-0.9468	≤	φ	<	-0.4866	1	-∞	≤	φ	<	0.1975
4	-0.4866	≤	φ	<	-0.1426	2	0.1975	≤	φ	<	0.3818
5	-0.1426	≤	φ	<	0.1464	3	0.3818	≤	φ	<	0.6696
6	0.1464	≤	φ	<	0.4115	4	0.6696	≤	φ	<	2.0000
7	0.4115	≤	φ	<	0.6762	5	2.0000	≤	φ	<	∞
8	0.6762	≤	φ	<	0.9706						
9	0.9706	≤	φ	<	1.3684	INV**					
10/11 *	1.3684	≤	φ	<	∞	0 - 2	-∞	≤	φ	<	0.0956
						2 - 4	0.0956	≤	φ	<	0.1479
CEM						4 - 6	0.1479	≤	φ	<	0.1888
1	-∞	≤	φ	<	-6.9643	6 - 8	0.1888	≤	φ	<	0.2338
2	-6.9643	≤	φ	<	-2.7094	8 - 10	0.2338	≤	φ	<	0.2891
3	-2.7094	≤	φ	<	-1.0685	10 - 12	0.2891	≤	φ	<	0.3636
4	-1.0685	≤	φ	<	0.0000	12 - 14	0.3636	≤	φ	<	0.4750
5	0.0000	≤	φ	<	0.9317	14 - 16	0.4750	≤	φ	<	0.6685
6	0.9317	≤	φ	<	2.0000	16 - 18	0.6685	≤	φ	<	1.1098
7	2.0000	≤	φ	<	3.9307	18 - 20	1.1098	≤	φ	<	3.3208
8	3.9307	≤	φ	<	∞	20	3.3208	≤	φ	<	∞

MPL: Multiple Price List. SCL: Single Choice List. CEM: Certainty Equivalent Method. INV: Investment Game.

* MPL choice 11: never switch from "safe" *Option A* to "risky" *Option B*

** INV choices are binned in $2. Each bin represents choices to "keep" the endowment away from the risky project, low <= choice < high

The implied range of risk parameters are estimated with the CRRA function: $u(x) = \begin{cases} \frac{x^{1-\varphi}}{(1-\varphi)} & if\ \varphi \neq 1 \\ \ln(x) & if\ \varphi = 1 \end{cases}$

Appendix H: Self-reported risk by cohort, gender, and timeframe

This appendix includes auxiliary comparisons of self-evaluated domain-specific risk measures (Section 3.6.5) by cohort and gender (Table H.1 and Table H.2), by gender and cohort (Table H.3 and Table H.4) and cohort and timeframe (Table H.5 and Table H.6).

Table H.1 Self-evaluated risk aversion between genders, controlled by cohort (present)

Risk Domain (present)	Control					Incident Commander				
	Male		Female		Wilcoxon	Male		Female		Wilcoxon
	N	Mean	N	Mean	p-value	N	Mean	N	Mean	p-value
Career	47	4.32	71	4.42	0.9408	56	4.59	5	5.00	0.1713
Financial	47	3.83	71	4.38	0.0046	56	3.93	5	4.20	0.4690
Health	47	4.15	71	4.32	0.5623	56	3.55	5	3.60	0.6510
Recreational	47	3.66	71	3.90	0.1479	56	3.27	5	4.40	0.0289
Safety	47	4.06	71	3.80	0.2662	56	3.50	5	3.80	0.6442
Social	47	3.62	71	3.68	0.9000	56	3.34	5	3.60	0.6090

Table H.2 Self-evaluated risk aversion between genders, controlled by cohort (past)

Risk Domain (past)	Control					Incident Commander				
	Male		Female		Wilcoxon	Male		Female		Wilcoxon
	N	Mean	N	Mean	p-value	N	Mean	N	Mean	p-value
Career	47	1.77	71	1.72	0.6667	56	1.88	5	1.40	0.3784
Financial	47	1.96	71	1.68	0.1699	56	2.38	5	2.00	0.4212
Health	47	2.19	71	1.96	0.2031	56	3.25	5	3.60	0.5407
Recreational	47	2.34	71	2.56	0.6024	56	3.64	5	2.20	0.0132
Safety	47	2.38	71	2.54	0.5967	56	3.38	5	3.20	0.6930
Social	47	2.11	71	2.23	0.2740	56	2.50	5	2.20	0.5783

Referencing Table H.1 and Table H.2, we observe that among the Control group, Females only seem to be significantly different from Males in stating higher *Financial* risk aversion in the present day ($p < 0.01$); among Incident Commanders, Females are significantly more risk averse in present *Recreation*, though less risk averse in past *Recreation* ($p < 0.05$).

Table H.3 Self-evaluated risk aversion between cohorts, controlled by gender and cohort (present)

Risk Domain (present)	Male					Female				
	Control		IC		Wilcoxon	Control		IC		Wilcoxon
	N	Mean	N	Mean	p-value	N	Mean	N	Mean	p-value
Career	47	4.32	56	4.59	0.3406	71	4.42	5	5.00	0.1045
Financial	47	3.83	56	3.93	0.8615	71	4.38	5	4.20	0.5034
Health	47	4.15	56	3.55	0.0019	71	4.32	5	3.60	0.1694
Recreational	47	3.66	56	3.27	0.0556	71	3.90	5	4.40	0.3865
Safety	47	4.06	56	3.50	0.0039	71	3.80	5	3.80	0.8443
Social	47	3.62	56	3.34	0.0857	71	3.68	5	3.60	0.6952

Table H.4 Self-evaluated risk aversion between cohorts, controlled by gender (past)

Risk Domain (past)	Male					Female				
	Control		IC		Wilcoxon	Control		IC		Wilcoxon
	N	Mean	N	Mean	p-value	N	Mean	N	Mean	p-value
Career	47	1.77	56	1.88	0.1838	71	1.72	5	1.40	0.6885
Financial	47	1.96	56	2.38	0.0302	71	1.68	5	2.00	0.4527
Health	47	2.19	56	3.25	0.0000	71	1.96	5	3.60	0.0105
Recreational	47	2.34	56	3.64	0.0000	71	2.56	5	2.20	0.7214
Safety	47	2.38	56	3.38	0.0000	71	2.54	5	3.20	0.2535
Social	47	2.11	56	2.50	0.0432	71	2.23	5	2.20	0.9129

Referencing Table H.3 and Table H.4, we see that among male participants, ICs are significantly less risk averse than Control in present-day *Health* and *Safety* $(p < 0.01)$, and marginally less risk averse in *Recreational* and *Social* $(p < 0.10)$; in their adult past, male ICs were more risk averse than male Control in *Health*, *Recreational*, *Safety* $(p < 0.01)$ as well as in *Financial* and *Social* $(p < 0.05)$. These observations are close to the observed differences in Table 3.21, in which male and female were pooled in Control/IC groups were pooled. Differences between Control and IC are less pronounced among females; female ICs are only significantly different from Control in higher risk aversion in the past *Health* domain $(p < 0.05)$.

Table H.5 Self-evaluated risk aversion between timeframe, controlled by cohort (females)

Risk Domain	Control					Incident Commander				
	Now		Past		Wilcoxon	Now		Past		Wilcoxon
	N	Mean	N	Mean	p-value	N	Mean	N	Mean	p-value
Career	71	4.42	71	1.72	0.0000	5	5.00	5	1.40	0.0046
Financial	71	4.38	71	1.68	0.0000	5	4.20	5	2.00	0.0250
Health	71	4.32	71	1.96	0.0000	5	3.60	5	3.60	0.8266
Recreational	71	3.90	71	2.56	0.0000	5	4.40	5	2.20	0.0135
Safety	71	3.80	71	2.54	0.0000	5	3.80	5	3.20	0.3886
Social	71	3.68	71	2.23	0.0000	5	3.60	5	2.20	0.1037

Table H.6 Self-evaluated risk aversion between timeframe, controlled by cohort (males)

Risk Domain	Control					Incident Commander				
	Now		Past		Wilcoxon	Now		Past		Wilcoxon
	N	Mean	N	Mean	p-value	N	Mean	N	Mean	p-value
Career	47	4.32	47	1.77	0.0000	56	4.59	56	1.88	0.0000
Financial	47	3.83	47	1.96	0.0000	56	3.93	56	2.38	0.0000
Health	47	4.15	47	2.19	0.0000	56	3.55	56	3.25	0.1418
Recreational	47	3.66	47	2.34	0.0000	56	3.27	56	3.64	0.0598
Safety	47	4.06	47	2.38	0.0000	56	3.50	56	3.38	0.4530
Social	47	3.62	47	2.11	0.0000	56	3.34	56	2.50	0.0003

When comparing present-day and past risks among genders of the same cohort, or cohorts of the same gender, we observe the significant differences between sub-groups are very similar to those between gender-pooled Control/IC groups in Table 3.22. However, it is interesting to note, when comparing past *Recreational* risks to current levels, female ICs today are significantly more risk averse $(p < 0.05)$, while their male counterparts exhibit marginally significant decreased risk aversion $(p < 0.10)$.

Appendix I: Instructions for the Risk Elicitation Economics Experiment

This appendix includes screenshots from the risk elicitation experiment undertaken from March to April 2021. I coded all components independently in the *oTree* framework (Chen et al., 2016); tasks MPL, SCL and CEM were adapted from code found in the online Supplementary Material folders of Holzmeister (2017).

UNIVERSITY OF ALBERTA
DEPARTMENT OF RESOURCE ECONOMICS
AND ENVIRONMENTAL SOCIOLOGY

Alberta Wildfire

Information Page

Title of study: Risk Elicitation Economic Experiments

Research Team:

Principal Investigator	Research Assistant
Bruno Wichmann	Michael Huang
Associate Professor, Department of Resource Economics & Environmental Sociology	MSc Student, Department of Resource Economics & Environmental Sociology
University of Alberta	University of Alberta
Email: bwichmann@ualberta.ca	Email: michael.huang@ualberta.ca

Invitation to Participate: You are invited to participate in this online research study about economic behaviour and lottery preferences.

Purpose: The purpose of today's study is to learn about economic behaviour and lottery preferences.

Study procedures: This is an experiment in individual decision-making. The instructions are simple, and if you follow them carefully you will have the opportunity to earn real money. You will be paid an initial $10 payment for your participation. You will have the opportunity to earn additional money, depending on the decisions you make, as explained later in the session. Payments for this session will be made as online *Amazon.ca* gift cards. This session will last approximately 30 minutes.

Confidentiality: The information that you share through participation will be maintained anonymous. Your identification number will be used to determine how much you are paid; your decisions will not be associated with your name. In order to maintain confidentiality and anonymity, it is important that you do not communicate with your peers and/or other individuals during the experiment, and that you refrain from talking about the contents of this experiment with other potential participants.

Your name may be used in order to associate the data you will have submitted in this experiment with secondary data on wildfire suppression operations. The decisions you make in this experiment *will not* be associated with your name. The decisions you make, as well as your personal choice to participate, will never be disclosed to anyone.

Data Management: This online experiment is hosted on the Heroku platform, whose secure servers are based in and subject to the privacy laws of the USA. Your name and email address are not stored on the Heroku platform; identifying information that is used for payment and further research will be securely stored on University of Alberta servers.

Benefits: The information collected from this session will be used is a study that examines how economic behaviour and lottery preferences affect decision-making. For participating in this study, you will receive via email an online *Amazon.ca* gift card with payment corresponding to your performance in the tasks that follow.

Risks: There are no known risks to you from participating in this study.

Voluntary participation: Your participation in this study is voluntary.

Freedom to withdraw: You are free to withdraw from the study anytime. If you withdraw before the end of the experiment, you will retain the $10 participation payment, but will not receive additional money. You cannot withdraw your data after you have completed the study and submitted your responses.

Funding: This study is being funded by the Wildfire Management Branch of Alberta Department of Agriculture and Forestry (University of Alberta Research Services Office Proposal ID *RES0053118*). This project is not funded by any non-governmental organizations.

Use of information: The information you provide may be used for research projects, including, for example, research reports, a MSc thesis, and articles in academic journals. Your name and identification will remain confidential.

Contact Information: If you have any questions or require more information about the study, you may contact the research team (please refer to contact information above).

Research Ethics: The plan for this study has been reviewed by a Research Ethics Board at the University of Alberta. If you have questions about your rights or how research should be conducted, you can call the Research Ethics Office directly at (780) 492-2615, or reoffice@ualberta.ca, and refer to the Protocol number of this experiment, *Pro00106176*. This office is independent of the researchers.

Next

Consent Form

In order to proceed with this experiment, you will need to consent to the following:

- I understand that I have been asked to participate in an experiment about economic behavior and lottery preferences.
- I understand that the information collected is part of a research being developed by the researchers listed above.
- I have read and understood the Information Page.
- I understand the benefits and risks involved in taking part in this experiment.
- I understand that my participation is voluntary and that I can choose to withdraw at any point during the experiment.
- I understand that if I withdraw from the study before submitting the data, the information I provide up to the point of withdrawal will not be used by the researchers.
- I understand that If I withdraw before the end of the experiment, I will retain the $10 participation payment, but will not receive additional money.
- I understand that the information that I provide will be kept in strict confidence and that any link between my answers and my name will be destroyed.
- I give the research team permission to use information that I provide for the purposes specified in the Information Page.

I understand the purpose, risks, and benefits of this study.
By clicking "Yes" below, I agree to participate in the study.

○ Yes ○ No

Back Submit

Welcome

Thank you for participating in this experiment.

In this session (approx. 30 minutes), you will complete four decision-making tasks and a survey. These activities are designed to engage your economic behaviour.

Tasks: In a typical setting, individuals pay for a lottery ticket and may get lucky or not. In the settings that will be presented to you, we will give you a free ticket, but you will have to choose which lottery on the ticket you want to play. These are real lotteries over real money.

Earnings: After all four tasks have been completed, one of the tasks will be randomly selected to determine your earnings for today's session. Even though you will make decisions for four tasks, only the decision from one task will end up affecting your earnings. You will not know in advance which task that will be; each task has an equal chance of being selected.

Withdrawal: If you withdraw before the end of the experiment, you can keep the $10 participation payment, but will not receive additional money. Any information you have submitted up to this point cannot be withdrawn. To withdraw from the experiment, exit your internet browser and notify the research team by email. You must notify the research team of your withdrawal in order to receive the $10 participation payment.

No communication: It is important that you do not communicate with your peers and/or other individuals during the experiment. After this session, do not talk about the contents of this experiment with other potential participants.

Click "Next" to proceed to the four decision-making tasks.

Next

Instructions Decision

Task:

Task "Green" Instructions

In the following, you are endowed with $20.00 that you can choose to keep or invest in a risky project.

The risky project has a 50% chance of success.

- If the project is successful, you will earn $2.50 for every $1.00 invested
- If the project is unsuccessful, you will lose the amount invested

Task Payoff

The task payoff will include:

- the amount not invested in the risky project, plus
- the return on the amount invested in the risky project

Next

4/14/2021 Task:

Instructions Decision

Task:

Task "Green" Decision

Please choose how much you would like to invest.

Note:
Enter any value between $0.00 to $20.00, including $0.00 and $20.00.
You may enter a value to the nearest $.01

$

Back Submit

https://experiment-rees.herokuapp.com/p/phgzyhhm/tasks/Green/5/ 1/1

Task:

Task "Green" is completed

Click "Next" to continue on to the next Task.

Next

Instructions Decision

Task:

Task "Violet" Instructions

In the following, you will be presented with a list of paired lotteries.
Each lottery has two possible outcomes, determined by a virtual coin toss.

Your Decision
Choose the one lottery you prefer from the list.

Task Payoff
A virtual coin toss will be played for the lottery you selected. "Heads" or "Tails" will determine the task payoff.

Next

Instructions Decision

Task:

Task "Violet" Decision

Choose your preferred lottery (1, 2, 3, 4, or 5), and then click "Submit".

	Heads		Tails		
Lottery No.	Prob.	Payoff	Prob.	Payoff	Your Choice
1	50%	$16	50%	$16	○
2	50%	$24	50%	$12	○
3	50%	$32	50%	$8	○
4	50%	$40	50%	$4	○
5	50%	$48	50%	$0	○

Back Submit

Task:

Task "Violet" is completed

Click "Next" to continue on to the next Task.

Next

Instructions Decision

Task:

Task "Orange" Instructions

In the following, you will be presented with a list of paired lotteries: *Option A* and *Option B*.

Values:

- *Option A* will have the same high/low values on every line
- *Option B* will have the same high/low values on every line (different from *Option A*).

Probabilities:

- There will be **different** probabilities for high/low values from line to line.

Your Decision

Consider the values and probabilities of Options A & B. In each line you must choose which option you prefer. Your possible choices:

- select Option A at Line 1, and stay with Option A down all lines of the list
- select Option A at Line 1, and switch to Option B at some line down the list
- select Option B at Line 1, and stay with Option B down all lines of the list

Task Payoff

To determine the task payoff, one line will be randomly selected. A lottery will be played for the Option you selected in this round. You will receive the high or low value in the Option you chose, depending on the probabilities.

Next

Instructions Decision

Task:

Task "Orange" Decision

Consider the high and low values of Options A & B and their probabilities. Your possible choices:

- select Option A on the first line, and stay with Option A down all lines of the list
- select Option A on the first line, and switch to Option B at some line down the list
- select Option B on the first line, and stay with Option B down all lines of the list

Line	Option A	Choice	Option B
1	10% probability of $20.00 90% probability of $16.00	○ ○	10% probability of $38.50 90% probability of $1.00
2	20% probability of $20.00 80% probability of $16.00	○ ○	20% probability of $38.50 80% probability of $1.00
3	30% probability of $20.00 70% probability of $16.00	○ ○	30% probability of $38.50 70% probability of $1.00
4	40% probability of $20.00 60% probability of $16.00	○ ○	40% probability of $38.50 60% probability of $1.00
5	50% probability of $20.00 50% probability of $16.00	○ ○	50% probability of $38.50 50% probability of $1.00
6	60% probability of $20.00 40% probability of $16.00	○ ○	60% probability of $38.50 40% probability of $1.00
7	70% probability of $20.00 30% probability of $16.00	○ ○	70% probability of $38.50 30% probability of $1.00
8	80% probability of $20.00 20% probability of $16.00	○ ○	80% probability of $38.50 20% probability of $1.00
9	90% probability of $20.00 10% probability of $16.00	○ ○	90% probability of $38.50 10% probability of $1.00
10	100% probability of $20.00 0% probability of $16.00	○ ○	100% probability of $38.50 0% probability of $1.00

Back Submit

Task:

Task "Orange" is completed

Click "Next" to continue on to the next Task.

Next

Instructions Decision

Task:

Task "Blue" Instructions

In the following, you'll face a list of paired choices between:

- playing a *50/50 Coin Toss* lottery, or
- receiving a *Sure Payment*.

For each line in the list:

- *If you choose the* Coin Toss,
  - there is 50% probability of receiving a high value,
  - there is 50% probability of receiving a low value.
- *If you choose the* Sure Payment,
  - there is 100% probability of receiving the value on this line.

Your Decision

The first and final line have been pre-selected (because in these lines the selected option will always give you a better payoff than the other option).

Your task is to choose at which line you want to switch from the *Coin Toss* option to the *Sure Payment* option.

Task Payoff

After you have completed your decision, one line will be randomly selected to determine your task payoff.

- If you chose *Coin Toss* in this line, a virtual coin toss will determine a high/low value for your payoff
- if you chose *Sure Payment* in this line, its value will be your task payoff.

Next

Instructions Decision

Task:

Task "Blue" Decision

The options on Line 1 and Line 9 have already been selected.
choose at which line you want to switch from the *Coin Toss* option to the *Sure Payment*

Click "Submit" when you are ready.

Line	50/50 Coin Toss	Choice	Sure Payment
1	$30.00 *(if Heads)* $10.00 *(if Tails)*	◉ ○	$10.00
2	$30.00 *(if Heads)* $10.00 *(if Tails)*	○ ○	$12.50
3	$30.00 *(if Heads)* $10.00 *(if Tails)*	○ ○	$15.00
4	$30.00 *(if Heads)* $10.00 *(if Tails)*	○ ○	$17.50
5	$30.00 *(if Heads)* $10.00 *(if Tails)*	○ ○	$20.00
6	$30.00 *(if Heads)* $10.00 *(if Tails)*	○ ○	$22.50
7	$30.00 *(if Heads)* $10.00 *(if Tails)*	○ ○	$25.00
8	$30.00 *(if Heads)* $10.00 *(if Tails)*	○ ○	$27.50
9	$30.00 *(if Heads)* $10.00 *(if Tails)*	○ ◉	$30.00

Back Submit

Task:

Task "Blue" is completed

Click "Next" to continue on to the next Task.

Next

Surveys

We are interested in knowing a bit about you. Please complete the following short surveys.

Next

Survey: Self-Evaluation

We are interested in everyday risk-taking. Please could you tell us if any of the following have ever applied to you *now* or in your adult *past*?

Please use the scales as follows:

1	2	3	4	5
never	rarely	quite often	often	very often

	Now 1 2 3 4 5	Past 1 2 3 4 5
recreational risks *(e.g. rock-climbing, scuba diving)*	○ ○ ○ ○ ○	○ ○ ○ ○ ○
health risks *(e.g. smoking, poor diet, high alcohol consumption)*	○ ○ ○ ○ ○	○ ○ ○ ○ ○
career risks *(e.g. quitting a job without another to go to)*	○ ○ ○ ○ ○	○ ○ ○ ○ ○
financial risks *(e.g. gambling, risky investments)*	○ ○ ○ ○ ○	○ ○ ○ ○ ○
safety risks *(e.g. fast driving, city cycling without a helmet)*	○ ○ ○ ○ ○	○ ○ ○ ○ ○
social risks *(e.g. standing for election, publicly challenging a rule or decision)*	○ ○ ○ ○ ○	○ ○ ○ ○ ○
	1 2 3 4 5 Now	1 2 3 4 5 Past

When you are finished, please click "Next" to submit and continue.

Next

Survey

We would like to know a bit about who you are.

What is your gender?	---------
In which year were you born?	---------
What is your highest level of education certification?	---------
Over the past five years, have you had (or do you currently have) dependent family? *(e.g. family members who are financially dependent on you)*	---------
In which year did you first join Alberta Wildfire?	---------
Since then, how many complete fire seasons have you worked?	---------
What is your current Incident Commander certification?	---------
What is your current title in Alberta Wildfire?	
How many fire seasons have you completed in your current role?	---------

Next

Survey Completed

This is the end of the survey.

Please select "Next" to submit your results, and find out how much you have earned from this experiment.

Next

Your Task Payoffs

In the next page, you will see a summary of the four tasks and their outcomes.

- **Task "Orange"**
- **Task "Green"**
- **Task "Blue"**
- **Task "Violet"**

Take a look at your results.

Next

Results of Tasks

Here is a summary of the four tasks and their outcomes.

At the bottom of the page, click "Next" to see how much you have earned.
Your total payment will be:

- $10 participation payment, plus
- the payoff from *one* of these four tasks (to be selected at random)

Task "Orange"

You were tasked to choose between paired lotteries.
The following line was randomly selected, and this was your decision:

Line	Option A			Option B
5	50 % probability of $20.00 50 % probability of $16.00	◉	○	50 % probability of $38.50 50 % probability of $1.00

On this line, you selected **Option A**.

A lottery has played for these probabilities.
The task payoff is **$20.00**.

Task "Green"

You were given a $20.00 endowment, and tasked to choose how much you wanted to invest. Your choice: invest $0.00.

You chose to keep: $20.00.

You chose to invest nothing, so you will keep your initial $20.00 as payoff.

The task payoff is **$20.00**.

Task "Blue"

You were tasked to choose between a "50/50 Coin Toss" and a Sure Payment. The following line was randomly selected, and this was your decision:

Line	50/50 Coin Toss			Sure Payment
2	$30.00 *(if heads)* $10.00 *(if tails)*	◉	○	$12.50

On this line, you selected **50/50 Coin Toss**.
A coin was flipped, and the result is *Heads*

The task payoff is **$30.00**.

Task "Violet"

You were tasked to choose between five lotteries. This was your decision:

	Heads		Tails		
Lottery No.	Prob.	Payoff	Prob.	Payoff	Your Choice
5	50%	$48.00	50%	$0.00	◉

A virtual coin toss has resulted in: *Tails*

The task payoff is **$0.00.**

Next

This concludes the session

The task that was randomly selected is:

Task "Violet"

You were tasked to choose between five lotteries. This was your decision:

	Heads		Tails		
Lottery No.	Prob.	Payoff	Prob.	Payoff	Your Choice
5	50%	$48.00	50%	$0.00	◉

A virtual coin toss has resulted in: *Tails*

The task payoff is **$0.00.**

Including the $10 participation payment, your total payment is:

$10.00

Your payment in the form of an *Amazon.ca* gift card will be sent to your email address in three business days. Thank you for your participation!

You can click on the "Finish" button and then close this window.

Finish

Appendix J: Supplementary literature review

This appendix includes literature with methods, both within the economics discipline and beyond, that were not used in the project. This supplementary literature review provides a more complete perspective on current research in risk behaviour research, both within economics and in other disciplines.

J.1. Additional experimental measures in the economics laboratory

Lottery lists and their variations are popular forms of economic revealed preference elicitation methods. Distilled from economic theory and readily adaptable for econometric analysis, this format is intuitive to economists, however, it may be daunting for subjects who are uncomfortable with hypothetical concepts of probability. Cognizant of these challenges, some economists have adapted EMs to be more palatable for their targeted participants by framing tasks in familiar themes and adding visual cues to assist understanding. Such examples are:

"Circles" in multiple price list (Hey and Orme, 1994b): Subjects in face 100 pairwise lotteries, in which payoff probabilities of each gamble is represented in a pie chart. Likelihood ratio tests are applied to results to determine which framework out of a wide-ranging series best represents each subject's decision-making: Expected Utility, Disappointment Aversion, Prospective Reference, Quadratic Utility, Regret, Rank Dependence, Weighted Utility, Yaari's probability/utility dual model. While no framework is conclusively superior, Hey and Orme find that Expected Utility Theory tops other frameworks as being "no worse off than other models" for 39% participants; they encourage more research into uncovering the noise parameter of the utility function.

Balloon analogue risk task (BART) (Lejuez et al., 2002): subjects are presented with a series computer simulated balloons and tasked with ”pumping” balloons for the opportunity to earn money; each additional pump marginally increases participant gains, but also monotonically increases the probability of the balloon popping, in which case earnings will be null, and the participant is presented with a new balloon. However, if the participant halts inflation the balloon prior to the “pop”, earnings from that round will be transferred to a permanent account. The authors found that revealed risk preferences correlated with self-evaluated addiction, health and safety behaviours.

Bomb risk elicitation task (BRET) (Crosetto and Filippin, 2013): this task parallels the balloon task, by incentivizing participants to select open computer simulated boxes, in which a randomly contains a “bomb” that will wipe out all earnings. A key difference with the balloon task lies in the fact that participants are aware there are 100 numbered boxes, and thus, they can more readily infer changing risk probabilities with each additional box opening.

In order to assist subjects in understanding lottery probabilities, many experimental economists have used physical props during experiments. Often, experimentalists use sets of dice to explain probability and to randomize outcomes of lotteries (Anderson and Mellor, 2008; Charness et al., 2018; Goeree et al., 2002). Tanaka et al. (2010), who adapted the Holt-Laury Multiple Price List into multiple series totalling 35 lines, employ a bingo cage with 35 numbered balls to determine randomized outcomes.

J.2. Eliciting risk attitudes outside of the economics laboratory

Experimental economists seek to apply findings from the laboratory to real-life behaviour; yet occasionally, there are natural experiments from which economists are able to analyze the behaviour of subjects without experimental intervention. Risk preference can be observed in the real world, and can explain esoteric phenomena like professional golfers' choice of putting, as analyzed in a Prospect Theory framework (Pope and Schweitzer, 2011), or stock trading frequencies due to the distorting effect of investors' risk-seeking behaviour on expected utility calculation (Barber and Odean, 2001; Odean, 1999, 1998).

However, the form of natural experiment that most readily lends itself to contemporary lottery list elicitation methods is televised game shows. Participants on shows like *Deal or No Deal* are offered series of risky choices, often increasing in risk, inducing risk-seeking behaviour (and gaudy television excitement). *Deal or No Deal,* along with its international variations, are real-world formats of the Certainty Equivalent method; as participants face increasingly risky lotteries between high/low payoffs, they are countered by the host who offers a marginally rising value of sure payment. As a natural experiment, the game show provides economists with rich data on risk preferences in the context of large monetary stakes (Bombardini and Trebbi, 2012; Post et al., 2008). This format of EM has also been adapted for an experimental laboratory in which Guiso et al. to measured a change in bank savings account holders' risk aversion after the 2008 financial crisis (2018).

J.3. Risk behaviour research in other disciplines

Further, researchers in neuroscience and cognitive science focus on biophysical indicators of risk behaviour in the brain, through using technology such as functional magnetic resonance imaging (fMRI); this method is sometimes operationalized in conjunction with those developed by economists and psychologists, towards developing an interdisciplinary understanding of human risk behaviour (Peterson, 2007; Schonberg et al., 2011).

Neuroscience researchers often employ functional fMRI to measure biophysical indicators in the brain that signal a propensity for risk aversion (Christopoulos et al., 2009), to determine which cortexes of the brain represent domain-specific risk preferences (Levy and Glimcher, 2011), and to investigate changes in neurotransmitters (e.g. dopamine, serotonin) and hormones (e.g. oxytocin, testosterone) can affect risk behaviour (for a review, see: Crockett and Fehr 2014). Neuroscience and psychophysics literature validates some revealed preference elicitation methods proposed by economists (Schonberg et al., 2011; Trepel et al., 2005), and the findings encourage further interdisciplinary collaboration between behavioural economists and cognitive scientists (Frydman and Camerer, 2016).

J.4. Comparing elicitation methods between economics and cognitive science

Cognitive science researchers study the mind through an interdisciplinary lens, incorporating elements from neuroscience, psychology and anthropology, among other physical and social sciences (Thagard, 2005). For instance, cognitive neuroscience researchers often use functional magnetic resonance imaging (fMRI) technology to discover the links between risk-seeking behaviour with activity in certain cortexes of the brain, and validate these findings using experimental methods originating in economics. Schonberg et al. (2011) review a series of such literature in an attempt towards bridging the inherent gap between economic and cognitive science perspectives on risk aversion. The authors summarize correlation analyses in which certain cortexes are found to be more active in subjects exhibiting economic risk-seeking behaviour in incentivized EMs, but, brain activity is not determinant of naturalistic risk-taking. However, when subjects encounter complex EMs such as BART and the Iowa Gambling Task, economic risk-seeking does indeed align with naturalistic risk-taking indices but not necessarily with brain activity.

Peterson (2007) reviews neuroscience literature on the influence of moods and emotions on financial investment decision-making, which includes research showing the effect of increased serotonin uptake on risk aversion (Arnold et al., 2004; Flory et al., 2004), as well as two studies in which researchers predict subjects' choices in incentivized economic EMs based on anticipatory brain activity, as measured by fMRI (Kuhnen and Knutson, 2005; Paulus et al., 2003).

Samanez-Larkin et al. (2010) motivate their study by asking how decisions by aging financial investors impacts the global economy. Towards addressing this question from a neuroeconomic perspective, the authors task subjects with a dynamic incentivized investment game[37], while using fMRI to evaluate brain activity during the task. During the task, they observe older subjects being more likely to make "suboptimal mistakes" (selecting a risky option when the sure payment was optimal), as well as "confusion mistakes" (selecting one risky option when another risky option was optimal). Regression results on age and measured fMRI activity show a positive relationship between age and neural decline. These findings lead the researchers to support existing neuroscience literature that surmise older subjects are, contrary to popular stereotype, less risk averse than younger counterparts (Mather 2006), and are more likely to make decision-making errors when facing risk. (Denburg et al., 2007; Mohr and Nagel, 2010; Peters et al., 2016).

[37] The game resembles Certainty Equivalent Method: on each decision line there is a choice between playing a lottery ("Stocks") or taking a sure payment ("Bond").

Appendix K: Structural estimation of risk by IC level

This appendix includes a structural estimation of Incident Commanders' (ICs) risk preferences, in which ICs are analyzed as subsamples split by certification levels (IC1 being highest and IC4 being lowest[38]). As seen in Table K.1 and Figure K.1, below, there is no clear relationship between IC certification seniority and risk aversion across four tasks.

[38] There is one IC5 in our sample; a single observation provides insufficient variation for a Maximum Likelihood estimation.

Table K. 1 Estimates of risk and noise parameters (φ, σ) by IC levels

	MPL	SCL	CEM	INV
Panel A: IC1				
φ	0.476	0.585***	1.000***	0.147***
	(0.308)	(0.181)	(0.000)	(0.040)
σ	2.551	1.498*	0.264***	0.191***
	(1.697)	(0.872)	(0.051)	(0.026)
N	60	24	48	60
N_clust	6	6	6	6
Panel B: IC2				
φ	0.257	0.259***	0.216	0.126***
	(0.240)	(0.064)	(0.497)	(0.028)
σ	4.777	0.842***	1.952	0.221***
	(3.140)	(0.256)	(2.663)	(0.032)
N	110	44	88	110
N_clust	11	11	11	11
Panel C: IC3				
φ	0.738***	0.545***	-0.146	0.190***
	(0.133)	(0.092)	(0.402)	(0.028)
σ	1.412***	1.332***	7.622	0.264***
	(0.387)	(0.314)	(9.473)	(0.047)
N	320	128	256	320
N_clust	32	32	32	32
Panel D: IC4				
φ	0.531**	0.417***	0.357	0.218***
	(0.248)	(0.104)	(0.530)	(0.049)
σ	2.343*	0.718***	1.376	0.173***
	(1.318)	(0.252)	(2.425)	(0.055)
N	90	36	72	90
N_clust	9	9	9	9

Robust standard errors clustered at the subject level in reported parentheses.
* p<0.10, ** p<0.05, *** p<0.01

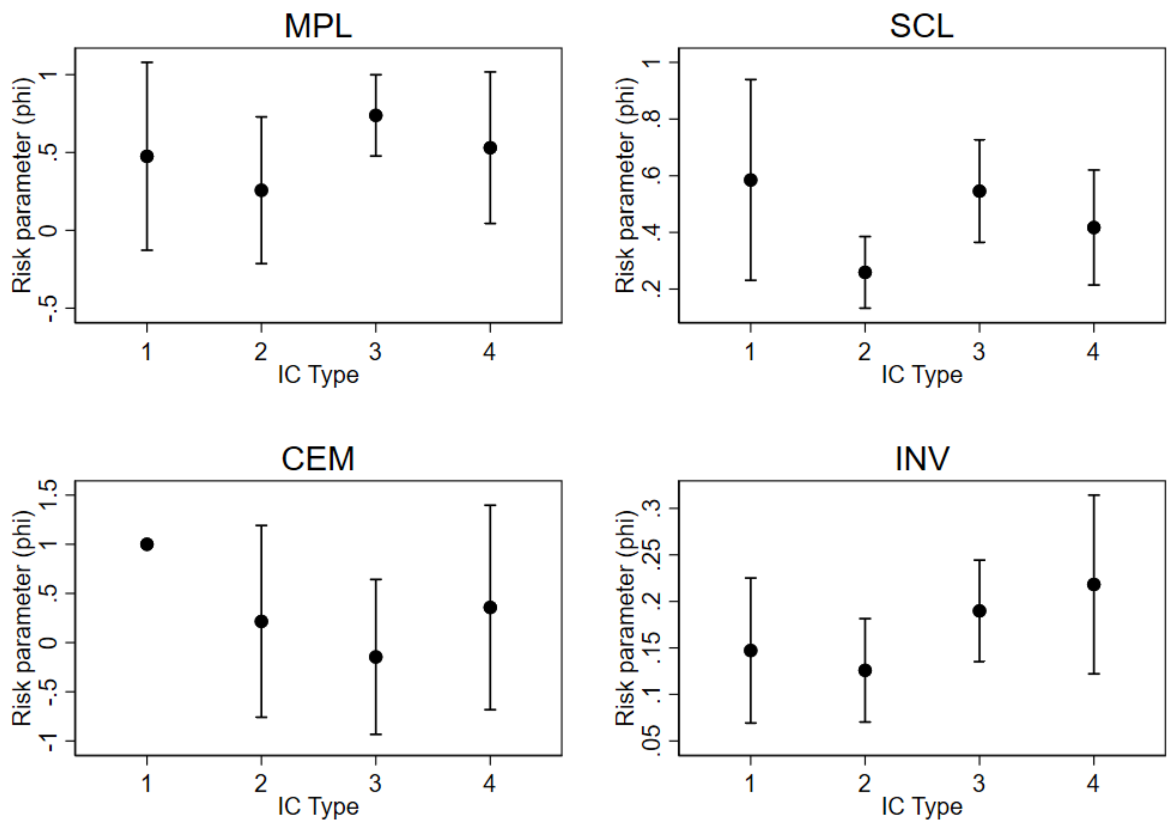

Figure K. 1 Risk aversion estimators by Incident Commander certification levels