VNIVERSITAT DE VALÈNCIA Facultat d'Economia

Departament d'Anàlisi Econòmica

The human factor in cybersecurity: An experimental approach to cyber-risk and cyberinsurance

DOCTORAL DISSERTATION
Programme in Industrial Economics

By
Yolanda Gómez González

Supervisor:
José Vila Gisbert

January 2021

A mi abuela.

Agradecimientos

En primer lugar, quiero expresar mi agradecimiento a mi director, Dr. José Vila, por la dedicación y el gran apoyo que me ha brindado no sólo durante el tiempo de desarrollo de esta tesis, si no también durante mi desarrollo personal y profesional. Sin duda todo un ejemplo a seguir y una gran influencia en mi vida.

A continuación, me gustaría agradecer a Dr. David Rios Insua, Dra. Pam Briggs y Dra. Dawn Branley-Bell su colaboración y ayuda. I really appreciate your professionalism and commitment and I hope this to be only the beginning of a long journey together. Agradecer también a todo el equipo del Proyecto Europeo Horizonte 2020 (740920) CYBECO, *Supporting Cyberinsurance from a Behavioural Choice Perspective* por el esfuerzo conjunto realizado y que ha servido de base a esta tesis.

También me gustaría mostrar mi agradecimiento al proyecto *Transmisión de información entre empresas y discriminación de precios. Políticas de "opt-out" y "opt-in" y bienestar social. Diseño óptimo. Ciberseguridad. Género y brecha digital* (PID2019-110790RB-I00) financiado por el Ministerio de Ciencia, Innovación y Universidades, en especial a Dra. Amparo Urbano, quién firmó mi primer contrato profesional y a Dr. Iván Arribas por aguantar mis interminables consultas en estadística.

Por otro lado, este trabajo no hubiera sido posible sin el apoyo de todo el equipo de DevStat, donde me gustaría destacar a Pablo Ferrer por su incansable ayuda y a José Luis Cervera por la confianza depositada en mí.

Finalmente agradecer a mi familia y amigos el estar siempre ahí, la paciencia y las horas dedicadas a escucharme hablar de este proyecto.

Acronyms

ARA Adversarial Risk Analysis.

ARU Axiomatic Random Utility.

ASMs Advance Security Measures.

BEE Behavioural Economic Experiment.

BSMs Basic Security Measures.

CFA Confirmatory Factor Analysis.

DM Decision Maker.

EUM Expected Utility Maximisation.

ICT Information and Communication Technologies.

ISP Information Security Policy.

MNL Multinomial logit.

NCSC National Cyber Security Centre.

PMT Protection Motivation Theory.

RPS Risk Propensity Scale.

SEJ Structured Expert Judgementy.

SEM Structural Equation Modelling.

SM Security Measures.

SMEs Small and Mid-size Enterprises.

TPB Theory of Planned Behaviour.

VC Virtual Currency units.

Contents

List of Tables

List of Figures

Introducción

Este trabajo de investigación tiene como objetivo el desarrollo y validación experimental de modelos conductuales, con sólido fundamento teórico, capaces de explicar y prever la adopción de ciberseguro, así como los de los elementos clave de ciberseguridad detrás de dicha adopción. Con este fin, la presente disertación se centra en tres dimensiones clave en ciberseguridad: *ciberseguro* (adopción de productos de seguros que cubren parcialmente el impacto de posibles ataques), *ciberprotección* (adopción de medidas capaces de reducir el riesgo de sufrir un ataque) y *comportamiento online* (nivel de riesgo asumido por los usuarios cuando navegan en Internet). Estas dimensiones recogen aspectos conductuales relevantes que condicionan la adopción de ciberseguro, tales como: (i) la racionalidad en el reparto del presupuesto disponible para ciberseguridad entre productos de ciberprotección y seguros, (ii) posibles efectos negativos causados por la asimetría de información intrínseca a cualquier tipo de seguro (incluido el ciberseguro) y; (iii) formación de creencias sobre cibervulnerabilidad, especialmente en la percepción del nivel riesgo de recibir un ataque intencional, así como los métodos de elicitación de dichas creencias. Cumpliendo este objetivo, nuestra investigación contribuye a llenar un vacío en la literatura sobre la toma de decisión de compra de ciberseguros y la formación de percepciones sobre el ciber-riesgo. Tal y como se muestra en la sección de discusión de esta disertación, esta aportación tiene un papel relevante tanto científico como de formulación de políticas y de desarrollo empresarial.

Los fundamentos teóricos de nuestro trabajo se basan en un enfoque multidisciplinar que reúne elementos de las últimas investigaciones en el campo del ciberseguro, integrándolas con (i) elementos de la literatura en economía sobre elección racional, utilidad aleatoria y economía conductual; (ii) mecanismos de formación de creencias en interacciones estratégicas que no se basen en el supuesto de conocimiento común y (iii) palancas psicológicas y conductuales que condicionan la toma de decisiones en ciberseguridad. La validación empírica de nuestros modelos se ha realizado aplicando un enfoque de economía conductual-experimental. Los datos experimentales se han obtenido a través de: (i) un experimento online a gran escala con 4,800 sujetos y realizado en 4 países europeos (Alemania, Polonia, España y Reino Unido) que se describe en el Capítulo 1 y (ii) un experimento presencial en laboratorio realizado en España con 100 sujetos que se describe en el Capítulo 3.

La ciberseguridad se percibe cada vez más como un problema global (World Economic Forum, 2020), más relevante aún a medida que las empresas, las administraciones y las personas están más interconectadas, lo que aumenta las oportunidades y facilita la propagación de ciberamenazas. Ejemplos de ciberataques relevantes son: (i) el ataque a Equifax de 2017, con el robo de datos de más de 140 millones de clientes, incluidos sus números de seguridad social y tarjetas de crédito; (ii) el ataque WannaCry de 2017, que derribó, entre otros, al Servicio Nacional de Salud del Reino Unido, Telefónica y FedEx, obstaculizando gravemente sus operaciones y provocando pérdidas estimadas en $4 mil millones (Berr, 2017); y (iii) el malware NotPetya, que afectó a miles de organizaciones en todo el mundo con un coste estimado de $10 mil millones (Greenberg, 2018). Pese a que las organizaciones implementan soluciones tecnológicas con antivirus, cortafuegos y sistemas de detección de intrusiones, numerosos fallos de ciberseguridad se deben a una componente humana no tecnológica. Por ejemplo, el error humano representó un 24% de los problemas de seguridad según en el último informe de Ponemon Institute and IBM Security (2020). Así, el comportamiento humano se puede considerar como el eslabón más débil de la cadena de ciberseguridad, convirtiendo el cambio de comportamiento en una estrategia fundamental para la mejora de los niveles de seguridad. Sin embargo, para abordar de forma efectiva la transformación del componente humano de la ciberseguridad, debemos comprender primero cómo funcionan los elementos racionales y no racionales (aunque predecibles) que afectan al comportamiento y la formación de creencias en ciberseguridad.

Aunque la interrupción del negocio por ciberataques es una preocupación crecientemente reconocida, la aceptación del ciberseguro ha sido relativamente baja. En su último informe, Ponemon Institute and IBM Security (2020) encontró que solo un pequeño porcentaje de empresas están adoptando el ciberseguro. Low (2017) indica que menos del 10% de las empresas del Reino Unido lo tienen. Este número es considerablemente bajo y, por tanto, se espera un importante crecimiento del sector de ciberseguros en los próximos años. Por ejemplo, Lloyd's of London detectó en 2016 un aumento del 50% en las empresas ciberaseguradas y recientemente han introducido 15 tipos diferentes de productos de ciberseguro para cubrir la creciente demanda (Sanchez, 2017). En este marco, se prevé que el mercado de ciberseguros crecerá a un valor total de \$14 mil millones para 2022 (Sharma, 2018).

Una masiva adopción de ciberseguro tiene el potencial de facilitar un reparto de ciber-riesgos a través de los mecanismos de mercado. También tiene el potencial de actuar como incentivo para la inversión en ciberseguridad. La adopción masiva de ciberseguro también podría ayudar a la agregación de datos sobre ataques y ayudar a compartir buenas prácticas y herramientas para evaluar los niveles de ciber-riesgo, elementos que actualmente se echan en falta en el sector del ciberseguro. En resumen, el ciberseguro podría fortalecer la seguridad de los sistemas informáticos de toda la sociedad (Baer and Parkinson, 2007; Kuru and Bayraktar, 2017). Sin embargo, el desarrollo del sector del ciberseguro conlleva riesgos que pueden llegar a comprometer la ciber-resiliencia de nuestra sociedad y mercados digitales.

Entre los potenciales peligros de la adopción del ciberseguro, deben destacarse dos cuestiones. En primer lugar, al adoptar el ciberseguro, existe la preocupación de que los responsables de la toma de decisiones puedan comportarse de manera irracional, lo que les impediría identificar y adquirir un nivel óptimo de cobertura. Dicha irracionalidad puede estar impulsadas por varios factores como la falta de conocimientos sobre seguros y ciberseguridad necesarios para comprender las características de las pólizas de ciberseguro, dificultades para estimar su nivel real de cibervulnerabilidad (mala percepción del riesgo de recibir un ataque y de su potencial impacto) o la presencia de sesgos cognitivos y heurísticas de decisión. En segundo lugar, el mercado de ciberseguros opera en un estado de asimetría de información, donde la aseguradora no tiene acceso a toda la información sobre el nivel de preparación cibernética de una empresa. Si bien las aseguradoras pueden observar algunos elementos como la presencia y características de software antispam, firewalls, planes de recuperación cibernética o sistemas de detección de intrusiones), otros les resultan sin embargo inobservables (nivel de seguridad del comportamiento online de los usuarios). Esta asimetría de información podría derivar en problemas selección adversa o de riesgo moral, con consecuencias nocivas para la resiliencia de los mercados y sociedades digitales. Por ejemplo, el riesgo moral podría tener un efecto perjudicial si la comprar un ciberseguro se redujesen las medidas de protección en ciberseguridad y aumentase el riesgo durante la navegación online.

Nuestra investigación contribuye a llenar un vacío en la literatura sobre ciberseguros al proporcionar y validar modelos conductuales para su adopción. El **Capítulo 1** de esta disertación está dedicado precisamente a comprender y modelizar dicho comportamiento. En concreto, comenzamos analizando la existencia de posibles desviaciones de la racionalidad perfecta, así como identificando las principales características del comportamiento humano en la compra de pólizas de ciberseguro. A continuación, analizamos cómo la adopción del ciberseguro puede afectar al comportamiento en otras dimensiones de la ciberseguridad, como ciberprotección y nivel de seguridad en la navegación online. La primera cuestión, entender el funcionamiento del mecanismo de adopción, es fundamental para el diseño de políticas públicas e intervenciones de las empresas para promover el mercado de ciberseguros. La segunda es fundamental para que los responsables políticos prevean y mitiguen el potencial impacto negativo de un crecimiento masivo de este mercado. Ejemplos de potenciales efectos negativos a identificar, capaces de comprometer la ciber-resiliencia no solo de los agentes individuales sino incluso del Mercado Digital Único, son el riesgo moral que puede hacer que los agentes reduzcan sus niveles de protección o posibles distorsiones en la percepción de su propia cibervulnerabilidad.

La validación empírica se ha realizado mediante un experimento de economía del comportamiento a gran escala, con la participación de 4,800 sujetos en 4 países europeos. En este experimento, se pide a los sujetos que inviertan una dotación inicial en la compra de productos de ciberprotección (capaces de reducir su cibervulnerabilidad) y pólizas de ciberseguro (que cubran las pérdidas derivadas de un potencial ciberataque). Después de eso, se les solicita que naveguen online en un entorno controlado, mientras realizan varias tareas relacionadas con la ciberseguridad (elegir una contraseña, revelar información privada y cerrar sesión, entre otras). Al final del experimento, los sujetos pueden sufrir un ciberataque con una probabilidad que depende de su nivel de protección seleccionado y de la seguridad de su comportamiento online. El pago variable viene dado por la dotación no invertida en productos de ciberseguridad, el valor de sus activos después de posibles pérdidas por un ataque y el reembolso por la póliza de ciberseguro adquirida.

En cuanto al proceso de toma de decisiones, nuestra principal pregunta de investigación se centra en su nivel de racionalidad. Así, nuestro objetivo es validar empíricamente si un modelo de elección racional (basado en la Maximización de la Utilidad Esperada) es capaz de explicar las observaciones sobre el nivel de compra de ciberseguros. Esta pregunta es crítica, ya que, si los modelos de elección racional fueran capaces de explicar los datos observacionales, no quedaría espacio para el desarrollo de modelos de ciberseguridad conductual. Hay que tener en cuenta que, dado que la restricción presupuestaria limita el nivel de protección y cobertura que se puede llegar a adquirir, las decisiones sobre la compra de ambos tipos de productos deben analizarse de forma conjunta. Para validar la racionalidad, y siguiendo la metodología validada presentada en Holt and Laury (2002), calibramos una función de utilidad con Aversión al Riesgo Relativo Constante para cada sujeto, utilizando nuestros datos experimentales. A partir de las utilidades calibradas, podemos determinar cuál es la decisión racional (combinación de protección y cobertura que maximiza la utilidad esperada del tema) para cada participante y compararla con la decisión que realmente tomó en el experimento. Esta comparación nos permite concluir que el modelo de elección racional no es capaz de explicar las decisiones tomadas por los sujetos que, en general, tienden a sobreprotegerse y sobreasegurarse.

La adopción del ciberseguro afecta también a las demás componentes de la ciberseguridad consideradas en esta investigación. En particular, el análisis del Capítulo 1 valida la hipótesis de que los productos de ciberseguro y ciberprotección son complementarios y no sustitutivos. De hecho, los datos experimentales muestran la existencia de dos tipos de agentes en función de su actitud hacia la ciberseguridad. El primer tipo es más sensible a este tema y tiende tanto a sobreasegurarse como a sobreprotegerse, además de comportarse de forma más segura cuando navega en Internet. Por otro lado, el segundo tipo exhibe un peor comportamiento las tres dimensiones de ciberseguridad. Estos resultados, y la falta de poder predictivo del modelo racional de elección, sugieren la conveniencia de desarrollar modelos de economía conductual capaces de predecir la adopción de seguros e identificar y cuantificar el impacto de posibles palancas conductuales que influyan en su compra. Esta tarea se aborda en el Capítulo 2. La evidencia empírica rechaza también que los sujetos con mayor nivel de cobertura se comporten de forma menos segura al navegar por Internet. Por otra parte, el comportamiento online es significativamente más seguro entre los sujetos con un mayor nivel de ciberprotección.

La investigación previa en ciberseguros se ha centrado más en el lado de la oferta. Sin embargo, el lado de la demanda (incluida la formación de creencias sobre cibervulnerabilidad) también debe abordarse para una visión completa del sector (Campbell et al., 2011; Weinstein, 1980). Las creencias sobre el nivel de susceptibilidad a un ataque tienen un impacto directo en la motivación para comportarse de forma más o menos segura (Furnell, 2007). Creencias erróneas en este punto pueden contribuir a la baja aceptación del ciberseguro y la protección identificadas en la literatura (Marotta et al., 2017). En este contexto, analizamos cómo la intencionalidad de un ciberataque (es decir, si el ataque es dirigido a un sujeto de forma estratégica) afecta a las creencias de dicho sujeto. Aunque estas creencias no son observables, su impacto se revela en las diferencias entre los niveles de ciberprotección y ciberseguro ante ataques intencionales (por ejemplo, víctimas seleccionadas intencionalmente por un ciberdelincuente) o no intencionales (por ejemplo, víctimas seleccionadas aleatoriamente por un virus que se propaga al azar a través de Internet). Los resultados del Capítulo 1 muestran que los sujetos se protegen y aseguran más bajo la amenaza de ciber-riesgos aleatorios que ante la amenaza de riesgos intencionales, aunque la probabilidad de sufrir ambos ataques sea la misma. Este comportamiento irracional motiva un análisis más profundo de los métodos de formación y elicitación de creencias sobre riesgo en situaciones adversariales, como la presentada en el Capítulo 3.

Para hacer frente a la falta de capacidad explicativa del Modelo de Elección Racional presentada en el primer capítulo, el **Capítulo 2** propone y valida un modelo predictivo de comportamiento de compra de ciberseguros, incorporando elementos de la Teoría de la Motivación a la Protección (PMT) y de la Teoría de Acción Planeada (TPB). Estos elementos se integran en un Modelo de Ecuaciones Estructurales (SEM), que se calibra utilizando datos del experimento online a gran escala descrito en el Capítulo 1.

Más detalladamente, la PMT propone un modelo de toma de decisión a partir de la evaluación de una amenaza y de cada una de las posibles estrategias que un agente puede utilizar para afrontarla. La evaluación de la amenaza depende tanto de su gravedad percibida (en este caso, un ciberataque), como de la percepción de vulnerabilidad ante ella que tiene el sujeto (en este caso, probabilidad percibida de sufrir el ciberataque). Por otro lado, la evaluación de cada posible estrategia para afrontar la amenaza se basa tanto en eficacia percibida de dicha estrategia (en este caso, la eficacia del ciberseguro), así como en la autopercepción del nivel de capacidad del individuo para gestionar la estrategia (en este caso, su capacidad para comprender y elegir adecuadamente un ciberseguro).

Además del enfoque PMT, nuestro modelo SEM integra elementos de la TPB, incluyendo factores adicionales que pueden influir en la decisión de adquirir un seguro. La TPB considera que la intención de realizar una acción es el mejor predictor de dicha acción. Además, propone que la intención está influida por las actitudes del individuo, las normas subjetivas y su percepción de control de la situación. Esta teoría sugiere que el fortalecimiento de actitudes positivas hacia el ciberseguro (fortaleciendo la creencia de que las compañías de seguros pagarían en caso de un ciberincidente) podría aumentar la aceptación del ciberseguro. Asimismo, fortalecer las normas sociales percibidas en torno al ciberseguro puede ayudar a aumentar la aceptación (por ejemplo, fortalecer la percepción de que otros creen que el ciberseguro es un producto valioso).

Las medidas de comportamiento (adopción de ciberprotección y de ciberseguro, así como la seguridad del comportamiento online) y las variables psicológicas propuestas en la PMT y la TPB se han integrado en el SEM que se muestra en la Figura 1. Se ha seleccionado este método econométrico porque los modelos SEM permiten analizar las relaciones de causalidad entre variables latentes, como las propuestas por la PMT y la TPB. Esta variables latentes no se pueden medir directamente, pero se pueden incluir en el modelo a través de variables observables correlacionadas con ellas. En nuestro caso, las variables observables utilizadas se han obtenido a partir de las respuestas conductuales del experimento y de escalas psicológicas validadas en la literatura.

La estimación del modelo muestra que todos los factores considerados en la TPB y algunos factores en la PMT son buenos predictores de la compra de ciberseguros. La adopción de medidas de seguridad avanzadas también se relaciona de forma positiva con el nivel de seguridad del comportamiento online; aquellos sujetos que adoptaron medidas de seguridad avanzadas también se comportan de forma segura online.

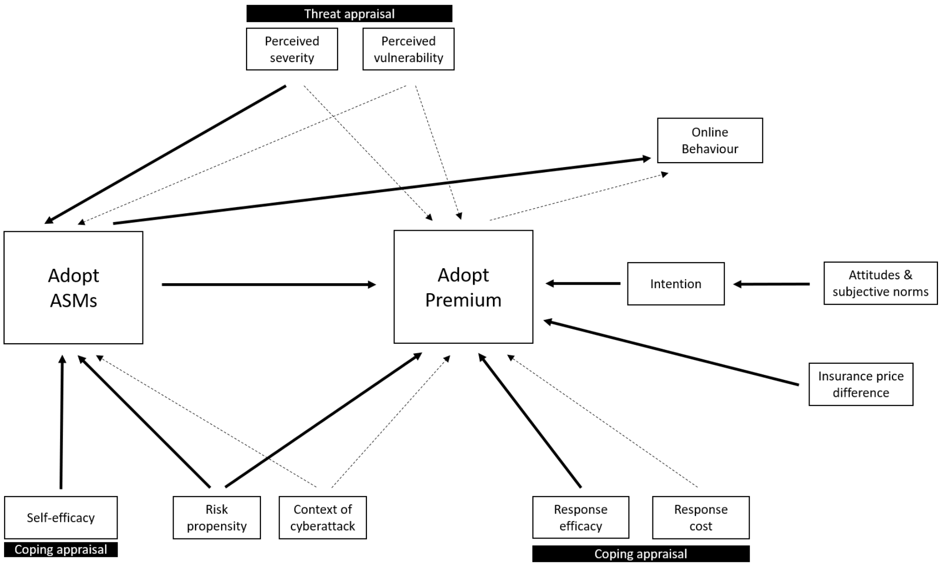


Figure 1: Modelo SEM de la adopción Ciberseguridad and Ciberprotección.

La eficacia de la respuesta y de los factores de la TPB (actitudes y normas) están positivamente relacionados con la adopción de seguros premium; los sujetos que los consideran más efectivos y muestran una actitud positiva hacia ellos, son más propensos a comprar un ciberseguro premium. La autoeficacia y gravedad percibidas de la amenaza influyen positivamente en la adopción de medidas de seguridad avanzadas. Los sujetos con mejor valoración de su capacidad para implementar medidas de ciberseguridad y aquellos que perciben la amenaza del ciberataque como más severa, tienen más probabilidades de adoptar medidas de seguridad avanzadas. Como se menciona anteriormente, la adopción de medidas de seguridad motiva la compra del seguro premium.

Finalmente, las inconsistencias entre lo predicho por los modelos racionales de elección y el comportamiento del agente ante a un ataque intencional motiva un análisis más profundo de los mecanismos de formación de creencias sobre la acción seleccionada por un adversario estratégico. Este análisis se desarrolla en el **Capítulo 3**. Par ello, y relajando el supuesto de conocimiento común, el capítulo propone y valida un enfoque disruptivo que combina modelos de utilidad aleatorios y Análisis de Riesgo Adversarial (ARA) para construir un mecanismo de obtención de probabilidades adversariales complejas (es decir, de la probabilidad de que el adversario estratégico seleccione cada una de sus posibles acciones) de una manera más precisa que en la mera elicitación directa de dichas probabilidades.

El método ARA se introdujo originalmente para tratar problemas de teoría de juegos desde una perspectiva de la teoría de la decisión (Banks et al., 2015). En este marco, un juego se formula de manera bayesiana, como en Kadane and Larkey (1982) y Raiffa (1982), considerando métodos para pronosticar las acciones del adversario sin recurrir al supuesto de conocimiento común. Para describir el enfoque ARA, consideremos una defensora (ella) que despliega controles de ciberseguridad. Habiendo observado la decisión de la defensora, un ciberdelincuente (él) decide si lanzar un ciberataque con resultados inciertos, cuya probabilidad de éxito depende tanto de las decisiones de la defensora como del atacante. Para resolver su problema de decisión, la defensora necesita conocer la función de reacción del atacante. El enfoque ARA propone un enfoque de descomposición para la formación de creencias de la defensora con respecto a la función de reacción del atacante, basado en un análisis desde la perspectiva del atacante. Para ello, la defensora se pone en el lugar del atacante y determina distribución de probabilidad que resume toda la información que el defensor puede obtener sobre las probabilidades de éxito del ataque y las utilidades del atacante. Este enfoque relaja los supuestos de conocimiento común, que son poco realistas, utilizados en teoría de juegos (Hargreaves-Heap and Varoufakis, 2004).

El enfoque ARA estándar considera que ambos agentes son maximizadores de utilidad esperada. El Capítulo 3 propone una extensión de este enfoque al considerar modelos de comportamiento alternativos de toma de decisiones para el atacante, basados en la teoría de la utilidad aleatoria (Marschak, 1959; Block et al., 1959; McFadden, 1973). En estos modelos, la utilidad del decisor está sujeta a shocks aleatorios, interpretados como el resultado de sesgos cognitivos, heurísticas de decisión o errores de implementación de los agentes. (Hess et al., 2018). La maximización de la utilidad junto con la distribución de los shocks aleatorios conduce a una regla probabilística para la elección de una acción del atacante. Hay que tener en cuenta que aunque el enfoque ARA y los modelos de utilidad aleatoria se traducen en una regla probabilística para la selección de acciones, la fuente de incertidumbre es diferente en cada caso: (i) En ARA, la incertidumbre sobre las acciones seleccionadas por el adversario es consecuencia de las creencias probabilísticas del defensor sobre las posibilidades de éxito del ataque y la utilidad del atacante; sin embrago, (ii) en los modelos de utilidad aleatoria, la utilidad de cada agente es intrínsecamente incierta, y esta aleatoriedad se convierte en la fuente de incertidumbre para la acción seleccionada, incluso bajo supuestos de conocimiento común.

En el Capítulo 3, proponemos la integración de reglas de elección probabilística provenientes del enfoque de utilidad aleatoria con modelos ARA para definir lo que denominamos métodos de *recomposición conductual de creencias.* La precisión de estas técnicas conductuales de recomposición se compara con las obtenidas con los métodos estándar de ARA con maximización de utilidad esperada y de elicitación directa. Utilizando datos de un experimento económico presencial, mostramos que (de los tres enfoques considerados), la recomposición conductual es el método más efectivo para elicitar creencias. La evidencia empírica del experimento también sugiere que el proceso de reflexión requerido para hacer explícitas las creencias de los agentes no mejora la precisión de una elicitación directa.

Los resultados del Capítulo 3 destacan la gran carga cognitiva y el nivel de complejidad necesarios para pronosticar acciones adversariales en ciberseguridad. La formación de creencias en contextos adversariales es una tarea a desarrollar por el Sistema 2 (consciente y analítico) en el modelo de pensamiento dual de Kahneman (Kahneman, 2012). Dado que en la mayoría de los casos el Sistema 2 es sustituido por el Sistema 1 (rápido y automático), el resultado es que la elicitación directa no llega a considerar los elementos estratégicos del problema y proporciona peores estimaciones. Así, una forma más efectiva de abordar este problema es el uso de métodos de elicitación directos para los elementos básicos del problema de decisión y externalizar la tarea que hubiese correspondido a un perezoso Sistema 2.

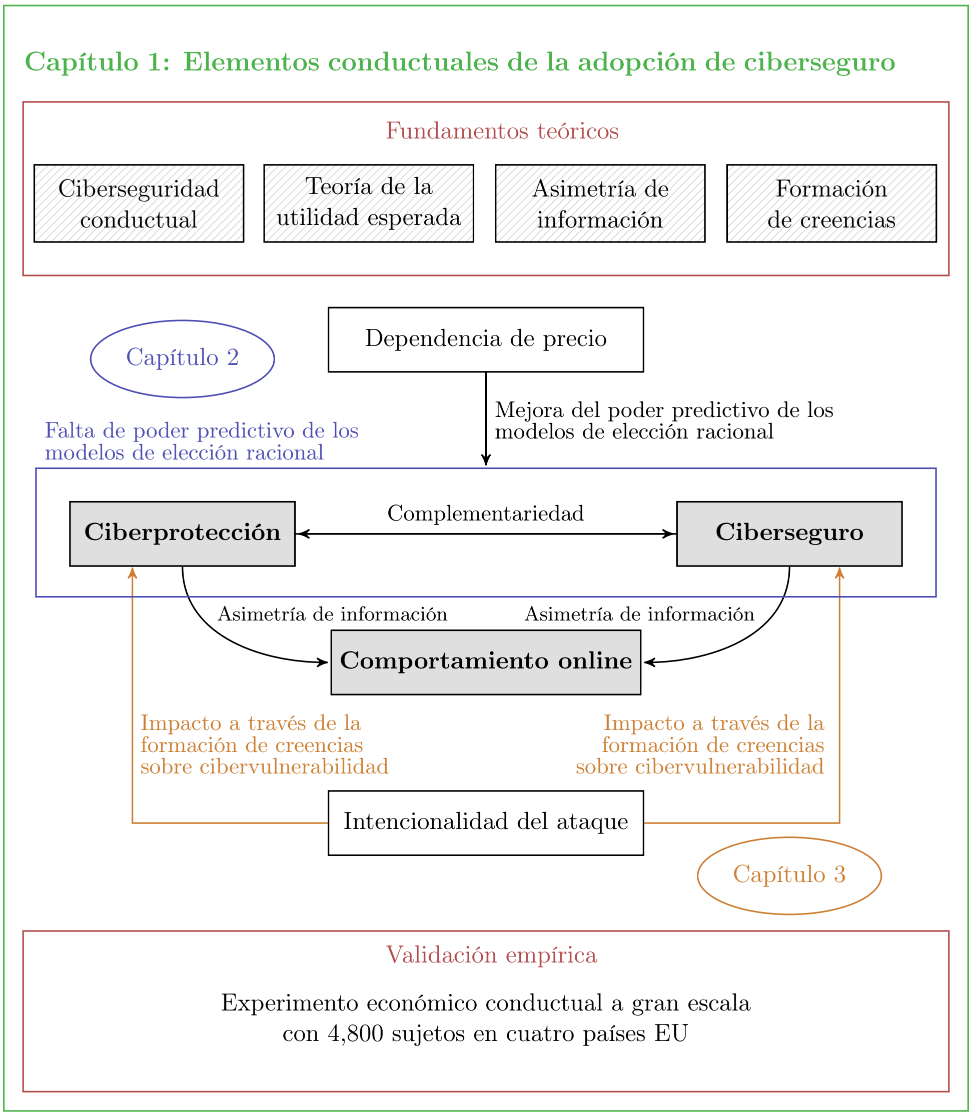
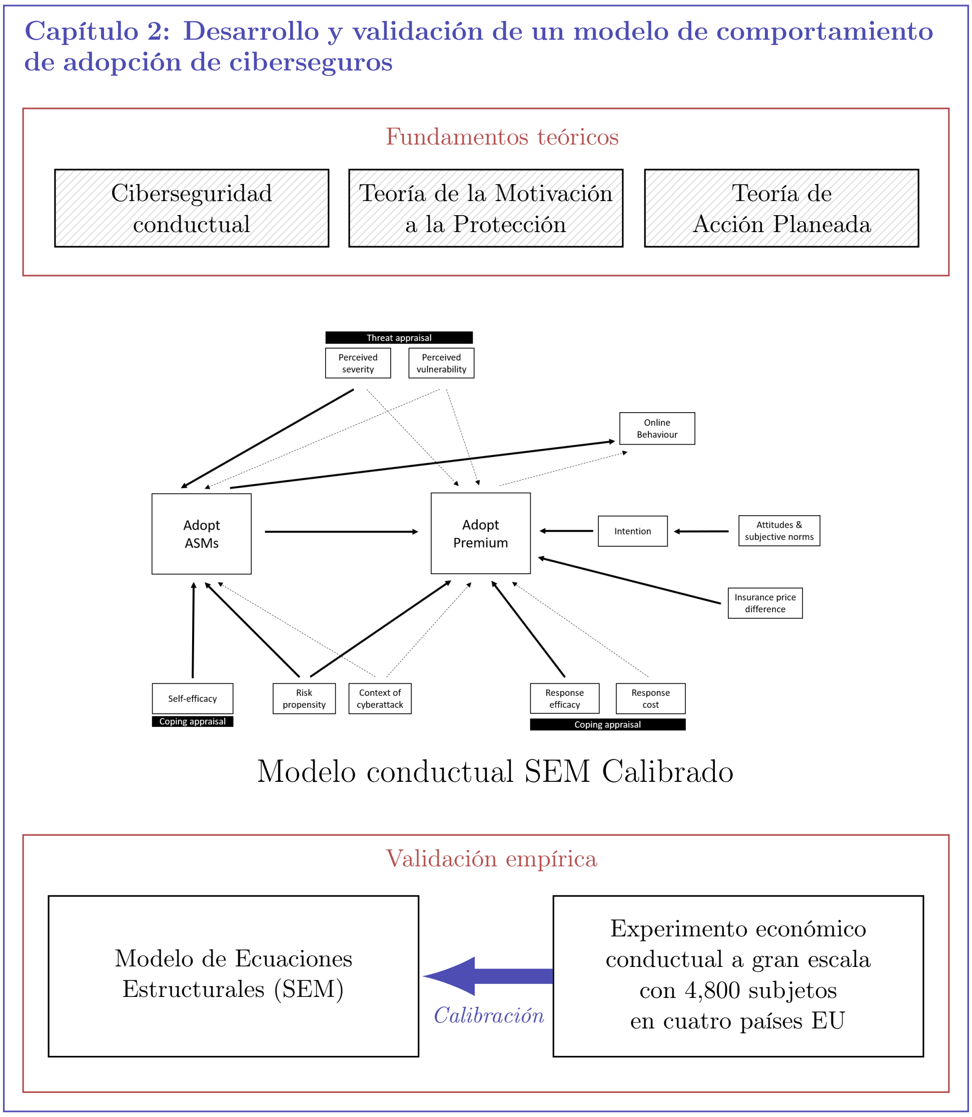

Capítulo 1: Elementos conductuales de la adopción de ciberseguro
Fundamentos teóricos
Ciberseguridad conductual
Teoría de la utilidad esperada
Asimetría de información
Formación de creencias
Capítulo 2
Dependencia de precio
Mejora del poder predictivo de los modelos de elección racional
Falta de poder predictivo de los modelos de elección racional
Ciberprotección
Complementariedad
Ciberseguro
Asimetría de información
Asimetría de información
Comportamiento online
Impacto a través de la formación de creencias sobre cibervulnerabilidad
Impacto a través de la formación de creencias sobre cibervulnerabilidad
Intencionalidad del ataque
Capítulo 3
Validación empírica
Experimento económico conductual a gran escala con 4,800 sujetos en cuatro países EU

Figure 2: Mapa conceptual del Capítulo 1.



Capítulo 2: Desarrollo y validación de un modelo de comportamiento de adopción de ciberseguros

Fundamentos teóricos

Ciberseguridad conductual

Teoría de la Motivación a la Protección

Teoría de Acción Planeada

Modelo conductual SEM Calibrado

Validación empírica

Modelo de Ecuaciones Estructurales (SEM)

Calibración

Experimento económico conductual a gran escala con 4,800 subjetos en cuatro países EU

Figure 3: Mapa conceptual del Capítulo 2.

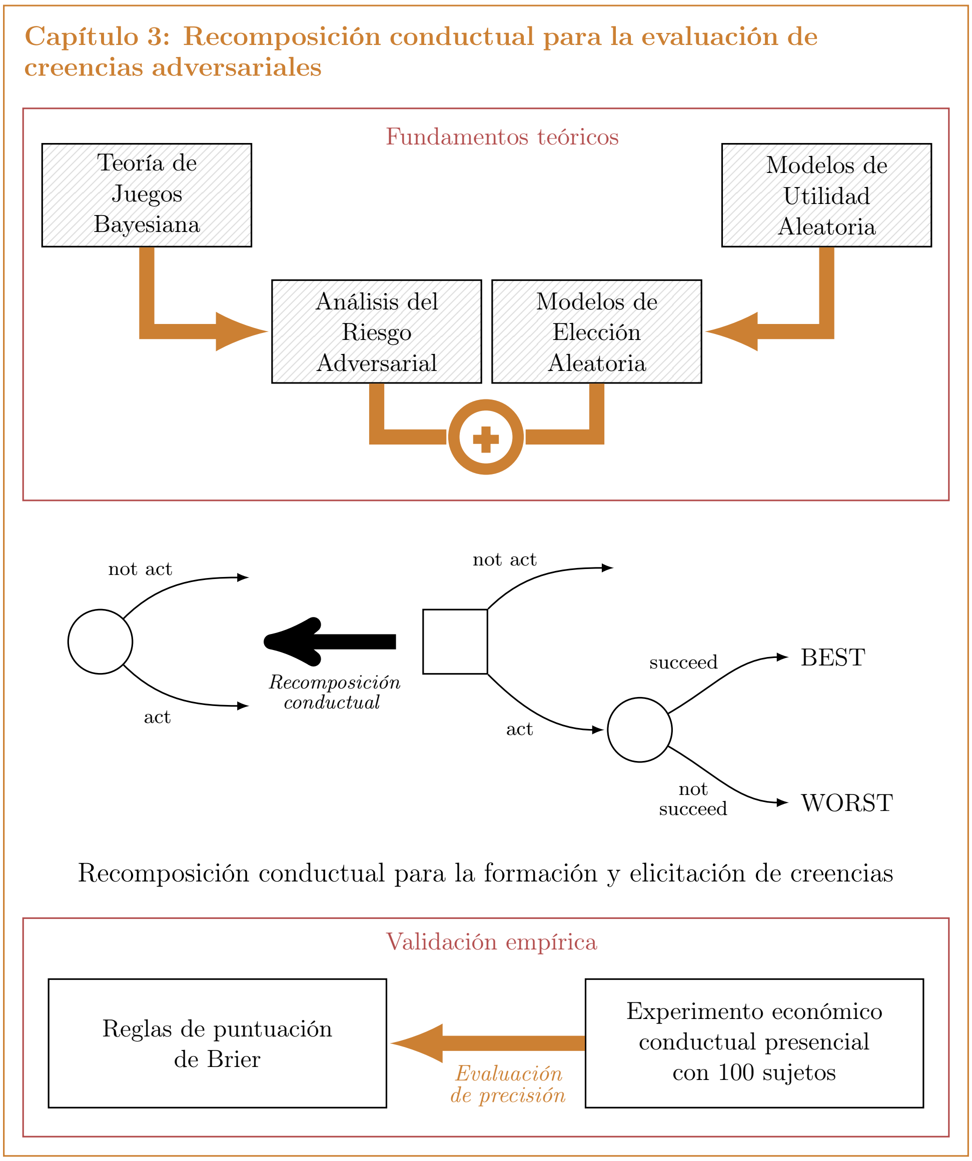

Capítulo 3: Recomposición conductual para la evaluación de creencias adversariales
Fundamentos teóricos
Teoría de Juegos Bayesiana
Modelos de Utilidad Aleatoria
Análisis del Riesgo Adversarial
Modelos de Elección Aleatoria
not act
act
Recomposición conductual
not act
act
succeed
BEST
not succeed
WORST
Recomposición conductual para la formación y elicitación de creencias
Validación empírica
Reglas de puntuación de Brier
Evaluación de precisión
Experimento económico conductual presencial con 100 sujetos

Figure 4: Mapa conceptual del Capítulo 3.

References

Baer, W. S. and Parkinson, A. (2007). Cyberinsurance in it security management. *IEEE Security & Privacy*, 5(3):50–56.

Banks, D. L., Aliaga, J. M. R., and Insua, D. R. (2015). *Adversarial risk analysis*. CRC Press.

Berr, J. (2017). Wannacry ransomware attack losses could reach $4 billion. `https://www.cbsnews.com/news/wannacry-ransomware-attacks-wannacry-virus-losses/`.

Block, H. D., Marschak, J., et al. (1959). Random orderings and stochastic theories of response. Technical report, Cowles Foundation for Research in Economics, Yale University.

Campbell, J., Ma, W., and Kleeman, D. (2011). Impact of restrictive composition policy on user password choices. *Behaviour & Information Technology*, 30(3):379–388.

Furnell, S. (2007). An assessment of website password practices. *Computers & Security*, 26(7-8):445–451.

Greenberg, A. (2018). The untold story of notpetya, the most devastating cyberattack in history.

Hargreaves-Heap, S. and Varoufakis, Y. (2004). *Game theory: A critical introduction*. Routledge.

Hess, S., Daly, A., and Batley, R. (2018). Revisiting consistency with random utility maximisation: theory and implications for practical work. *Theory and Decision*, 84(2):181–204.

Holt, C. A. and Laury, S. K. (2002). Risk aversion and incentive effects. *American economic review*, 92(5):1644–1655.

Kadane, J. B. and Larkey, P. D. (1982). Subjective probability and the theory of games. *Management Science*, 28(2):113–120.

Kahneman, D. (2012). *Think Fast and Slow.* Penguin. London.

Kuru, D. and Bayraktar, S. (2017). The effect of cyber-risk insurance to social welfare. *Journal of Financial Crime.*

Low, P. (2017). Insuring against cyber-attacks. *Computer Fraud & Security*, 2017(4):18–20.

Marotta, A., Martinelli, F., Nanni, S., Orlando, A., and Yautsiukhin, A. (2017). Cyber-insurance survey. *Computer Science Review*, 24:35–61.

Marschak, J. (1959). Binary choice constraints and random utility indicators. In *K. Arrow (Ed.), Stanford symposium on mathematical models in the social sciences.* Stanford, CA: Stanford University Press.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In *P. Zarembka (Ed.), Frontiers in econometrics.*, pages 2373–2375. New York: Academic Press.

Ponemon Institute and IBM Security (2020). Cost of a data breach report.

Raiffa, H. (1982). *The art and science of negotiation.* Harvard University Press.

Sanchez, A. (2017). Lloyd's predicts surge in cyber insurance uptake in 2017. `http://www.insurancebusinessmag.com/uk/news/breaking-news/lloyds-predicts-surge-incyber-insurance-uptake-in-2017-42266.aspx`.

Sharma, Y. (2018). Cyber insurance market to reach $14 billion, globally, by 2022.

Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of personality and social psychology*, 39(5):806.

World Economic Forum (2020). The global risks report.

Introduction

This research work aims at developing and experimentally validate theoretically-sound behavioural models capable to explain and foresee the adoption of cyberinsurance and related human cybersecurity behaviour. Specifically, this dissertation focus on three critical and interrelated dimensions of cybersecurity: *cyberinsurance* (adoption of insurance products partially covering the impact of potential attacks), *cyberprotection* (adoption of measures able to reduce the risk level of suffering an attack) and *online behaviour* (level of cyber-risk assumed by users when navigating online). Such dimensions take into the game most of the relevant behavioural issues related to cyberinsurance adoption, such as (i) the rationality of the allocation of the available cybersecurity budget between the adoption of protection and insurance products, (ii) the potential negative effect coming from the information asymmetry intrinsic to any field of insurance (including cyberinsurance) and; (iii) belief formation on cybervulnerability, especially on risk perception and risk assessment methods in case of intentional attacks. By achieving this objective, our research contributes to fill the critical existing gap on how agents do actually make their decisions on cyberinsurance adoption and form their perceptions on their own cyber-risks, which has relevant scientific as well as policy-making and business development role as shown in the discussion section of this dissertation.

The theoretical foundations or our work take advantage from a multidisciplinary approach, getting together elements from the latest research in the field of cyberinsurance and integrating them with (i) economic elements from rational choice, random utility and behavioural choice models; (ii) mechanisms of belief formation without the strong assumption of common knowledge in strategic interactions and (iii) psychological levers conditioning cybersecurity decision-making. The empirical validation of the models has been performed applying behavioural-experimental economics approach. Experimental data are collected from a large-scale online experiment with 4,800 subjects run in 4 European countries (Germany, Poland, Spain and United Kingdom) as described in Chapter 1 and a face-to-face laboratory experiment run in Spain with 100 subjects as described in Chapter 3.

Cybersecurity is increasingly perceived as a major global problem (World Economic Forum, 2020) becoming even more relevant as companies, administrations and individuals get more interconnected, thereby increasing opportunities for, and facilitating the spread of, cyberthreats. Three widely recognised examples of cyberattacks include: (i) the 2017 Equifax breach, which resulted in stolen data from over 140 million customers – including social security and credit card numbers; (ii) the 2017 WannaCry attack, which took down the UK National Health Service, Telefonica, and FedEx, among others, severely hindering their operations and entailing losses estimated to have reached $4 billion (Berr, 2017); and (iii) the NotPetya malware, which affected thousands of organisations worldwide with an estimated cost of $10 billion (Greenberg, 2018). Whilst organisations typically employ technical security solutions, including antivirus software, firewalls and intrusion detection systems, cybersecurity failures are often attributed to the human component of cybersecurity, which has been previously described as the weakest link in the cybersecurity chain. In this context, behavioural change arises as a critical strategy to improve cyber preparedness. In addition to malicious, targeted attacks (which included social engineering and phishing), human error accounted for a further 24% of breaches in the latest Ponemon Institute and IBM Security (2020) report. To address the human component of cybersecurity we need to understand the rational and non-rational (although predictable) factors that affect behaviour and belief formation, specifically through the application of theoretically sound and empirically validated models.

Although business disruption from cyberattacks is a recognised and growing concern, the uptake of cyberinsurance has been relatively low. In their latest report, Ponemon Institute and IBM Security (2020) found that only a small percentage of companies are adopting cyberinsurance. Low (2017) found that less than 10% of UK companies report holding cyberinsurance. This number is considerably low and, therefore, an important growth of the cyberinsurance sector is expected in the coming years. For instance, Lloyd's of London reported an increase in uptake of 50% in 2016 and they have recently introduced 15 different types of cyberinsurance products for a predicted boom in uptake (Sanchez, 2017). In this context, the cyberinsurance market is predicted to grow to a total value of $14 billion by 2022 (Sharma, 2018).

If widely adopted and well-functioning, cyberinsurance has the potential to encourage market-based risk management for information security, with a mechanism for spreading the risk across multiple stakeholders. It also has the potential to act as an incentive towards organisational investments in information security; which would reduce risk for the investing organisation and for their wider network. Cyberinsurance uptake could also lead to data aggregation on best practices and better tools for assessing security – something that is currently lacking in relation to cyberinsurance. In summary, cyberinsurance could strength IT security for society as a whole (Baer and Parkinson, 2007; Kuru and Bayraktar, 2017). However, a cyberinsurance rocketing is not free of risks and may even compromise the cyber-resilience not only of individual agents but also that of our digital society and markets.

Among the potential dangers of cyberinsurance adoption, two critical issues should be highlighted. Firstly, when adopting cyberinsurance, there are some concerns that decision-makers may behave in an irrational way, resulting in them being unable to identify and/or purchase their optimal level of cybersecurity coverage. Irrational decisions may be driven by several factors including a lack of insurance and/or cybersecurity literacy required to understand the features of cyberinsurance policies, failing to estimate their actual level of cybervulnerability (i.e., in terms of their risk of receiving an attack and/or their perceptions of an attack's potential impact), or the presence of cognitive biases or decision heuristics affecting the decision procedure. Secondly, the cyberinsurance market operates in a state of information asymmetry, where the insurer does not have access to all information regarding a company's cyber-preparedness level. Although some elements of the cybersecurity position of the company can be observed by insurers risk audits (such as the existence and features of anti-spam software, firewalls and cyber-recovery plans or intrusion-detection systems), others cannot (such as the safety level of human online behaviour). This information asymmetry could potentially result in adverse selection and/or moral hazard issues, which can have dramatic consequences for the resilience of digital markets and societies. For instance, moral hazard could have a hugely detrimental effect if it resulted in those who adopted cyberinsurance showing a significant reduction in cybersecurity protection measures and/or an increase in risky online behaviour.

Our research contributes to fill a gap in the cyberinsurance literature by providing and validating behavioural models of ciberinsurance adoption. **Chapter 1** of this dissertation is devoted to understanding and modelling critical behavioural insights in the process of cyberinsurance adoption. Specifically, we start by analysing potential deviations from perfect rationality and the main behavioural features in the purchase of cyberinsurance polices. After that, we analyse how the adoption of cyberinsurace may affect agents' behaviours in other dimensions of their cybersecurity strategy, such as cyberprotection and safety level when navigating online. The first question, understanding the mechanism of adoption, is critical to support policy-making and industry decisions to promote the existing underdeveloped cyberinsurance market. The second one is fundamental for policy-makers to foresee and mitigate potential negative impact of the growth of this market. Examples of potential negative effects to be identified, capable to compromise the cyberesiliance not only of individual agents but even that of the Unique Digital Market, are the presence of a moral hazard effect making agents to reduce their protection and navigation safety levels or potential distortions in the beliefs of their own cybervulnerability that may be induced by the context in which an attack may take place.

The empirical validation has been done using a large-scale behavioural economics experiment with the participation 4,800 subjects in 4 European countries. In this experiment, subjects are asked to invest their initial endowment in purchasing cyberprotection products (capable to reduce their cybervulnerability) and a cyberinsurance policies (covering the losses coming from a potential cyberattack). After that, subjects are asked to navigate online in a controlled environment, while performing several cybersecurity-related tasks (choosing a password, revelling private information and logging out, among others). At the end of the experiment, subjects may suffer a cyberattack with a probability depending on their selected protection level and the safety of their online behaviour. Variable payment is given by the endowment not invested in cybersecurity products, the value of their assets after potential losses from an attack and the payback of the cyberinsurance policy (if previously purchased).

As regards with the decision-making process, our main research question focuses on its level of rationality. Specifically, we aim at testing empirically if a rational choice model (based in Expected Utility Maximisation) is capable to explain actual cyberinsurance adoption. This question is critical in our research, since if rational choice models were able to explain observational data, no room would be left for the development of behavioural cybersecurity models. Note that, since the budget constrain limits the level of protection and coverage to be purchased, the decisions on the purchase of both types of products must be analysed together. To test rationality, and following the validated methodology presented in Holt and Laury (2002), we calibrate a Constant Relative Risk Aversion utility function for each subject using our experimental data. Using the calibrated utilities, we can compute which is the perfectly rational decision for this subject (i.e. the combination of protection and coverage that maximise the expected utility of the subject) and compare it with the decision actually made in the experiment. This comparison shows that the rational choice model fails to explain the decisions made by the subjects and, in general, subjects tend to overprotect and overinsure themselves for the case of a potential cyberattack.

The adoption of cyberinsurance is no free of effects on the other components of cybersecurity. In particular, the analysis in Chapter 1 validates the research hypothesis stating that cyberinsurance and cyberprotection products are complementary and no substitutive. In fact, experimental data show the existence of two clear types of agents in terms of their attitude towards security. The first type seems to be concern on security and tends both to overinsure and overprotect herself (as well as behaving more safely when navigating online). On the other hand, the second type exhibits a lousier behaviour in all the three cybersecurity dimensions. These results, and the failure of predictive power of the rational choice model, suggests the convenience to develop new models with a sound behavioural foundation capable to explain insurance adoption, as well as identify and quantify the impact of potential irrational levers conducting to cyberinsurance purchase. This task will be approached in Chapter 2. The empirical evidence supports also rejecting that subjects with a higher level of coverage behave in a less safe way when navigating on the Internet.

oreover, online behaviour is significantly safer among subjects with a higher cyberprotection level.

The limited research into cyberinsurance has tended to focus upon the supply side of insurability, however the demand side (including formation of risk beliefs on cybervulnerability) is also vital (Campbell et al., 2011; Weinstein, 1980). Individuals' beliefs about their own susceptibility to an attack directly impact upon their motivation to behave securely (Furnell, 2007), meanwhile inaccurate risk beliefs may contribute to low uptake of cyberinsurance and protection (Marotta et al., 2017). In this context, we analyse how the intentionality of a cyberattack (i.e., whether the attack was intentionally targeted on their specific business) affects the beliefs of the agents. Although beliefs are not observable, this impact can be revealed by the differences in their cyberprotection and cyberinsurance uptake under the presence of intentional (e.g., victims intentionally selected by a cybercriminal) or unintentional (e.g., victims are random, such as a virus spreading randomly through the internet) attacks. Results in Chapter 1 shows that subjects protect and ensure themselves more under the menace of random cyber-risks than under that of intentional ones, although the probability of suffering such an attack was the same in both contexts. This irrational behaviour suggests the interest of a deeper analysis of risk belief formation and elicitation methods in adversarial situations, as presented in Chapter 3.

To cope with the lack of capacity of the Rational Choice Model to explain the empirical evidence presented in the first chapter, **Chapter 2** proposes and validates a behavioural predictive model of cyberinsurance adoption, incorporating elements of Protection Motivation Theory (PMT) and the Theory of Planned Behaviour (TPB). Such elements are integrated in a Structural Equation Modelling (SEM), which is calibrated using data from the large-scale online experiment described in Chapter 1.

In more detail, PMT proposes that people protect themselves by making both a threat and a coping appraisal. The threat appraisal is dependent upon both the perceived severity of a threatening event (the cyberattack) and their perceived vulnerability to the event (perceived probability of that event occurring). The coping appraisal reflects the perceived efficacy of the recommended protective behaviour (cyberinsurance adoption) and the individual's perceived self-efficacy (ability to understand and properly purchasing cyberinsurance products). In PMT, the threat of a particular behaviour is weighed up against the rewards of that behaviour and the costs of the coping action are also a factor. Therefore, in our application of PMT, an individual considering whether to invest in cyberinsurance may firstly weigh up the likelihood that they will receive a cyberattack of a particular severity against (a) The cost of taking out cyberinsurance (finances, time, effort) and (b) How effective they believe that insurance will be and/or how much confidence they have in their own ability to put insurance measures into place.

In addition to PMT, our SEM model integrates elements from the TPB, which taps into one of the same constructs as PMT (as perceived self-efficacy and perceived behavioural control are thought to measure the same construct). However, TPB also highlights additional factors which may influence the decision whether to purchase insurance. TPB states that intention to perform a behaviour is the most immediate and important determination of behaviour. Intention is influenced by the individual's attitudes towards the behaviour, subjective norm, and perceived control over the situation. This theory suggests that strengthening positive attitudes towards cyberinsurance (strengthening the belief that insurance companies would pay out in the event of a cyber-incident) could increase cyberinsurance uptake. Likewise strengthening the perceived social norms around cyberinsurance may help to increase uptake (e.g., strengthening the perception that others believe cyberinsurance to be a worthwhile product).

The behavioural measures (cyberprotection and cyberinsurance adoption, as well as safety of online behaviour) and psychological variables from PMT and TPB have been integrated in the SEM shown in Figure 5. SEM models allow the analysis of causality relations among latent variables. Latent variables cannot be measured directly but can be observed through other correlated measurable variables obtained from behavioural measures from the experiment and validated psychological scales.

The results of the estimation of the model show that all TPB factors, but only some PMT factors positively predicted adoption of premium cyberinsurance. Specifically, model shows a significant positive pathway that links the adoption of advanced security measures to the adoption of premium insurance. The adoption of advanced security measures was also significantly positively related to security of online behaviour; those who adopted advanced security measures were also more likely to behave securely online. The pathway between insurance adoption and online behaviour, although positive, fails to

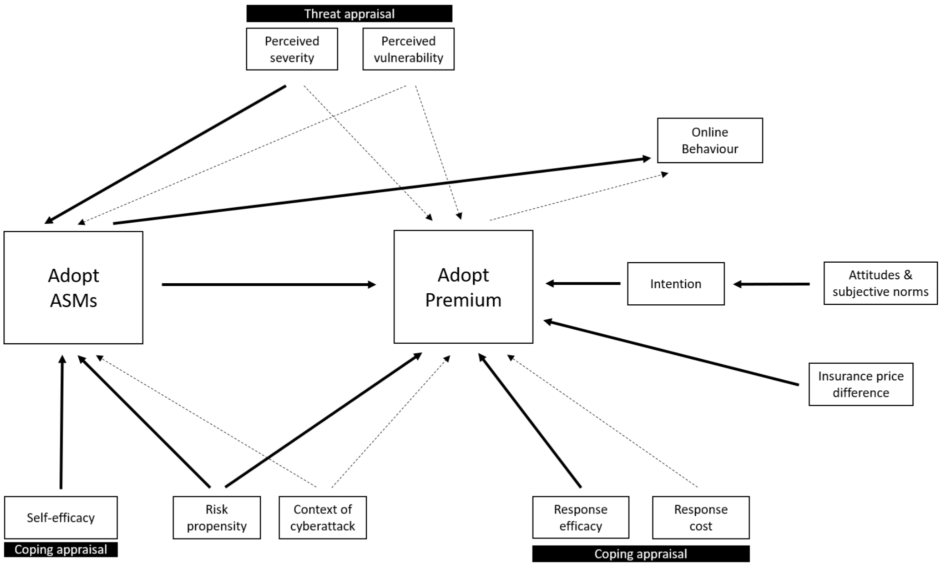


Figure 5: SEM Model of Cyberinsurance and Cyberprotection Adoption.

reach significance once adoption of security measures was introduced into the model.

Response efficacy and the TPB factors (attitudes and norms) are positively related to adoption of premium insurance; those who perceived insurance to be more effective, and those who had positive attitudes and positive subjective norms, were more likely to adopt premium cyberinsurance. Perceived self-efficacy and perceived threat severity both positively fed into the adoption of advanced security measures rather than adoption of premium insurance directly. Those who had higher perceptions of their ability to put cybersecurity measures into place, and those who perceived the threat of the cyberattack as more severe, were more likely to adopt advanced security measures. As aforementioned, the adoption of security measures then subsequently fed into premium insurance adoption.

Finally, the inconsistences between the rational choice models prediction and agent's behaviour in front of an intentional attack, motivates the analysis of the mechanisms of belief formation and elicitation methods on the action to be selected by a strategic adversary, such as a cybercriminal. This analysis is undertaken in **Chapter 3**. To this end, and relaxing the assumption of common knowledge, the chapter proposes and validates a disruptive approach combining both Adversarial Risk Analysis (ARA) and Random Utility Models to build a mechanism of elicitation of complex adversarial probabilities (i.e., the probability of a strategic adversary to select an action) in a more accurate way than just a direct elicitation of such probabilities.

ARA was originally introduced to deal with game theoretic problems studied from a decision analytic perspective (Banks et al., 2015). Games are formulated in a Bayesian manner, as in Kadane and Larkey (1982) and Raiffa (1982) and operationalised through the provision of procedures to forecast the actions of the adversary with the aim of mitigating common knowledge assumptions standard in game theory. Imagine a defender (she) which deploys cybersecurity controls. Then, having observed what the defender has deployed, a cybercriminal attacker (he) decides whether to launch a cyberattack with uncertain results, whose probability of success depends on both the defender and the attacker's decisions. To solve her decision problem, the defender needs to know the reaction function of the attacker. ARA usefully suggests a decomposition approach for the formation of the defender's belief respect to the attacker's reaction function based on analysing the problem from the attacker's perspective. To this end, the defender puts herself in the attacker's shoes, considering the probability distribution summarising all the information she can obtain about the probabilities of the attack to succeed and attacker's utilities. This approach weakens the standard, but unrealistic, common knowledge assumptions in game theoretic approaches (Hargreaves-Heap and Varoufakis, 2004). In the ARA approach, the defender has uncertainty about the information known by the attacker.

The standard ARA approach assumes that both agents are expected utility maximisers. Chapter 3 proposes an extension of this approach by considered alternative behavioural models of decision making for the attacker based on Random Utilility Theory (Marschak, 1959; Block et al., 1959; McFadden, 1973). In these models, the decision maker's utility is subject to random shocks, typically interpreted as the result of agents' cognitive biases, decision heuristics or implementation errors (Hess et al., 2018). Utility maximization together with the distribution of shocks leads to a probabilistic criterion for a particular alternative to be selected over the others. Note that although ARA and random utility modelling translate into a probabilistic rule for action selection, the source of the uncertainty is different in each model approaches: (i) In ARA, uncertainty about the actions selected by the adversary is a consequence of the probabilistic beliefs of the defender on the chances of success of the attack and the attacker's utility; Meanwhile, (ii) in random utility models the utility of each agent is intrinsically uncertain, its randomness becoming the source of uncertainty for the selected action, even under common knowledge assumptions.

In Chapter 3, we propose the integration of probabilistic choice rules coming from the random utility approach into ARA models to define what we named as *behavioural recomposition* methods. The accuracy of these two behavioural recompositions techniques is tested against those obtained with standard expected utility ARA and direct elicitation methods. Using data from a face-to-face behavioural economics experiment, we show that (out of the three approaches considered), behavioural recomposition is the most accurate method belief elicitation in strategy setting. The empirical evidence from the experiment does also suggests that the reflection process required to make explicit the agents' beliefs on the probabilities of success of the attack and the adversary's utilities does not improve the accuracy of direct elicitation.

The results in Chapter 3 make explciit the large cognitive burden and complexity level required to forecast adversarial actions in cybersecurity. Therefore, adversarial beliefs formation seems to be a task for the conscious and analytic System 2 in Kahneman's dual thinking model (Kahneman, 2012). Since in most of the cases System 2 is overcome by the fast and automatic System 1, the result is that direct beliefs elicitation does not take into account the strategic aspects, providing then lousier estimations. An effective way to address this issue may be to rely on direct elicitation methods just for the belief on basic elements and perform externally the task that would have correspond to the lazy System 2.

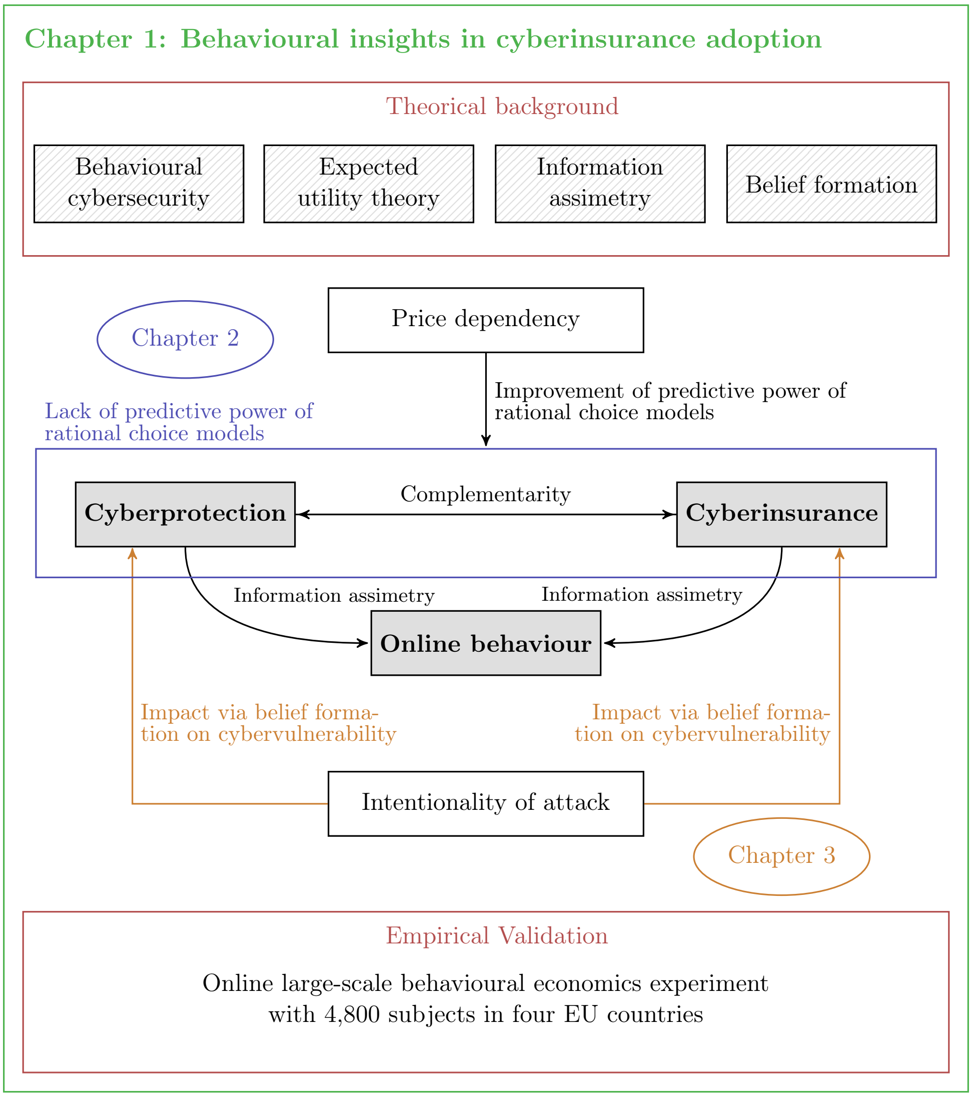

Chapter 1: Behavioural insights in cyberinsurance adoption
Theorical background
Behavioural cybersecurity
Expected utility theory
Information assimetry
Belief formation
Price dependency
Chapter 2
Improvement of predictive power of rational choice models
Lack of predictive power of rational choice models
Cyberprotection
Complementarity
Cyberinsurance
Information assimetry
Information assimetry
Online behaviour
Impact via belief formation on cybervulnerability
Impact via belief formation on cybervulnerability
Intentionality of attack
Chapter 3
Empirical Validation
Online large-scale behavioural economics experiment with 4,800 subjects in four EU countries

Figure 6: Conceptual map of Chapter 1.

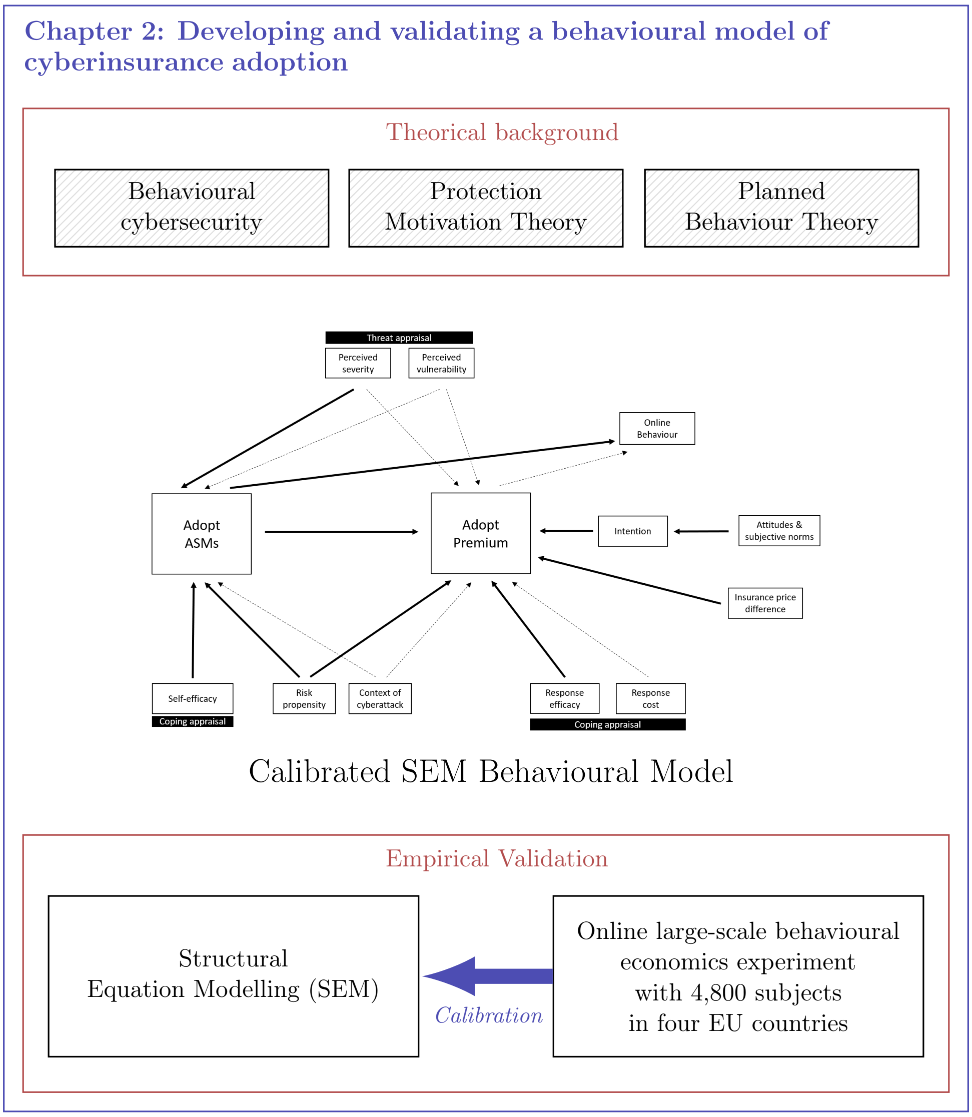

Chapter 2: Developing and validating a behavioural model of cyberinsurance adoption
Theorical background
Behavioural cybersecurity
Protection Motivation Theory
Planned Behaviour Theory
Threat appraisal
Perceived severity
Perceived vulnerability
Online Behaviour
Adopt ASMs
Adopt Premium
Intention
Attitudes & subjective norms
Insurance price difference
Self-efficacy
Coping appraisal
Risk propensity
Context of cyberattack
Response efficacy
Response cost
Coping appraisal
Calibrated SEM Behavioural Model
Empirical Validation
Structural Equation Modelling (SEM)
Calibration
Online large-scale behavioural economics experiment with 4,800 subjects in four EU countries

Figure 7: Conceptual map of Chapter 2.

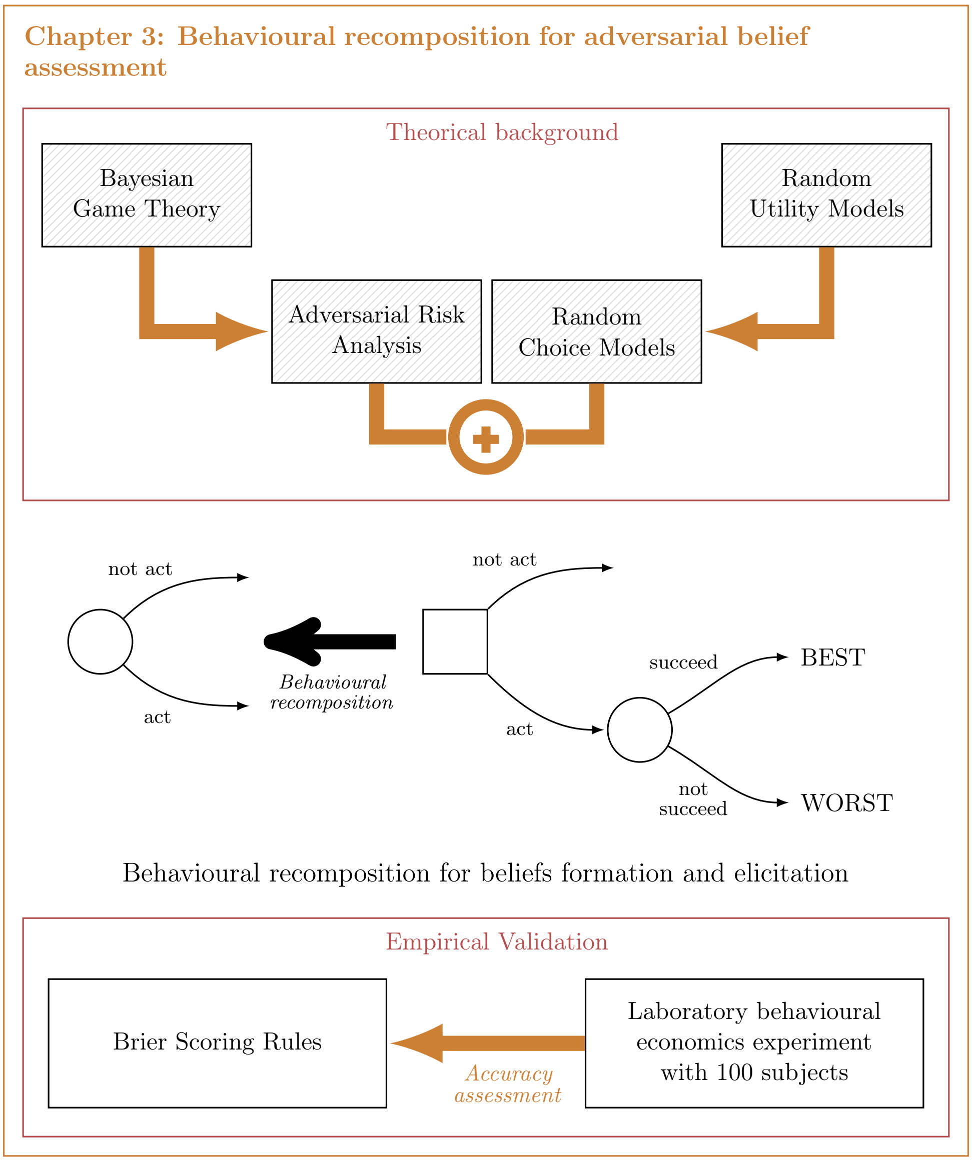

Chapter 3: Behavioural recomposition for adversarial belief assessment
Theorical background
Bayesian Game Theory
Random Utility Models
Adversarial Risk Analysis
Random Choice Models
not act
act
Behavioural recomposition
not act
act
succeed
BEST
not succeed
WORST
Behavioural recomposition for beliefs formation and elicitation
Empirical Validation
Brier Scoring Rules
Accuracy assessment
Laboratory behavioural economics experiment with 100 subjects

Figure 8: Conceptual map of Chapter 3.

References

Baer, W. S. and Parkinson, A. (2007). Cyberinsurance in it security management. *IEEE Security & Privacy*, 5(3):50–56.

Banks, D. L., Aliaga, J. M. R., and Insua, D. R. (2015). *Adversarial risk analysis*. CRC Press.

Berr, J. (2017). Wannacry ransomware attack losses could reach $4 billion. `https://www.cbsnews.com/news/wannacry-ransomware-attacks-wannacry-virus-losses/`.

Block, H. D., Marschak, J., et al. (1959). Random orderings and stochastic theories of response. Technical report, Cowles Foundation for Research in Economics, Yale University.

Campbell, J., Ma, W., and Kleeman, D. (2011). Impact of restrictive composition policy on user password choices. *Behaviour & Information Technology*, 30(3):379–388.

Furnell, S. (2007). An assessment of website password practices. *Computers & Security*, 26(7-8):445–451.

Greenberg, A. (2018). The untold story of notpetya, the most devastating cyberattack in history.

Hargreaves-Heap, S. and Varoufakis, Y. (2004). *Game theory: A critical introduction*. Routledge.

Hess, S., Daly, A., and Batley, R. (2018). Revisiting consistency with random utility maximisation: theory and implications for practical work. *Theory and Decision*, 84(2):181–204.

Holt, C. A. and Laury, S. K. (2002). Risk aversion and incentive effects. *American economic review*, 92(5):1644–1655.

Kadane, J. B. and Larkey, P. D. (1982). Subjective probability and the theory of games. *Management Science*, 28(2):113–120.

Kahneman, D. (2012). *Think Fast and Slow.* Penguin. London.

Kuru, D. and Bayraktar, S. (2017). The effect of cyber-risk insurance to social welfare. *Journal of Financial Crime.*

Low, P. (2017). Insuring against cyber-attacks. *Computer Fraud & Security*, 2017(4):18–20.

Marotta, A., Martinelli, F., Nanni, S., Orlando, A., and Yautsiukhin, A. (2017). Cyber-insurance survey. *Computer Science Review*, 24:35–61.

Marschak, J. (1959). Binary choice constraints and random utility indicators. In *K. Arrow (Ed.), Stanford symposium on mathematical models in the social sciences.* Stanford, CA: Stanford University Press.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In *P. Zarembka (Ed.), Frontiers in econometrics.*, pages 2373–2375. New York: Academic Press.

Ponemon Institute and IBM Security (2020). Cost of a data breach report.

Raiffa, H. (1982). *The art and science of negotiation.* Harvard University Press.

Sanchez, A. (2017). Lloyd's predicts surge in cyber insurance uptake in 2017. `http://www.insurancebusinessmag.com/uk/news/breaking-news/lloyds-predicts-surge-incyber-insurance-uptake-in-2017-42266.aspx`.

Sharma, Y. (2018). Cyber insurance market to reach $14 billion, globally, by 2022.

Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of personality and social psychology*, 39(5):806.

World Economic Forum (2020). The global risks report.

Chapter 1

May cyberinsurance compromise the resilience of digital markets?

Abstract

This study is devoted to understanding and modelling critical behavioural insights in the process of cyberinsurance adoption. Specifically, we start by analysing potential deviations from perfect rationality and the main behavioural features in the purchase of cyberinsurance polices. We also analyse how the adoption of cyberinsurace may affect agents' behaviours in other dimensions of their cybersecurity strategy, such as cyberprotection and safety level when navigating online. To validate our findings, we run an online economic experiment with 4,800 subjects in four EU countries. Our main conclusion is that Rational

Choice Models fail to predict agents' cybersecurity decision. Specifically, we found that individuals show a tendency to opt for an overprotective cybersecurity strategy by ensuring higher protection levels and insurance coverage than those maximising their expected utility. This result motivates the application of a behavioural economics approach to analyse the cyberinsurance, motivating the development of alternative behavioural models not assuming perfect rationality and capable to explain our observational data. Moreover, this result highlights the focus on the human component of cybersecurity and the need to develop behavioural-oriented interventions based in sound behavioural insights and capable to take advantage of the non-rational component of cybersecurity decision-making.

A conceptual map of this chapter is presented in Figure 6.

1.1 Introduction

Cybersecurity is increasingly perceived as a major global problem (World Economic Forum, 2020) becoming even more relevant as companies, administrations and individuals get more interconnected, thereby increasing opportunities for, and facilitating the spread of, cyberthreats. Three widely recognised examples of cyberattacks include: (i) the 2017 Equifax breach, which resulted in stolen data from over 140 million customers - including social security and credit card numbers; (ii) the 2017 WannaCry attack, which took down the UK National Health Service, Telefonica, and FedEx, among others, severely hindering their operations and entailing losses estimated to have reached $4 billion (Berr, 2017); and (iii) the NotPetya malware, which affected thousands of organisations worldwide with an estimated cost of $10 billion (Greenberg, 2018).

Whilst organisations typically employ technical security solutions, including antivirus software, firewalls and intrusion detection systems, it is impossible to achieve perfect security protection (Pal et al., 2017). Cybersecurity failures are often attributed to its human component, which has been previously described as the weakest link in the cybersecurity chain. No matter how well IT systems and company protocols are designed, the company remains vulnerable if employees do not follow the protocols and/or engage in behaviour that weakens the security level. In this context, behavioural change arises as a critical strategy to improve cyber preparedness. Attackers are increasingly aware that employees can provide the most effective entry point into company systems, even if sophisticated security technologies are in place. In addition to malicious, targeted attacks (which included social engineering and phishing), human error accounted for a further 24% of breaches in the latest Ponemon Institute and IBM Security (2020) report. To address the human component of cybersecurity we need to understand the rational and non-rational (although predictable) factors that affect behaviour specifically through the application of theoretically sound and empirically validated models from behavioural economics.

The potential harm caused by security breaches creates the market for cyberinsurance. The global cyberinsurance business was worth around $3.89 billion in 2017 (Androit, 2019). In comparison, the total cost of security breaches worldwide is around $445 billion (Pal et al., 2017), which leaves room for an exponential growth of the cyberinsurance market. In their latest report, Ponemon Institute and IBM Security (2020) found the global average cost of a cyber breach to be $3.86 million, rising to an average of $7.13m for breaches within the healthcare industry. However, the same report found that cyberinsurance alone reduced the total costs of a cyberbreach by an average of almost $200,000. Despite this, studies have shown that only a small percentage of companies are adopting cyberinsurance. Low (2017) found that less than 10% of UK companies report holding cyberinsurance; whilst the Cybersecurity Breaches Survey 2017 found that almost two-fifths (38%) of UK businesses reported having insurance (Klahr et al., 2017). Either way, this number is considerably lower than would be expected. However, Lloyd's of London reported an increase in uptake of 50% in 2016 and they have recently introduced 15 different types of cyberinsurance products for a predicted boom in uptake (Sanchez, 2017) – suggesting that the market is increasing but at a slower rate than predicted.

Current cyberinsurance policies tend to provide three basic types of coverage: Liability as a result of data theft; a means to remedy the breach; and legal and regulatory fines (Bandyopadhyay et al., 2009; Romanosky et al., 2017). The ideal scenario is that organisations will invest in both self-protection (e.g., firewalls and up to date antivirus software) and cyberinsurance (Pal et al., 2017), as well as in educational and behaviour change interventions to promote more secure practices (van Bavel et al., 2019). Cyberattacks can include many different types of risk, e.g., hacking, phishing, DDoS attacks, worms and viruses (Pal et al., 2017). One of the most common sources of security breaches are fraudulent emails sent to staff (Klahr et al., 2017) and other social engineering attacks, further highlighting the need to ensure that staff are aware and capable of detecting and dealing with attacks.

If widely adopted and well-functioning, cyberinsurance has the potential to encourage market-based risk management for information security. It also has the potential to act as an incentive towards organisational investments in information security; which would reduce risk for the investing organisation and for their wider network. Cyberinsurance uptake could also lead to data aggregation on best practices and better tools for assessing security – something that is currently lacking in relation to cyberinsurance. In principle, cyberinsurance could strength IT security for society as a whole (Baer and Parkinson, 2007; Kuru and Bayraktar, 2017). However, massive cyberinsurance adoption is not free of risks and may even compromise the cyber-resilience not only of individual agents but also of the Digital Single Market and digital societies. Among the potential dangers of cyberinsurance adoption, two critical issues should be investigated. Firstly, when adopting cyberinsurance, there are some concerns that decision-makers may behave in an irrational way, resulting in them being unable to identify and purchase their optimal level of cybersecurity coverage. Irrational decisions may be driven by several factors including a lack of insurance and/or cybersecurity literacy required to understand the features of cyberinsurance policies, failing to estimate their actual level of cybervulnerability (i.e., in terms of their risk of receiving an attack and/or their perceptions of an attack's potential impact), and/or the presence of cognitive biases or decision heuristics affecting the decision procedure. Secondly, the cyberinsurance market operates in a state of *information asymmetry*, where the insurer does not have access to all information regarding a company's cyber-preparedness level. Although some elements of the cybersecurity position of the company can be observed by insurers risk audits (such as the existence and features of anti-spam software, firewalls and cyber-recovery plans or intrusion-detection systems), others cannot (such as risky human behaviour for online navigation or offline password management). This information asymmetry could potentially result in *adverse selection* or *moral hazard issues*, which can have dramatic consequences for the resilience of digital markets and societies. For instance, moral hazard could have a hugely detrimental effect if it resulted in those who adopted cyberinsurance showing a significant reduction in cybersecurity protection measures and/or an increase in risky online behaviour.

This paper explores whether concerns around irrational cyberinsurance decisions, information asymmetry and/or moral hazard appear to be justified. Behavioural data were obtained via a large-scale online behavioural economic experiment conducted across four EU countries (Germany, Poland, Spain and UK) with the participation of 4,800 subjects. The conclusions obtained from these data help to give light not only on the potential challenges associated to cyberinsurance adoption, but also on the reasons after them and the policy behavioural interventions required to deal with them.

1.2 Theoretical framework

Cybersecurity is a complex multidimensional issue, involving heterogenous factors such as the diversity of threats, the range of potential impacts, and the many products to choose from for the protection of cyber assets (from a plethora of security controls to relatively new products such as cyberinsurance). In this context, cyberinsurance, cyberprotection and online human behaviour become critical dimensions of the cybersecurity strategy to be set by companies and individuals. Since these dimensions are highly interdependent (Bolot and Lelarge, 2009), the analysis of how they interact is critical to foresee the implications and/or potential risks of the predicted growth of the cyberinsurance market, including its potential impact on the cyber-resilience of digital markets and social systems. As a first step of this analysis, we must first understand each component of cybersecurity. To this end, we should first focus on decisions on cyberprotection and cyberinsurance, whose purchase requires the allocation of a limited budget (which is likely to be split amongst the two). On the other hand, we must consider decisions on online behaviour, both at the organisational level (e.g., cybersecurity policies) and the individual level (e.g., staff behaviour including security compliance). Decisions during online navigation are usually made after cyberprotection and cyberinsurance have been already purchased, and so do not have a direct budgetary implication[1].

Moreover, online behaviour is often non-observable to the insurer, this differs from the physical and technological cyberprotection of the insured, which can often be audited using appropriate monitoring mechanisms. Observability of the cyberprotection level allows the insurer to apply strategies to guarantee or promote reasonable levels of protection, for example by requiring a minimum level of protection to be eligible for insurance or applying price bundling mod-

[1] These decisions may have an economic impact in some cases, such as purchasing a digital product from a secure site instead of downloading it for free from an unknown source. Additionally, other types of cost beyond monetary costs (such as cognitive charge of dealing with complex but secure passwords) may play a relevant role in decision-making (van Bavel et al., 2019).

els. Non-observability of online behaviour generates an information asymmetry between the insurer and the insurance taker, which may translate into moral hazard and adverse selection phenomena. *Adverse selection* comes from the inability of an insurer to distinguish between different client types, i.e., those who have risk-appropriate behaviours and those who do not (Young et al., 2016). As a consequence, the insurer cannot discriminate those agents with a higher risk of suffering an attack, which may be more prone to purchase the insurance. The risk to insurers is also increased by the opportunity for another effect of information asymmetry: *moral hazard*, i.e., the change of behaviour by the insured after purchasing insurance such as reduced incentive to invest in self-protection measures or necessary updates (Eling and Schnell, 2016; Young et al., 2016). This change may be due to dishonesty or alternatively due to behaviour from the client that unintentionally increases the chance and/or severity of loss (Young et al., 2016). As insurers will not run at a loss, this leads to a stalemate situation whereby insurance companies increase their policy prices in an attempt to mitigate risk, however this then deters consumers from purchasing these policies. Moral hazard has been demonstrated in relation to many other types of insurance, e.g., use of health services following health insurance adoption (Sapelli and Vial, 2003) and game play following virtual insurance adoption (Tolvanen, 2015).

Relevant works have disputed these claims, suggesting that moral hazard may not exist in some circumstances (e.g. Chiappori and Salanie, 2000; Zavadil, 2015). Some literature goes even further, and suggests that in contrast to moral hazard and adverse selection, *advantageous* selection can occur. Advantageous selection is possible if individuals who opt to purchase cyberinsurance tend to be more risk averse and seek to reduce risk across all domains of their decision-making and behaviour (e.g. Hudson et al., 2017). These results suggest that the purchase of a cyberinsurance policy may influence other decisions of the insurance-takers, such as the adoption of cyberprotection measures (firewall, antivirus, etc.) or their online behaviour (password management, e-privacy, etc.) in both a positive or negative way.

The theoretical framework and research hypotheses related to decision-making on cyberprotection and cyberinsurance are presented in section 1.2.1. The cybersecurity level of online behaviour is discussed in section 1.2.2 and section 1.2.3 discusses the implication of the decision context in the definition of these three components of cybersecurity strategy. Figure 1.1 presents a map of all the research hypotheses.

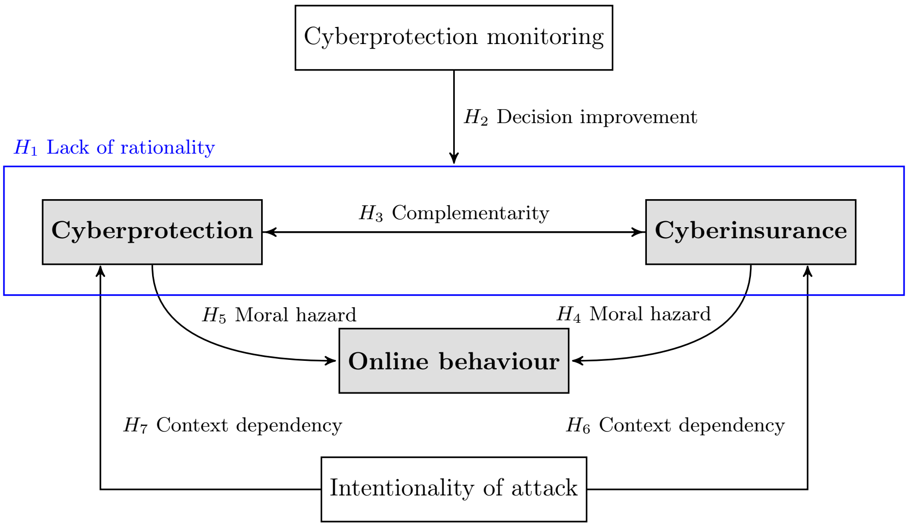


Figure 1.1: Map of the relation of cyberinsurance, cyberprotection and online behaviour.

1.2.1 Acquisition of cyberinsurance and cyberprotection

Companies invest all or part of their cybersecurity budget in the purchase of cyberprotection measures and cyberinsurance products. These purchase decisions can be modelled as standard consumer choice decision-making under uncertainty. From a rational choice perspective, company decision will be determined by the available budget, the prices of the cybersecurity elements, the risk of suffering the attack and the utility function of the company (Wakker, 2010). Specifically, the company will select the combination of cyberprotection and cybersecurity maximising its expected utility. Although many decision models rely upon the assumption that people are rational decision makers, these models are not always effective to predict observed behaviour (Hanoch et al., 2017). Decision-making is often influenced by biases and the use of heuristics (rule of thumb processes) that can lead to less than optimal choices (Gilovich et al., 2002). For example, low probability events are sometimes vastly outweighed or other times just ignored when making a decision whether to purchase insurance (Tversky and Kahneman, 1992). Due to loss aversion, an individual may interpret insurance as a certain expense for a non-certain benefit and then acquire suboptimal coverage (Baicker et al., 2012). Another example is the general lack of knowledge about insurance products in consumers, frequently making poor insurance decisions (Loewenstein et al., 2013). In this context, Behavioural Economics has questioned the capacity of the rational choice approach to explain actual cybersecurity-related decision-making and suggested the need of alternative approaches to model them in what is known as *Behavioural Cybersecurity.* The behavioral approach to cyberinsurance should be applied not only to the demand, but also to the supply side of this market. Results from Farahmand (2019) indicate that in the cyberinsurance industry, corporate managers are also likely rely upon a limited number of simplifying heuristics (i.e., mental shortcuts) rather than extensive algorithmic processing when assessing premiums and making decisions about purchasing cyberinsurance. Additionally, surveys of actuaries and underwriters over decades (Johnson et al., 1993) indicate that insurers price policies for ambiguous events, such as earthquakes and leakage of underground storage tanks, higher than would be suggested by expected utility theory or profit-maximization models. To validate the need of applying such a behavioural insurance approach, we propose our first research hypothesis, H_1.

H_1*: The purchasing decision of cyberprotection and cybersecurity products is not properly explained by rational choice models, i.e., the selection of the cybersecurity strategy is not completely driven by maximisation of the expected utility of the agents.*

As discussed at the beginning of this section, cyberprotection level could be observable by an insurance company auditing the protection elements present in the IT system of a company. Then, the insurer can require a minimum level of protection or apply different premiums in terms of the protection level. Mandatory regulations that stipulate certain self-protection measures (similar to mandatory seat belts in automobiles) is a general requirement in the case of many critical infrastructure operators (Young et al., 2016). The idea of a minimum level of observable protection is also suggested by Cyber Essentials in the UK. In 2014, the UK National Cyber Security Centre (NCSC) introduced

Cyber Essentials: a government backed cybersecurity certification scheme that sets out a good baseline of cybersecurity suitable for all organisations. The scheme addresses five key controls that, correctly implemented, can prevent around 80% of cyberattacks. However, beyond the basic Cyber Essentials or other insurer requirements, the insured can also invest in additional protection measures that are not compulsory.

Beyond the requirement of a minimal protection level, another common practice in the industry is the application of different pricing for the same insurance product depending on the organisation's self-protection level (Young et al., 2016). In this context, a relevant question is whether the application of these pricing strategies helps companies to make closer decisions to those established by rational rational choice theory (Gordon et al., 2003). Hypothesis 2, H_2, states that this is actually the case:

H_2: If cyberprotecion level can be observed by the insurer, appropriate variable pricing policies incentivizing cyberprotection with a cybersinsurance price reduction enhances the rationality level in the purchase of cyberprotection and cybersecurity products.

Businesses, particularly SMEs, can often be heavily restricted by the budget they have available for cybersecurity; because of this they are forced to make trade-offs regarding how they defend their systems (Fielder et al., 2016). When making this trade-off, the organisation has to make a decision based upon the direct cost of implementing a particular safeguard and the impact that the safeguard may have on the business (e.g., indirect costs such as a reduction in productivity speed, system performance speed, morale cost or re-training cost;

Fielder et al., 2016). At a certain level of protection, implementing additional controls/safeguards may only reduce vulnerability by a fraction of its maximum efficiency. Conversely, the cost of implementation remains the same, therefore there becomes a diminishing return for each control that you add to the system (Fielder et al., 2016).

In this context, our third research question is to determine the relationship between cyberinsurance and additional cyberprotection, i.e., if these two dimensions of cybersecurity are perceived as substitutive, independent or complementary. Many researchers believe that cyberinsurance can be an incentive to invest in self-protection, leading to an increase in the level of security and, thus, the level of the security of the Internet in general (Young et al., 2016). Accordingly, our third hypothesis, H_3, claims that protection and insurance are in fact complementary cybersecurity goods.

H_3: Cyberprotection and cybersinsurance products are complementary goods.

In other words, H_3 states that insurance is not considered a substitute for protection and higher levels of insurance are associated with higher levels of protection.

1.2.2 Online behaviour

Cyberthreats to organisations are constantly evolving. For instance, the past several years have also seen the growth of botnets applying social engineering to become more destructive than ever before, as they leverage the computing power of devices that are part of the burgeoning Internet of Things to take advantage of users' unsafe online behaviour. In this context, no matter the latest security products adopted by an organisation, employees' unsafe online navigation or intentional misbehaviour is a critical source of vulnerability (Pal et al., 2017). In an analysis of security breaches reported across different sectors, 64% of incidents were judged to be likely due to improper human behaviour (Evans et al., 2018) and such a 'weak link' in the security chain is increasingly becoming the target of intentional and random cyberattacks (ENISA, 2018).

Since online behaviour is in general unobservable by the insurer, the analysis of the relation between online behaviour and cybersecurity should be discussed in the frame of information asymmetry, focusing on the potential implications of adverse selection and moral hazard. For instance, Gordon et al. (2003) suggest that moral hazard could potentially be addressed by offering premium reductions for increases in security posture, and by imposing deductibles that ensure that the insured suffers some loss in the event of an incident. However, due to the general unobservability of the insured's behaviour, these measures cannot be easily implemented.

The seminal work of Rothschild and Stiglitz (1978) initiated and exemplified the prediction models of insurance markets under asymmetric information. Specifically, they show that those agents with private information and higher risk are more prone to select insurance policies with a higher coverage level than those also with private information but a lower risk. Departing from this work, theoretical research has long emphasized the potential importance of asymmetric information in impairing the efficient operation of insurance markets (Finkelstein and McGarry, 2006). However, there is empirical evidence on which appears to conflict with the major implications in terms of moral hazard and adverse selection of the standard economic model of insurance.For instance, 4.8% of UK credit cards are reported lost or stolen each year, whereas for insured cards the corresponding figure is only 2.7% (De Meza and Webb, 2001) or the mortality rate of U.S. males purchasing life insurance is below that of the uninsured (Cawley and Philipson, 1999). De Meza and Webb (2001) results suggest that individuals who adopt insurance may generally be more risk averse, whereas those who are reluctant to purchase insurance may be less risk adverse and therefore more likely to behave in a risky manner and less inclined to take precautionary security measures. Building from this empirical evidence in other insurance domains, our next research hypotheses, state that those who acquire cyberinsurance (H_4) and/or implement advanced cyberprotection (H_5) will also act more securely online. Specifically:

H_4*: Individuals who have acquired cyberinsurance policies with a higher coverage will behave more securely online.*

H_5*: Individuals who have acquired safer cyberprotection products, will behave more securely online.*

1.2.3 Intentionality of a cyberattack and vulnerability beliefs

The limited research into cyberinsurance has tended to focus upon the supply side of insurability, however the demand side (including formation of risk beliefs) is also vital (Campbell et al., 2011; Weinstein, 1980). Individuals' beliefs about their own susceptibility to an attack directly impact upon their motivation to protect themselves and behave securely (Furnell, 2007), meanwhile inaccurate risk beliefs may contribute to low protection or cyberinsurance uptake (Marotta et al., 2017). To illustrate this, (Davinson and Sillence, 2010) found that training interventions around intentional attacks (phishing) failed to improve secure behaviour unless people changed their views about their own vulnerability.

Aiming at contributing to fill this gap in the literature on cybervulnerability self-perception, our last research question focuses on how the intentionality of a cyberattack (i.e., whether the attack was intentionally targeted on their specific business) affects the beliefs of the agents. Although beliefs are not observable, this impact may be revealed by potential differences in the cyberprotection and cyberinsurance uptake under the presence of intentional (e.g., victims intentionally selected by a cybercriminal) or unintentional (e.g., victims are random, such as a virus spreading randomly through the internet) attacks.

Users' beliefs about their susceptibility to an attack directly impact their motivation to behave securely. In the absence of previous experiences with adverse events, the dramatic communication impact of recent large-scales random cyberattacks may reinforce the perception of the risk of suffering a random attack (availability bias), as happened with the 2017 security breach of Equifax (in which the data of over 140 million customers, including social security and credit card numbers was stolen) or the 2017 massive WannaCry attack. Moreover, Random cyberattacks and the protection against them are not something that can usually be contained within a single organisation and spread fast in the Internet, which may also increase the risk perception for such random attacks coming from multiple connections (Meland et al., 2015). On the other hand, risk beliefs on vulnerability from intentional attacks can be affected by cognitive biases such as optimism bias ("An attacker won't target my business"), which may result in some individuals/businesses assuming that intentional cyberattacks will not happen to them (Eling and Schnell, 2016). Advisen (2015) found that SMEs consider that intentional cyberattacks as less probable to target these companies and are thus less likely to engage with cyberinsurance. However, although contrary to popular this perception, the majority of cyberattacks target small to medium businesses and individuals (Meland et al., 2015; Sarah E., 2012). A critical consequence of the misbeliefs on the risk of intentional cyberattacks is that organisations are not investing time in understanding their vulnerabilities (Marsh, 2016) nor providing adequate funding for cybersecurity (Fielder et al., 2016).

Building from this discussion, our last two research hypotheses focus on existence of observable effects of the intentionality of the attack on the selected protection and insurance coverage levels as a consequence of the above mentioned self-perception of being lees vulnerable to intentional cybermenaces:

H_6*: Individuals will choose higher cyberinsurance coverage when threatened by a random attack than an intentional attack.*

H_7*: Individuals will choose a higher protection level when threatened by a random attack than an intentional attack.*

1.3 Method and experimental design

An online Behavioural Economic Experiment (BEE) was designed and implemented to measure participants' cybersecurity decisions in a controlled situation. The experiment is mainly composed of two tasks:

(i) Purchase decisions about cyberinsurance and security measures products (cybersecurity strategy).

(ii) Online behaviour whilst performing an online task.

Subjects were informed that they would receive payoffs defined in Virtual Currency units (VC) and the conversion rate from VC to cash which would be applied at the end of the experiment. The instructions clearly explained all tasks and decisions to be made during the experiment, as well as, their implications. Figure 1.2 shows the experiment blueprint.

At the beginning of the experiment, subjects were informed that they could suffer a cyberattack with a given initial probability, and the consequences if they did suffer the attack. They were, then, provided with an economic endowment in VC and offered the opportunity to spend part of this to purchase different types of protection security measures (guaranteeing different probability of suffering the attack) and cyberinsurance policies (with different prices and coverages in case of attack). Specifically, subjects could choose between Basic Security Measures (BSMs) and Advance Security Measures (ASMs). BSMs were provided at no cost to the subject, and opting for this option would see the subject retain the initial probability of suffering the attack. Purchasing

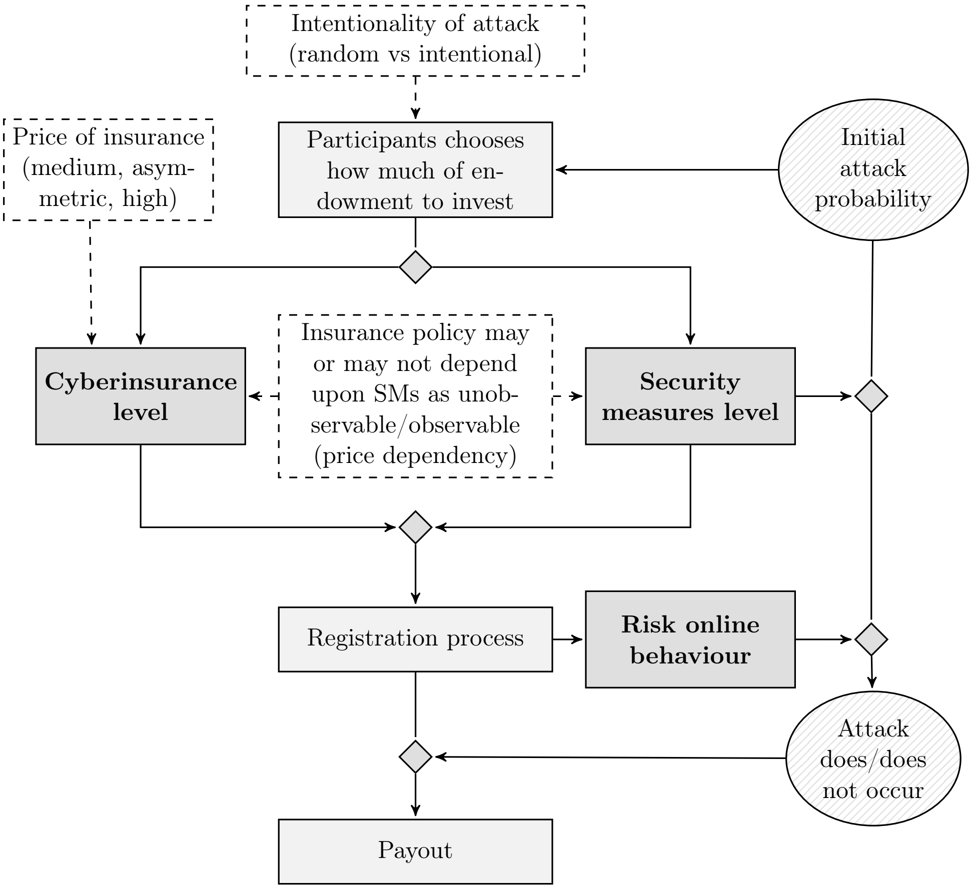
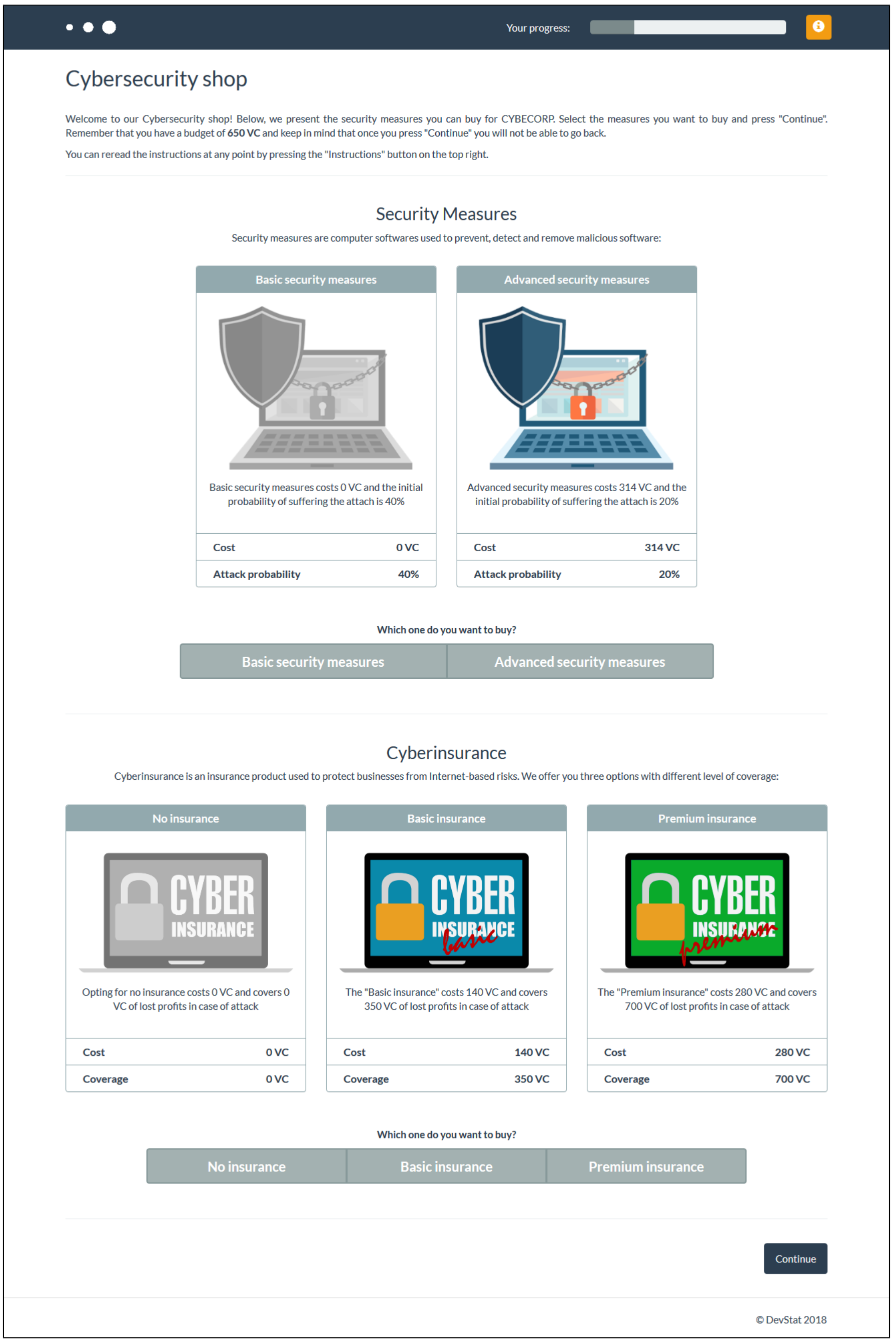

Intentionality of attack (random vs intentional)
Price of insurance (medium, asymmetric, high)
Participants chooses how much of endowment to invest
Initial attack probability
Insurance policy may or may not depend upon SMs as unobservable/observable (price dependency)
Cyberinsurance level
Security measures level
Registration process
Risk online behaviour
Attack does/does not occur
Payout

Figure 1.2: Experiment blueprint.

ASMs required the subjects to invest part of their initial endowment, in return their probability of suffering the cyberattack was reduced by half. In addition to security measures to protect against the attack, subjects also had the opportunity to investing part of their initial endowment in two different levels of cyberinsurance policy (basic or premium) to paying back a part of their losses in case of cyberattack. The basic policy carried a lower price but also offered lower coverage in the event of an attack (i.e., a lower payout). The premium policy had a higher price but providing higher coverage in the event of an attack (i.e., a higher payout). Non-buying any cyberinsurance (none) is a possible option too. Figure 1.3 shows a screenshot of the mock-up website offering the security measures and the cyberinsurance policies.

After purchasing their chosen cyberprotection (security measures) and cyberinsurance options, subjects were asked to complete an online task. The task was to register online for a conference. To complete the registration, each subject was required to perform some security-related decisions, i.e., security level of chosen password, disclosure of non-compulsory private information, viewing the terms and conditions and logging out. Subjects were informed that their probability of suffering a cyberattack would be affected by how securely they behaved whilst completing the online task.

Following completion of the online task, the experiment simulated whether the subject had suffered a cyberattack, or not (based upon the probability calculated from their chosen security measures and their online behaviour). The payoff of each subject was computed as the sum of the remaining endowment (i.e., the initial endowment minus the cost of the security measures and cy-



Cybersecurity shop

Welcome to our Cybersecurity shop! Below, we present the security measures you can buy for CYBECORP. Select the measures you want to buy and press "Continue". Remember that you have a budget of **650 VC** and keep in mind that once you press "Continue" you will not be able to go back.

You can reread the instructions at any point by pressing the "Instructions" button on the top right.

Security Measures

Security measures are computer softwares used to prevent, detect and remove malicious software:

Basic security measures

Basic security measures costs 0 VC and the initial probability of suffering the attach is 40%

Cost	**0 VC**
Attack probability	**40%**

Advanced security measures

Advanced security measures costs 314 VC and the initial probability of suffering the attach is 20%

Cost	**314 VC**
Attack probability	**20%**

Which one do you want to buy?

Cyberinsurance

Cyberinsurance is an insurance product used to protect businesses from Internet-based risks. We offer you three options with different level of coverage:

No insurance

Opting for no insurance costs 0 VC and covers 0 VC of lost profits in case of attack

Cost	**0 VC**
Coverage	**0 VC**

Basic insurance

The "Basic insurance" costs 140 VC and covers 350 VC of lost profits in case of attack

Cost	**140 VC**
Coverage	**350 VC**

Premium insurance

The "Premium insurance" costs 280 VC and covers 700 VC of lost profits in case of attack

Cost	**280 VC**
Coverage	**700 VC**

Which one do you want to buy?

© DevStat 2018

Figure 1.3: Mock-up online shop.

berinsurance products purchased by the participant) and the profit that the company can obtain from their commercial data (if the cyberattack does not occur) or the coverage by the insurance (if the cyberattack occurs). At the end of the experiment, each subject received a variable payoff which depended on her purchase decisions and the fact of suffering or not the cyberattack.

In addition, the experiment included a Holt and Laury test (Holt and Laury, 2002) test to estimate the utility function of each participant. To this end, they were required to make binary decisions between different pairs of random lotteries. At the end of the experiment, participants completed a questionnaire of psychological measures and received their payout.

1.3.1 Treatments

There were two experimental manipulations: the *intentionality of the cyberattack and the pricing strategy* applied to protection and insurance products. The intentionality of the cyberattack (C) has two levels: random attack (participants are informed that there is a virus that may randomly affect any internet user) and intentional attack (participants are informed that an attacker may specifically target their company). Pricing strategy has six levels obtained from a combination of the following two factors: *Ciberinsurance price level* (I) and *price dependency* (P). Insurance price has three levels: medium, asymmetric, and high. Price dependency has two levels: dependent price, where the insurance policy price reflects the chosen security measures; and independent price, where the chosen security measures has no effect on the price of the insurance policy. Therefore, if c^i_{11} is the price of an insurance i, given by its expected value (i.e., the product of the initial probability of a cyberattack and the coverage of the cyberinsurance), the different insurance prices are represented in Table 1.1.

I – Ciberinsurance price level	*P – Price dependency*	
	$P = 1$ – Independent	$P = 2$ – Dependent
$I = 1$ - Medium	c^i_{11}	c^i_{12}=(1-0.5)c^i_{11}
$I = 2$ – Asymmetric	c^1_{21} c^2_{21}=(1+0.2)c^2_1	c^1_{12}=(1-0.5)c^1_{11} c^2_{22}=(1-0.7)c^2_{11}
$I = 3$ – High	c^i_{31}=(1+0.2)c^i_1	c^i_{32}=(1-0.3)c^i_{11}

Note: Basic cyberinsurance: $i = 1$; Premium cyberinsurance $i = 2$.

Table 1.1: Cyberinsurance prices c^i_{IP}

The experiment implements a full-factorial design with the following three factors and 2 x 2 x 3 levels (Table 1.2). Participants were randomly allocated to each of the 12 combinations of the three factors.

Factor	**Levels**
C: Intentionality of the cyberattack	C1: The attack is random (there is a virus in the Internet that may affect randomly to any user). C2: The attack is intentional (the attack is intentionally launch by a cyber-criminal).
P: Price dependency	P1: The price of the insurance does not depend on the protection level. P2: The price of the insurance does depend on the protection level.
I: Cyberinsurance price level	I1: Medium price. I2: Asymmetric price. I3: High price.

Table 1.2: Experimental conditions

1.3.2 Behavioural measures

Three behavioural measures were obtained during the experiment: two purchase-based measures based upon *security measures adoption* (with two possible values, Basic or Advanced) and *insurance adoption* (with three possible values none, *basic or premium*). The third measure is based upon the individual's online behaviour (Box 1.1) during the conference registration task. Online behaviour is calculated as a continuous variable between 0 (safest behaviour) and 1 (riskiest behaviour) as a linear combination of the proxy security variables included in the experiment: *security level of chosen password*, *disclosure of non-compulsory private information*, *viewing the terms and conditions* and *logging out* after completing the registration.

1.3.3 Sample

Participants ($N = 4,800$) were recruited across four EU countries (Germany, Poland, Spain and UK) in June 2018. The distribution by age and gender reflects Eurostat's data from the 2017 survey on ICT[2] that was used to create the quota, Table 1.3. The experiment software was a web application developed using the Yii PHP framework. Following a between-participants design, 100 participants of each country were randomly assigned to each treatment. The actual payoff for each participant was the sum of a constant show-up fee and the sum of payoff obtained by the participant in each phase. Regarding the

[2] Data given in this domain are collected annually by the National Statistical Institutes and are based on Eurostat's annual model questionnaires on ICT (Information and Communication Technologies) usage in households and by individuals. https://ec.europa.eu/eurostat/cache/metadata/en/isoc_i_esms.htm

The risk level is computed from the following binary variables, which are equal to 1 if they verify the following statements or 0 otherwise:

- Password, x_i^{pass}: Password does not contain capital letters; Password does not contain lowercase letters; Password does not contain numbers; Password does not contain special characters (|['^£$%&*)(}{ @#~?,|><>=_+¬-); Password is short (less than 8 characters); Password includes the username (case-insensitive).
- Registration, x_i^{reg}: The participant has filled the "First name" field; The participant has filled the "Last name" field; The participant has filled the "Occupation" field; The participant has filled the "Phone Number" field; The participant has filled the "Address" field; The participant has filled the "City" field; The participant has filled the "Zip" field.
- Privacy policy, x_i^{pp}: The participant has not opened the "Privacy Policy" wind.
- Log out, x_i^{log}:The participant has not logged out of the website after the registration.

The security level of the online behaviour, OB, is obtained as a weighted average of the above variables:

$$OB = w_{pass} \sum_{i=1}^{6} x_i^{pass} + w_{reg} \sum_{i=1}^{7} x_i^{reg} + w_{pp} x_i^{pp} + w_{log} x_i^{log}$$

where w represents the weight of each binary variable, given by $w_{pass} = \frac{0.4}{6}$, $w_{reg} = \frac{0.3}{7}$, $w_{pp} = 0.15$ and $w_{log} = 0.15$.

Box 1.1: Online behaviour computation.

education of participants, most of them had either finished high school or had a university degree, as shown in Table 1.4.

	Germany		***Spain***		***Poland***		***UK***	
	n	%	n	%	n	%	n	%
Male	617	51.4	600	50.0	552	46.0	595	49.6
Female	583	48.6	600	50.0	648	54.0	605	50.4
16 – 34 years	932	77.7	842	70.2	713	59.4	844	70.3
35 – 74 years	268	22.3	358	29.8	487	40.6	356	29.7
Total	*1,200*	*100.0*	*1,200*	*100.0*	*1,200*	*100.00*	*1,200*	*100.0*

Table 1.3: Distribution of the participants by gender, age and country.

Education level	***n***	***%***
Compulsory	403	8,4
Further	1,446	30.1
Higher	2,951	61.5
Total	*4,800*	*100.0*

Work Status	***n***	***%***
Worker	2,808	58.5
Self-employed	452	9.4
Other	1,540	32.1
Total	*4,800*	*100.0*

Table 1.4: Distribution of the participants by level of education and work status.

1.4 Results

1.4.1 Cybersecurity strategy

The cybersecurity strategy (combination of cyberinsurance and cyberprotection/security measures) chosen by participants is represented in Figure 1.4. The results are shown for both price dependency groups (dependent and independent). In both groups, the most frequently selected strategy is the safest option: adoption of both advanced security measures (ASMs) and premium insurance.

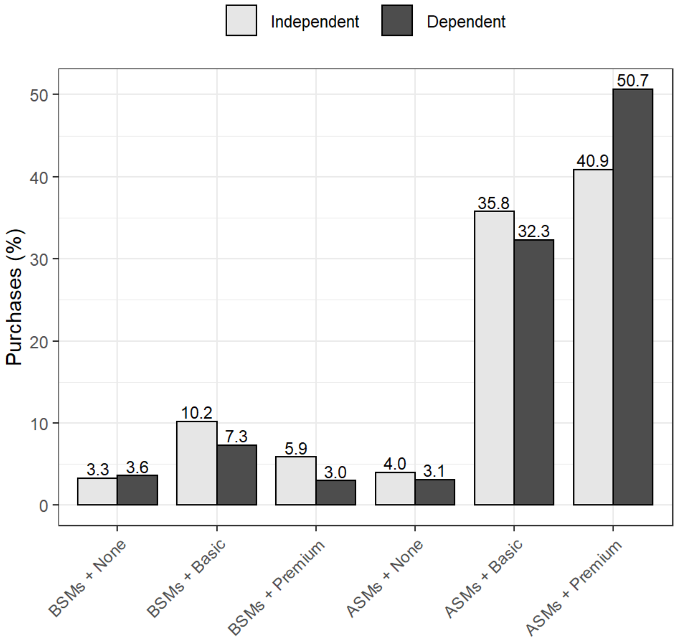


Figure 1.4: Cybersecurity strategy.

To test our first research hypothesis, we need to compare the cybersecurity strategy selected by an agent with that maximimizing her expected utility. To this end, we follow Holt and Laury (2002) methodology to determine which is the best of protection and insurance from a rational choice perspective for each subject. Specifically, we assume that the utility function of each subject follows a constant relative risk aversion specification, with relative risk aversion r, given by the expression $U(x) = x^{1-r}/(1-r)$. The relative risk aversion of each participant has been estimated from their selections of ten options from ten ordered pairs of incentivised lotteries including a safer (A) and riskier (B) option, as in Holt and Laury (2002) experiment. Table 1.5 presents the optimal cybersecurity strategy from a rational choice perspective for a subject moving from the safer to the riskier lottery B from the first to the tenth pair of lotteries in Holt and Laury (2002). The distribution of subjects moving from the safer (A) to riskier (B) options at each ordered pair of lotteries is presented in Figure 1.5.

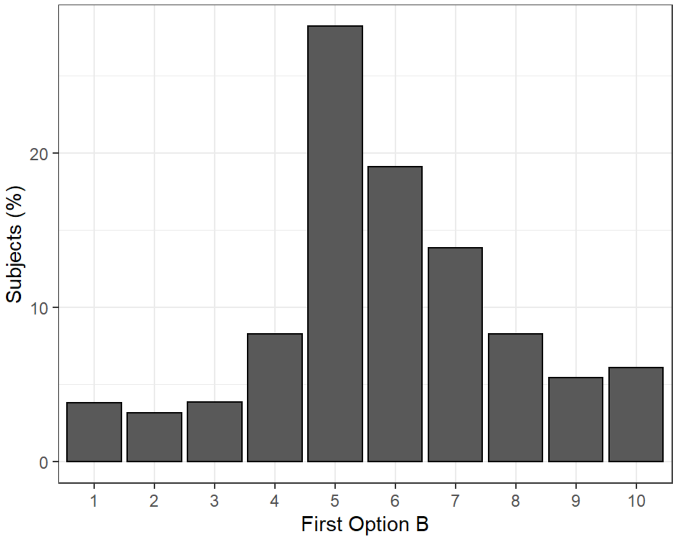


Figure 1.5: Distribution of participants moving to the riskier option B at each pair of ordered lotteries.

1st B	**Best cybersecurity strategy from a rational choice perspective**					
	Independent price			**Dependent price**		
	Rational	**Asymmetric**	**High**	**Rational**	**Asymmetric**	**High**
1	BSMs + None	BSMs + None	BSMs + None	BSMs + None	BSMs + None	BSMs + None
2						
3						
4						
5	BSMs + Premium	BSMs + Basic		BSMs + Premium	BSMs + Basic	
6						
7			BSMs + Premium			BSMs + Premium
8						ASMs + Premium
9		BSMs + Premium			ASMs + Premium	
10						

Table 1.5: Optimal cybersecurity strategy from a rational choice perspective by relative risk aversion.

As shown in Table 1.5, when the price is independent, the purchase of ASMs is never the best choice from a rational choice perspective. Taking into account these values, the best combination was purchased by only 5.3% of participants in the independent price group. This percentage increased to 8.0% in the dependent price group (Table 1.6). A chi square test shows that this increase is significant (*p-value* = 0.015).

	Cybersecurity strategy	**Purchases**	**p-value (χ^2)**
Independent	Best from a rational choice perspective	5.30%	0.015
	Other	94.70%	
Dependent	Best from a rational choice perspective	8.02%	
	Other	91.98%	

Table 1.6: Purchases of the best cybersecurity strategy from a rational choice perspective by price dependency

These results support our first and third hypothesis: Subjects do not make optimal purchases of cyberinsurance and cyberprotection from a rational choice perspective (H_1), however the application of appropriate pricing startegies depending on protection level helps subjects to make better decisions from a rational choice perspective (H_3).

The crossed distribution of the purchases of security measures and cyberinsurance products is presented in Figure 1.6. If we focus upon participants who purchased premium insurance, we notice that the majority (91.2%) of these participants also purchase the ASMs. In contrast, this share drops to 79.5% for those subjects who purchased only basic insurance, and 50.9% for those who did not purchase any insurance at all. A chi square test shows that this difference is significant (*p-value* < 0.001). The combination of the products therefore appears to be complementary, supporting our second hypothesis that insurance does not substitute protection, but individuals show a tendency to purchase both types of products (complementary goods).

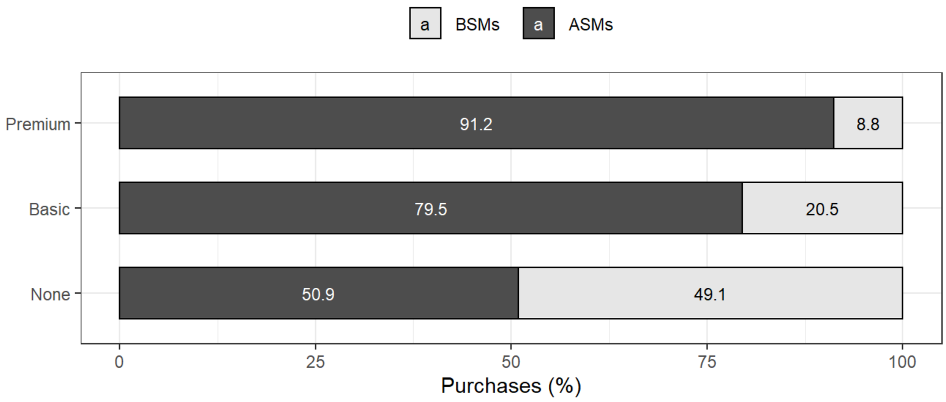


Figure 1.6: Security measures purchased by each cyberinsurance group.

1.4.2 Impact of cyberprotection in online behaviour

Figure 1.7 shows the level of risk assumed by the subjects in their online behaviour, with a breakdown in terms of the adopted security measures and cyberinsurance. Note that the value 0 depicts online behaviour that is completely safe, this number increases (up to a maximum of 1) depending upon the degree of risky behaviour by the subject during the online navigation task.

Since online behaviour is a continuous variable, two ANOVA models was estimated to study the effects of (i) the acquisition of SM and (ii) the purchase of cyberinsurance with online behaviour as the dependent variable, Table 1.7. As shown by the corresponding statistical test ($p\text{-}value < 0.001$), the acquisition of

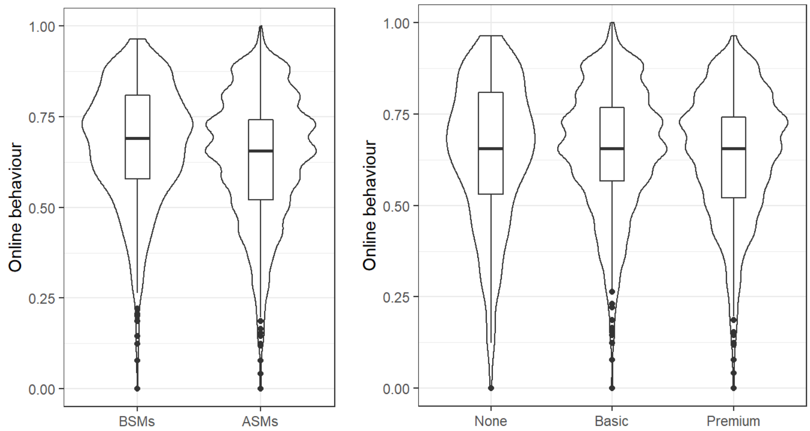


Figure 1.7: Online behaviour by SMs and cyberinsurance acquisitions.

ASMs has a significant effect on online behaviour; subjects who adopted ASMs behaved more securely during the online task. No significant effect was found for the purchase of cyberinsurance on online behaviour ($p\text{-}value = 0.200$).

	Product	**n**	**Mean**	**Std. Error**	**p-value**	
SMs	BSMs	799	0.674	0.162	0.000	***
	ASMs	1935	0.638	0.174		
Cyberinsurance	None	175	0.645	0.195	0.200	
	Basic	1103	0.649	0.172		
	Premium	1122	0.640	0.171		

. $p\text{-}value < 0.1$; * $p\text{-}value < 0.05$; ** $p\text{-}value < 0.01$; *** $p\text{-}value < 0.001$

Table 1.7: Online behaviour by products when SMs is no observable

Therefore, H_4 is supported – individuals who adoption advanced security measures behave more securely online. However, this behaviour is not affected by the acquisition of cyberinsurance (rejecting H_5).

1.4.3 Effects of the intentionality of the attack

This section is devoted to test our last two research hypotheses, which focus on the impact of the intentionality of the attacks on agents' cybersecurity behaviour. To isolate the impact of this variable from those of the other profile variable and treatments, the next two subsections present and interpret the estimation of two logistic models variable cyberinsurance and cyberprotection adoption as dependent variables. Beyond allowing for testing our research hypothesis H_6 and H_7, these estimations provide additional results on the effect of the price level (I) and price dependency (P).

Effects of intentionality on the acquisition of cyberinsurance

To evaluate the impact of the intentionality and avoid confounding effects with other profile variables and experimental factors, we estimate a logistic model of cybersinurance adoption, including all these variables as independent. Note that, since the agent can choose among three different levels of insurance (None, Basic and Premium), we apply a multinomial logistic model, whose estimations are presented in Table 1.8. Since all the profile variables have a significant effect of the selection of the insurance coverage, none of them has been removed from the model. The main conclusion from this model is that the intentionality of the attack has no significant impact on the coverage level selected by the agent, therefore rejecting our research hypothesis H_6.

	None cyberinsurance				***Premium cyberinsurance***			
	Estimate	***Std. Error***	***Odds***	***Pr(>|z|)***	***Estimate***	***Std. Error***	***Odds***	***Pr(>|z|)***
(Intercept)	-1.503	0.337	2.258	0.000 ***	-0.216	0.177	0.047	0.223
Age	-0.001	0.005	0.000	0.904	0.011	0.003	0.000	0.000 ***
Male	0.308	0.146	0.095	0.035 *	-0.166	0.072	0.028	0.021 *
Education: Further	0.075	0.272	0.006	0.783	0.142	0.148	0.020	0.337
Education: Higher	-0.006	0.167	0.000	0.971	-0.207	0.083	0.043	0.012 *
Employment: Self-employed	0.633	0.223	0.401	0.004 **	-0.030	0.128	0.001	0.815
Employment: Other	0.382	0.160	0.146	0.017 *	0.059	0.080	0.003	0.465
Country: Spain	-0.749	0.210	0.561	0.000 ***	0.445	0.109	0.198	0.000 ***
Country: Poland	-0.851	0.209	0.724	0.000 ***	0.276	0.108	0.076	0.010 *
Country: UK	-0.732	0.195	0.535	0.000 ***	0.207	0.105	0.043	0.049
Risk Aversion	0.202	0.077	0.041	0.009 **	0.123	0.038	0.015	0.001 **
C: Intentional	0.047	0.143	0.002	0.742	-0.093	0.071	0.009	0.189
I: Asymmetric	-0.177	0.175	0.031	0.313	-0.254	0.087	0.065	0.003 **
I: High	-0.097	0.173	0.009	0.576	-0.196	0.087	0.039	0.023 *
P: Dependent	-0.078	0.143	0.006	0.588	0.253	0.071	0.064	0.000 ***

. *p-value* < 0.1; * *p-value* < 0.05; ** *p-value* < 0.01; *** *p-value* < 0.001

Note: The reference levels of the multinomial model are basic cyberinsurance, woman, compulsory education, worker, German, C: Random and I: Medium.

Table 1.8: Estimation of cyberinsurance decision model.

The estimated model provides additional results on how other covariables do affect cyberinsurance adoption. For instance, cyberinsurance price level (I) and price dependency (P) have a significant effect on the decision of purchase basic or premium cyberinsurance, although not in the decision of whether to buy any cyberinsurance. To illustrate this, Table 1.9 and Table 1.10 present the marginal effects of factors I and P respectively. In average, and assuming that all other variables are constant, a change from Rational to Asymmetric prices reduces the probability of purchase the Premium insurance decrease in almost 6 points. On other hand, price dependence increases the probability of acquire Premium insurance more than 6 points.

	Predict probabilities (%)			*Differences with Medium*	
Insurance	*Medium*	*Asymmetric*	*High*	*Asymmetric*	*High*
None	6.64	6.40	6.70	-0.24	0.06
Basic	38.71	44.58	43.11	5.87	4.40
Premium	54.65	49.02	50.19	-5.63	-4.46

Table 1.9: Marginals effects of Factor I in cyberinsurance decision.

	Predict probabilities (%)		*Differences*
Insurance	*Independent*	*Dependent*	*Independent vs Dependent*
None	7.23	5.94	-1.29
Basic	44.71	39.55	-5.16
Premium	48.06	54.50	6.44

Table 1.10: Marginals effects of Factor P in cyberinsurance decision.

Effects of intentionality on the acquisition of cyberprotection

The test of hypothesis H_7 is also carried out by the estimation of a logistic regression (logit) model. To this end, we have estimated a logit model considering the purchase or not of advanced security measures (ASMs) as the dependent variable and the intentionality of the attack, pricing strategy and subject's profile as independent variables. Since the estimation of this first model shows that sex and education level have no significant predictive power for the purchase of protection measures, these covariables have been removed. The estimation of the model without these two variables is presented in Table 1.11.

	Estimate	***Std. Error***	***Odds***	***Pr(>|z|)***	
(Intercept)	0.914	0.212	0.836	0.000	***
Age	0.009	0.004	0.000	0.009	**
Employment: Self-employed	-0.488	0.152	0.238	0.001	**
Employment: Other	0.061	0.108	0.004	0.574	
Country: Spain	0.350	0.132	0.122	0.008	**
Country: Poland	0.413	0.135	0.171	0.002	**
Country: UK	0.480	0.136	0.230	0.000	***
Risk Aversion	0.136	0.051	0.018	0.007	**
C: Intentional	-0.189	0.095	0.036	0.046	*
I: Asymmetric	0.031	0.115	0.001	0.791	
I: High	0.081	0.116	0.007	0.483	
P: Dependent	0.514	0.096	0.264	0.000	***

. *p-value* < 0.1; * *p-value* < 0.05; ** *p-value* < 0.01; *** *p-value* < 0.001

Table 1.11: Estimation of ASMs purchases model.

According to the information in Table 1.11, the intentionality of the attack has a significant effect on the decision of purchase ASMs, supporting our hypotheses that individuals will choose a higher protection level when threatened by a random attack (H_7). The marginal effects of a cyberattack being intentional is a reduction in the probability of adopting advanced security measures of 2.4 points (Table 1.12).

Additionally, we can conclude that the cyberinsurance price level (I) does not have a significant effect on the decision to purchase ASMs ($p\text{-}value > 0.05$), whereas price dependency (P) has significant predictive power.

SMs	**Factor**	**Predict probabilities (%)**	**Differences**
ASMs	*Random*	85.98	-2.40
	Intentional	83.58	
	Independent	81.47	6.49
	Dependent	87.96	

Table 1.12: Marginals effects of Factors C and P in SMs decision model when SMs is no observable.

1.5 Discussion

Using a behavioural-experimental economics approach, this paper shows that Rational Choice Model fails to predict agents' cybersecurity decision, specifically their selection of protection level and cyberinsurance coverage. We found that individuals show a tendency to opt for an overprotective cybersecurity strategy by ensuring higher protection levels and insurance coverage than those maximising their expected utility. This result motivates the application of a behavioural economics approach to analyse the cyberinsurance sector, motivating the development of alternative behavioural models not assuming perfect rationality and capable to explain our observational data. The development, validation and interpretation of such models is presented in Chapter 2. Moreover, this result highlights the relevance of focusing on the human component of cybersecurity and the need to develop behavioural-oriented interventions (from public policies to improve the general level of cybersecurity of digital markets to the design of cybersecurity strategy in individual companies) based in sound behavioural insights and capable to take advantage of the non-rational component of cybersecurity decision-making. Our paper also shows that cybersecurity audits making the protection levels of the agent observable to the insurer and conditioning the policy primes to it, arises as an effective strategy to help agents to behave as expected utility maximisers. As an application, this finding supports public regulations and insurers pricing policies requiring a minimum level of protection as a condition to opt for a cyberinsurance policy.

Our paper also tests the existence of behavioural insights of cyberinsurance adoption previously proposed in the insurance and cyberinsurance literature. Specifically, our research allows for disregarding two critical features that, if existing, may translate the development of the cyberinsurance sector into a weakening of the resilience of global cyber systems. To this end, we show firstly that protection and insurance coverage are not substitutive but complementary goods. In other words, the adoption of a higher level of coverage is not associated to the purchase of a lower level of protection. In fact, the situation is the opposite, with strongly insured agents adopting also more advanced security measures. This finding suggests the existence of some kind of advantageous selection effect in the cyberinsurance market (in these cases where protection cannot be observed by the insurer) and the presence of an underlaying segmentation of the agents in terms of their awareness and aversion towards cyber-risks. The second critical feature is related to the impact of cyberinsurance adoption on the safety level of online behaviour. Our research provides empirical evidence on the fact that such an impact does not exist: even when online behaviour is not observable by the insurer, the coverage of the losses associated to a cyber attack is not associated with less safe behaviours increasing the chances to receive such an attack. These results suggest that, despite its information asymmetry, moral hazard does not appear in the cyberinsurance market. Beyond their scientific interest, these two findings are relevant from a policy and industrial perspective. Since they show that information asymmetry issues (adverse selection and moral hazard) seem not to be a problem for the cyberinsurance sector, the adoption of cyberinsurance in a company or policy-making to promote the adoption of this type of policies would non reduce the safety in cybersecurity related human behaviour.

Finally, we have analysed how agents react under the menace of intentional attacks. In our controlled experimental environment, the chances and the impact of both unintentional and intentional attacks are just the same under. Therefore, a rational agent is expected to purchase the same security measure and insurance policies in these two situations. However, our results show that this is not the case: Agents adopt a significantly higher level of cyberprotection in front of a random cyberattack. This finding suggests that the intentionality of an attack may influence agents' beliefs and vulnerability self-perception by potentiating cognitive biases such as optimism bias ("my data is not interesting enough for a cybercrimial") and resulting in agents assuming that intentional cyber attacks will not happen to them. This irrational behaviour motivates a deeper analysis of risk belief formation in adversarial situations, as presented in Chapter 3.

References

Advisen (2015). 2015 network security & cyber risk management: The fourth annual survey of enterprise-wide cyber risk management practices in europe.

Androit (2019). Global cyber security insurance market 2025.

Baer, W. S. and Parkinson, A. (2007). Cyberinsurance in it security management. *IEEE Security & Privacy*, 5(3):50–56.

Baicker, K., Congdon, W. J., and Mullainathan, S. (2012). Health insurance coverage and take-up: Lessons from behavioral economics. *The Milbank Quarterly*, 90(1):107–134.

Bandyopadhyay, T., Mookerjee, V. S., and Rao, R. C. (2009). Why it managers don't go for cyber-insurance products. *Communications of the ACM*, 52(11):68–73.

Berr, J. (2017). Wannacry ransomware attack losses could reach $4 billion. `https://www.cbsnews.com/news/wannacry-ransomware-attacks-wannacry-virus-losses/`.

Bolot, J. and Lelarge, M. (2009). Cyber insurance as an incentivefor internet security. In *Managing information risk and the economics of security*, pages 269–290. Springer.

Campbell, J., Ma, W., and Kleeman, D. (2011). Impact of restrictive composition policy on user password choices. *Behaviour & Information Technology*, 30(3):379–388.

Cawley, J. and Philipson, T. (1999). An empirical examination of information barriers to trade in insurance. *American Economic Review*, 89(4):827–846.

Chiappori, P.-A. and Salanie, B. (2000). Testing for asymmetric information in insurance markets. *Journal of political Economy*, 108(1):56–78.

Davinson, N. and Sillence, E. (2010). It won't happen to me: Promoting secure behaviour among internet users. *Computers in Human Behavior*, 26(6):1739–1747.

De Meza, D. and Webb, D. C. (2001). Advantageous selection in insurance markets. *RAND Journal of Economics*, pages 249–262.

Eling, M. and Schnell, W. (2016). What do we know about cyber risk and cyber risk insurance? *The Journal of Risk Finance*.

ENISA (2018). Cybersecurity culture guidelines: Behavioural aspects of cybersecurity.

Evans, M. G., He, Y., Yevseyeva, I., and Janicke, H. (2018). Analysis of published public sector information security incidents and breaches to establish the proportions of human error. In *HAISA*, pages 191–202.

Farahmand, F. (2019). Quantitative issues in cyberinsurance: Lessons from behavioral economics, counterfactuals, and causal inference. *IEEE Security & Privacy*, 18(2):8–15.

Fielder, A., Panaousis, E., Malacaria, P., Hankin, C., and Smeraldi, F. (2016). Decision support approaches for cyber security investment. *Decision support systems*, 86:13–23.

Finkelstein, A. and McGarry, K. (2006). Multiple dimensions of private information: evidence from the long-term care insurance market. *American Economic Review*, 96(4):938–958.

Furnell, S. (2007). An assessment of website password practices. *Computers & Security*, 26(7-8):445–451.

Gilovich, T., Griffin, D., and Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge university press.

Gordon, L. A., Loeb, M. P., and Sohail, T. (2003). A framework for using insurance for cyber-risk management. *Communications of the ACM*, 46(3):81–85.

Greenberg, A. (2018). The untold story of notpetya, the most devastating cyberattack in history.

Hanoch, Y., Barnes, A., and Rice, T. (2017). *Behavioral economics and healthy behaviors: Key concepts and current research*. Taylor & Francis.

Holt, C. A. and Laury, S. K. (2002). Risk aversion and incentive effects. *American economic review*, 92(5):1644–1655.

Hudson, P., Botzen, W. W., Czajkowski, J., and Kreibich, H. (2017). Moral hazard in natural disaster insurance markets: empirical evidence from germany and the united states. *Land Economics*, 93(2):179–208.

Johnson, E. J., Hershey, J., Meszaros, J., and Kunreuther, H. (1993). Framing, probability distortions, and insurance decisions. *Journal of risk and uncertainty*, 7(1):35–51.

Klahr, R., Shah, J., Sheriffs, P., Rossington, T., Pestell, G., Button, M., and Wang, V. (2017). Cyber security breaches survey 2017.

Kuru, D. and Bayraktar, S. (2017). The effect of cyber-risk insurance to social welfare. *Journal of Financial Crime.*

Loewenstein, G., Friedman, J. Y., McGill, B., Ahmad, S., Linck, S., Sinkula, S., Beshears, J., Choi, J. J., Kolstad, J., Laibson, D., et al. (2013). Consumers' misunderstanding of health insurance. *Journal of Health Economics*, 32(5):850–862.

Low, P. (2017). Insuring against cyber-attacks. *Computer Fraud & Security*, 2017(4):18–20.

Marotta, A., Martinelli, F., Nanni, S., Orlando, A., and Yautsiukhin, A. (2017). Cyber-insurance survey. *Computer Science Review*, 24:35–61.

Marsh (2016). Uk cyber risk survey report.

Meland, P. H., Tondel, I. A., and Solhaug, B. (2015). Mitigating risk with cyberinsurance. *IEEE Security & Privacy*, 13(6):38–43.

Pal, R., Golubchik, L., Psounis, K., and Hui, P. (2017). Security pricing as enabler of cyber-insurance a first look at differentiated pricing markets. *IEEE Transactions on Dependable and Secure Computing*, 16(2):358–372.

Ponemon Institute and IBM Security (2020). Cost of a data breach report.

Romanosky, S., Ablon, L., Kuehn, A., and Jones, T. (2017). Content analysis of cyber insurance policies: How do carriers write policies and price cyber risk? *Available at SSRN 2929137.*

Rothschild, M. and Stiglitz, J. (1978). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. In *Uncertainty in economics*, pages 257–280. Elsevier.

Sanchez, A. (2017). Lloyd's predicts surge in cyber insurance uptake in 2017. `http://www.insurancebusinessmag.com/uk/news/breaking-news/lloyds-predicts-surge-incyber-insurance-uptake-in-2017-42266.aspx`.

Sapelli, C. and Vial, B. (2003). Self-selection and moral hazard in chilean health insurance. *Journal of health economics*, 22(3):459–476.

Sarah E., N. (2012). Cybercriminals sniff out vulnerable firms. *Wall Street Journal*.

Tolvanen, J. (2015). Measuring moral hazard using insurance panel data.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and uncertainty*, 5(4):297–323.

van Bavel, R., Rodríguez-Priego, N., Vila, J., and Briggs, P. (2019). Using protection motivation theory in the design of nudges to improve online security behavior. *International Journal of Human-Computer Studies*, 123:29–39.

Wakker, P. P. (2010). *Prospect theory: For risk and ambiguity.* Cambridge university press.

Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of personality and social psychology*, 39(5):806.

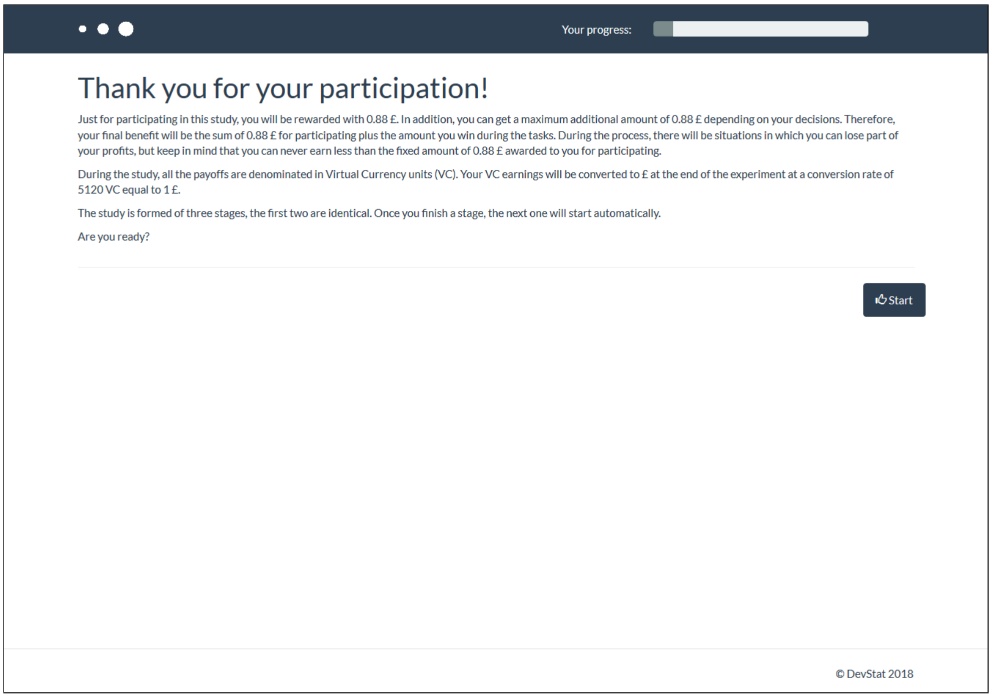
World Economic Forum (2020). The global risks report.

Young, D., Lopez Jr, J., Rice, M., Ramsey, B., and McTasney, R. (2016). A framework for incorporating insurance in critical infrastructure cyber risk strategies. *International Journal of Critical Infrastructure Protection*, 14:43–57.

Zavadil, T. (2015). Do the better insured cause more damage? testing for asymmetric information in car insurance. *Journal of Risk and Insurance*, 82(4):865–889.

Appendix

1.A Appendix: Screenshots from experiment



Thank you for your participation!

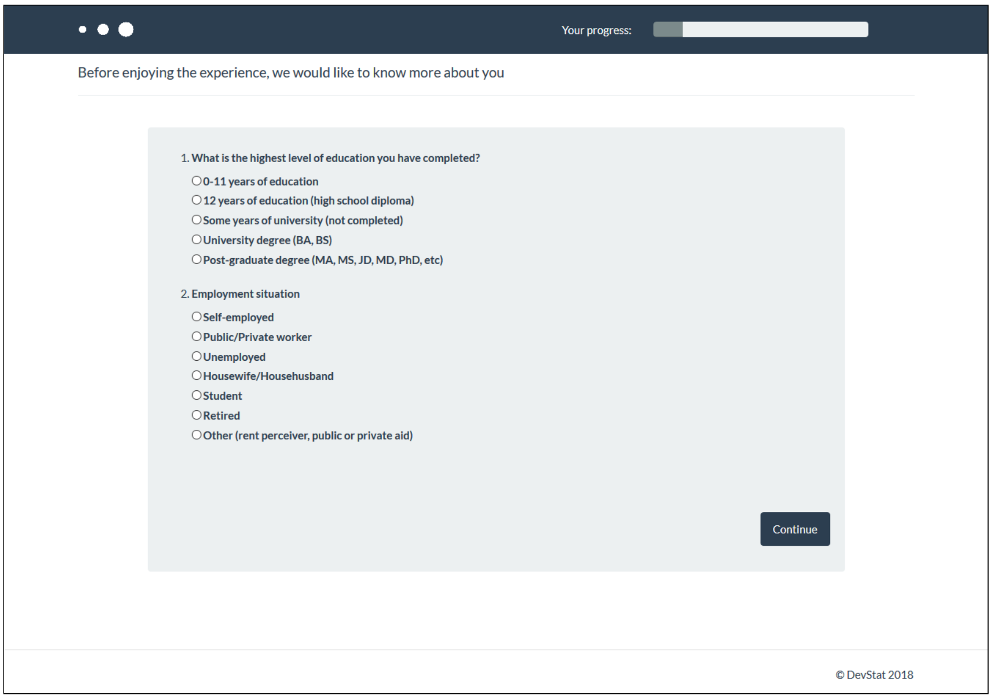
Just for participating in this study, you will be rewarded with 0.88 £. In addition, you can get a maximum additional amount of 0.88 £ depending on your decisions. Therefore, your final benefit will be the sum of 0.88 £ for participating plus the amount you win during the tasks. During the process, there will be situations in which you can lose part of your profits, but keep in mind that you can never earn less than the fixed amount of 0.88 £ awarded to you for participating.

During the study, all the payoffs are denominated in Virtual Currency units (VC). Your VC earnings will be converted to £ at the end of the experiment at a conversion rate of 5120 VC equal to 1 £.

The study is formed of three stages, the first two are identical. Once you finish a stage, the next one will start automatically.

Are you ready?

Figure 1.A.1: Welcome page.



Before enjoying the experience, we would like to know more about you

1. What is the highest level of education you have completed?

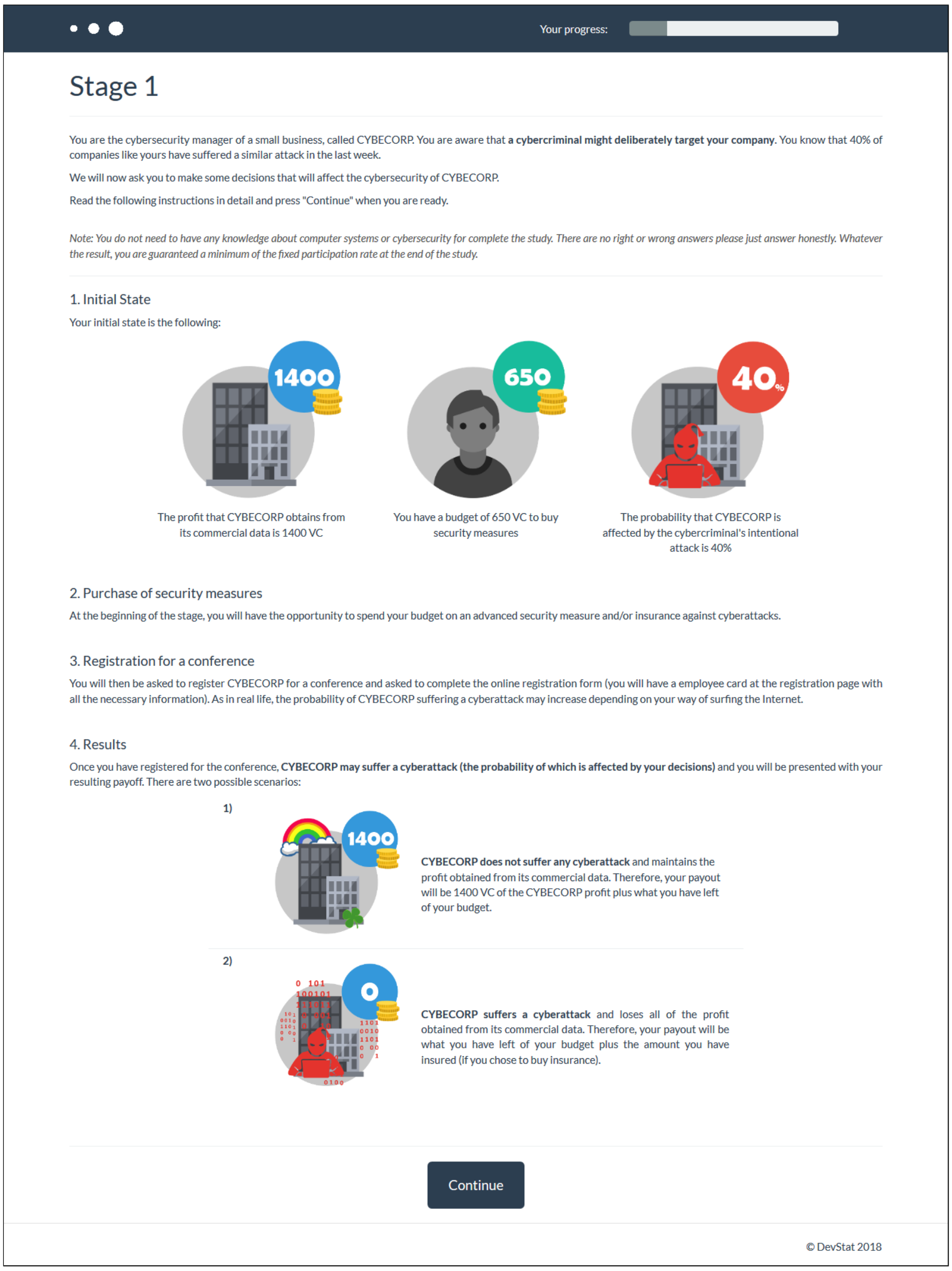
○ 0-11 years of education
○ 12 years of education (high school diploma)
○ Some years of university (not completed)
○ University degree (BA, BS)
○ Post-graduate degree (MA, MS, JD, MD, PhD, etc)

2. Employment situation

○ Self-employed
○ Public/Private worker
○ Unemployed
○ Housewife/Househusband
○ Student
○ Retired
○ Other (rent perceiver, public or private aid)

© DevStat 2018

Figure 1.A.2: Socio-demographic questionnaire.



Stage 1

You are the cybersecurity manager of a small business, called CYBECORP. You are aware that **a cybercriminal might deliberately target your company**. You know that 40% of companies like yours have suffered a similar attack in the last week.

We will now ask you to make some decisions that will affect the cybersecurity of CYBECORP.

Read the following instructions in detail and press "Continue" when you are ready.

Note: You do not need to have any knowledge about computer systems or cybersecurity for complete the study. There are no right or wrong answers please just answer honestly. Whatever the result, you are guaranteed a minimum of the fixed participation rate at the end of the study.

1. Initial State

Your initial state is the following:

The profit that CYBECORP obtains from its commercial data is 1400 VC

You have a budget of 650 VC to buy security measures

The probability that CYBECORP is affected by the cybercriminal's intentional attack is 40%

2. Purchase of security measures

At the beginning of the stage, you will have the opportunity to spend your budget on an advanced security measure and/or insurance against cyberattacks.

3. Registration for a conference

You will then be asked to register CYBECORP for a conference and asked to complete the online registration form (you will have a employee card at the registration page with all the necessary information). As in real life, the probability of CYBECORP suffering a cyberattack may increase depending on your way of surfing the Internet.

4. Results

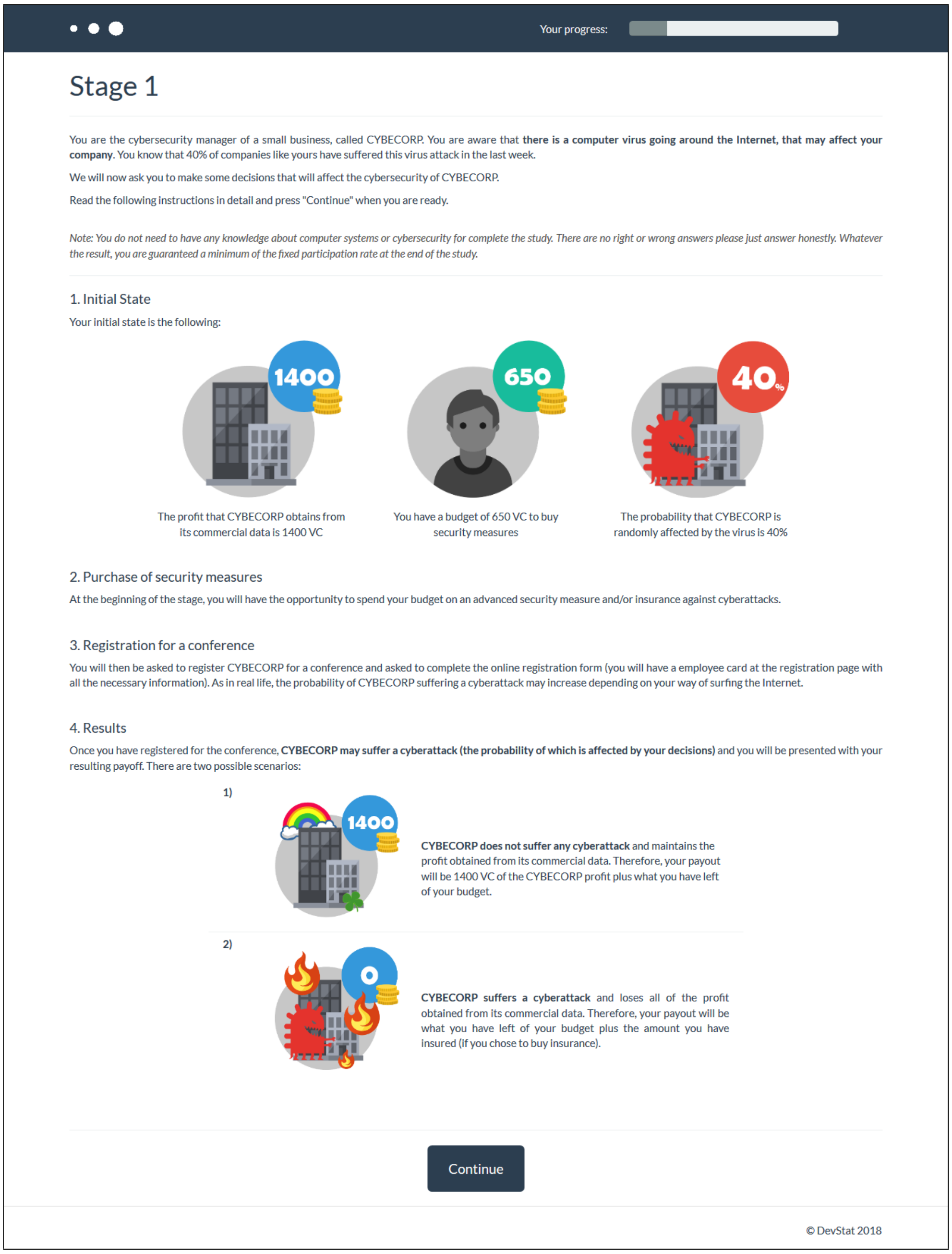
Once you have registered for the conference, **CYBECORP may suffer a cyberattack (the probability of which is affected by your decisions)** and you will be presented with your resulting payoff. There are two possible scenarios:

1) **CYBECORP does not suffer any cyberattack** and maintains the profit obtained from its commercial data. Therefore, your payout will be 1400 VC of the CYBECORP profit plus what you have left of your budget.

2) **CYBECORP suffers a cyberattack** and loses all of the profit obtained from its commercial data. Therefore, your payout will be what you have left of your budget plus the amount you have insured (if you chose to buy insurance).

© DevStat 2018

Figure 1.A.3: Instructions when the context is random, $C = 1$.



Stage 1

You are the cybersecurity manager of a small business, called CYBECORP. You are aware that **there is a computer virus going around the Internet, that may affect your company**. You know that 40% of companies like yours have suffered this virus attack in the last week.

We will now ask you to make some decisions that will affect the cybersecurity of CYBECORP.

Read the following instructions in detail and press "Continue" when you are ready.

Note: You do not need to have any knowledge about computer systems or cybersecurity for complete the study. There are no right or wrong answers please just answer honestly. Whatever the result, you are guaranteed a minimum of the fixed participation rate at the end of the study.

1. Initial State

Your initial state is the following:

The profit that CYBECORP obtains from its commercial data is 1400 VC

You have a budget of 650 VC to buy security measures

The probability that CYBECORP is randomly affected by the virus is 40%

2. Purchase of security measures

At the beginning of the stage, you will have the opportunity to spend your budget on an advanced security measure and/or insurance against cyberattacks.

3. Registration for a conference

You will then be asked to register CYBECORP for a conference and asked to complete the online registration form (you will have a employee card at the registration page with all the necessary information). As in real life, the probability of CYBECORP suffering a cyberattack may increase depending on your way of surfing the Internet.

4. Results

Once you have registered for the conference, **CYBECORP may suffer a cyberattack (the probability of which is affected by your decisions)** and you will be presented with your resulting payoff. There are two possible scenarios:

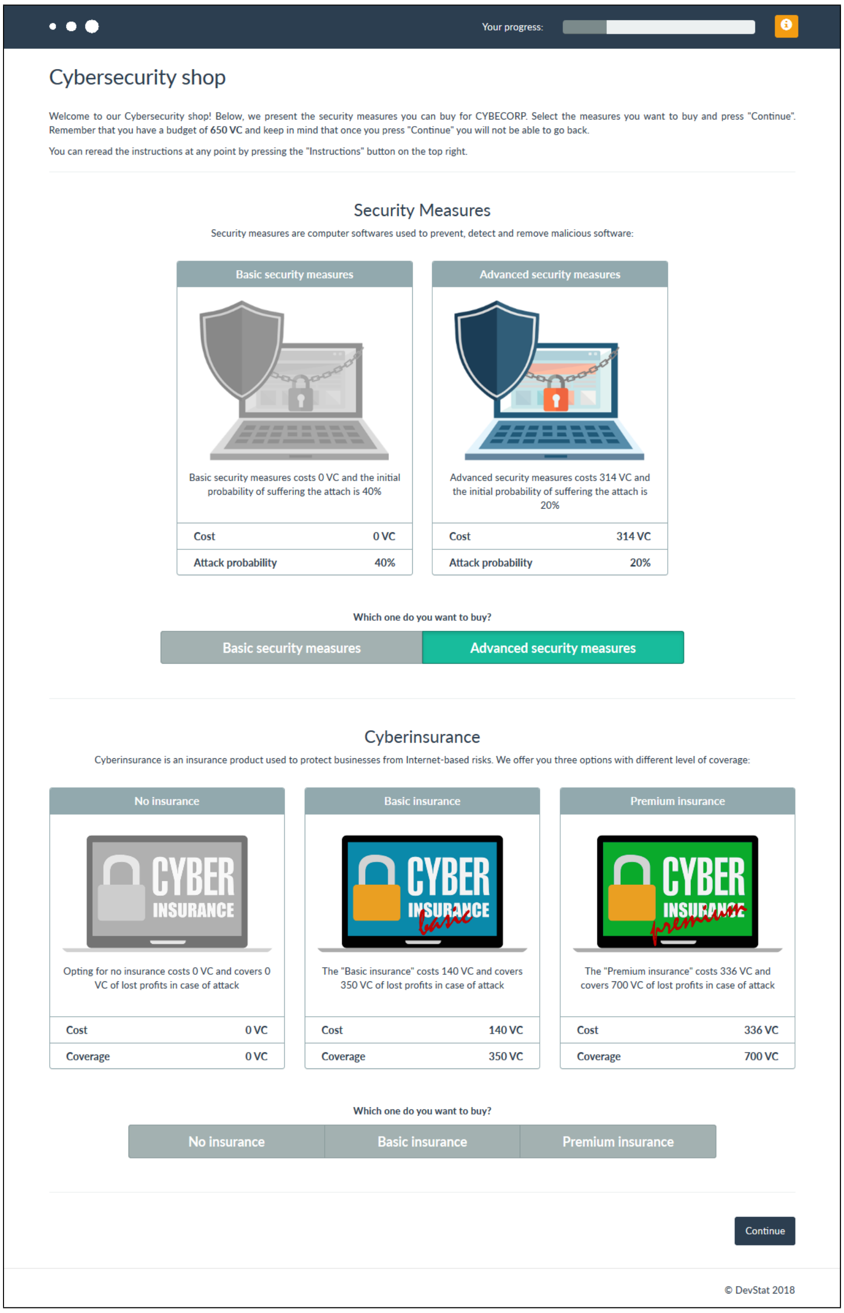
1) **CYBECORP does not suffer any cyberattack** and maintains the profit obtained from its commercial data. Therefore, your payout will be 1400 VC of the CYBECORP profit plus what you have left of your budget.

2) **CYBECORP suffers a cyberattack** and loses all of the profit obtained from its commercial data. Therefore, your payout will be what you have left of your budget plus the amount you have insured (if you chose to buy insurance).

Continue

© DevStat 2018

Figure 1.A.4: Instructions when the context is intentional, $C = 2$.



Cybersecurity shop

Welcome to our Cybersecurity shop! Below, we present the security measures you can buy for CYBECORP. Select the measures you want to buy and press "Continue". Remember that you have a budget of **650 VC** and keep in mind that once you press "Continue" you will not be able to go back.

You can reread the instructions at any point by pressing the "Instructions" button on the top right.

Security Measures

Security measures are computer softwares used to prevent, detect and remove malicious software:

Basic security measures

Basic security measures costs 0 VC and the initial probability of suffering the attach is 40%

Cost	0 VC
Attack probability	40%

Advanced security measures

Advanced security measures costs 314 VC and the initial probability of suffering the attach is 20%

Cost	314 VC
Attack probability	20%

Which one do you want to buy?

Cyberinsurance

Cyberinsurance is an insurance product used to protect businesses from Internet-based risks. We offer you three options with different level of coverage:

No insurance

Opting for no insurance costs 0 VC and covers 0 VC of lost profits in case of attack

Cost	0 VC
Coverage	0 VC

Basic insurance

The "Basic insurance" costs 140 VC and covers 350 VC of lost profits in case of attack

Cost	140 VC
Coverage	350 VC

Premium insurance

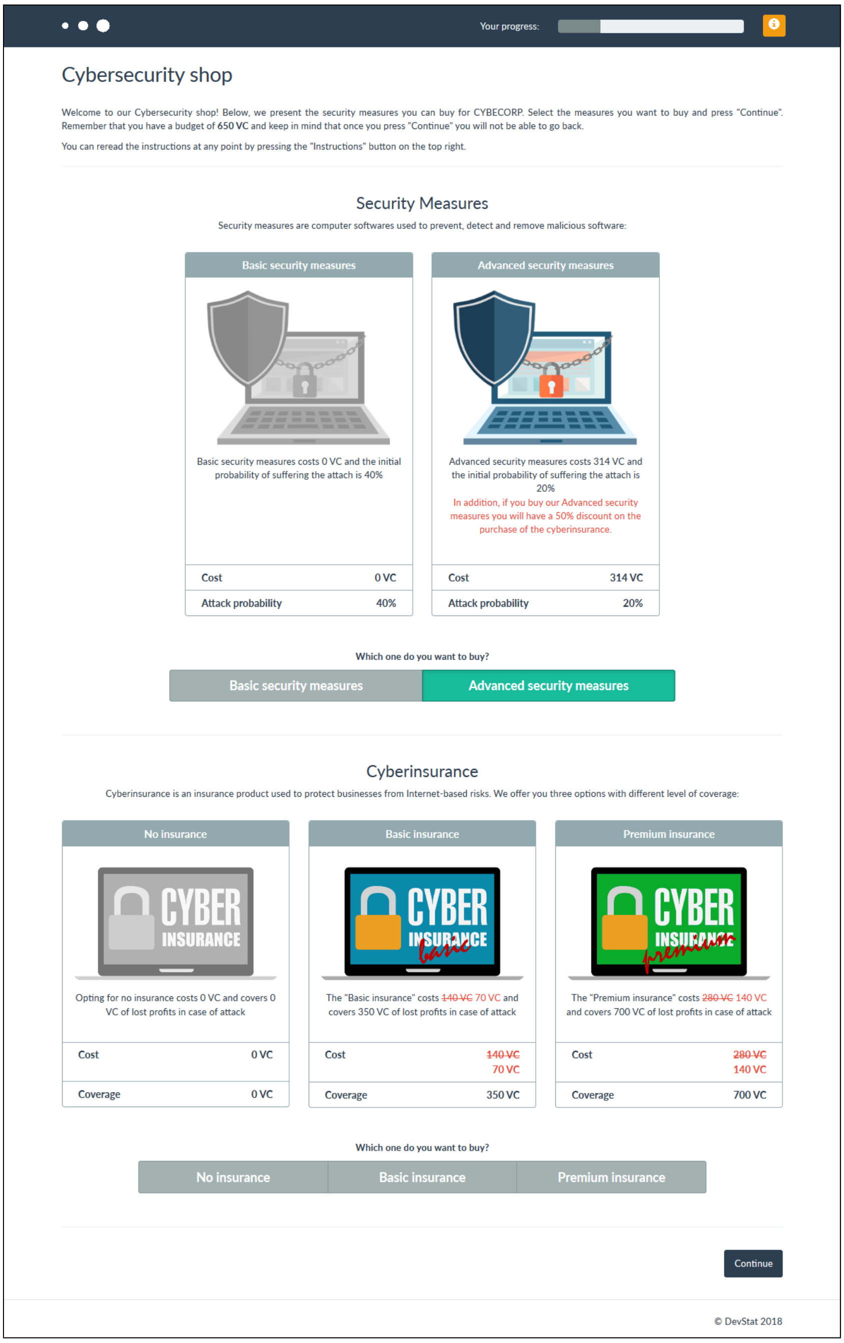
The "Premium insurance" costs 336 VC and covers 700 VC of lost profits in case of attack

Cost	336 VC
Coverage	700 VC

Which one do you want to buy?

© DevStat 2018

Figure 1.A.5: Cibersecurity shop when there are not price dependency and the prices of insurance are medium, $P = 1$ and $I = 1$.



Cybersecurity shop

Welcome to our Cybersecurity shop! Below, we present the security measures you can buy for CYBECORP. Select the measures you want to buy and press "Continue". Remember that you have a budget of **650 VC** and keep in mind that once you press "Continue" you will not be able to go back.

You can reread the instructions at any point by pressing the "Instructions" button on the top right.

Security Measures

Security measures are computer softwares used to prevent, detect and remove malicious software:

Basic security measures

Basic security measures costs 0 VC and the initial probability of suffering the attach is 40%

Cost	0 VC
Attack probability	40%

Advanced security measures

Advanced security measures costs 314 VC and the initial probability of suffering the attach is 20%

In addition, if you buy our Advanced security measures you will have a 50% discount on the purchase of the cyberinsurance.

Cost	314 VC
Attack probability	20%

Which one do you want to buy?

Cyberinsurance

Cyberinsurance is an insurance product used to protect businesses from Internet-based risks. We offer you three options with different level of coverage:

No insurance

Opting for no insurance costs 0 VC and covers 0 VC of lost profits in case of attack

Cost	0 VC
Coverage	0 VC

Basic insurance

The "Basic insurance" costs ~~140 VC~~ 70 VC and covers 350 VC of lost profits in case of attack

Cost	~~140 VC~~ 70 VC
Coverage	350 VC

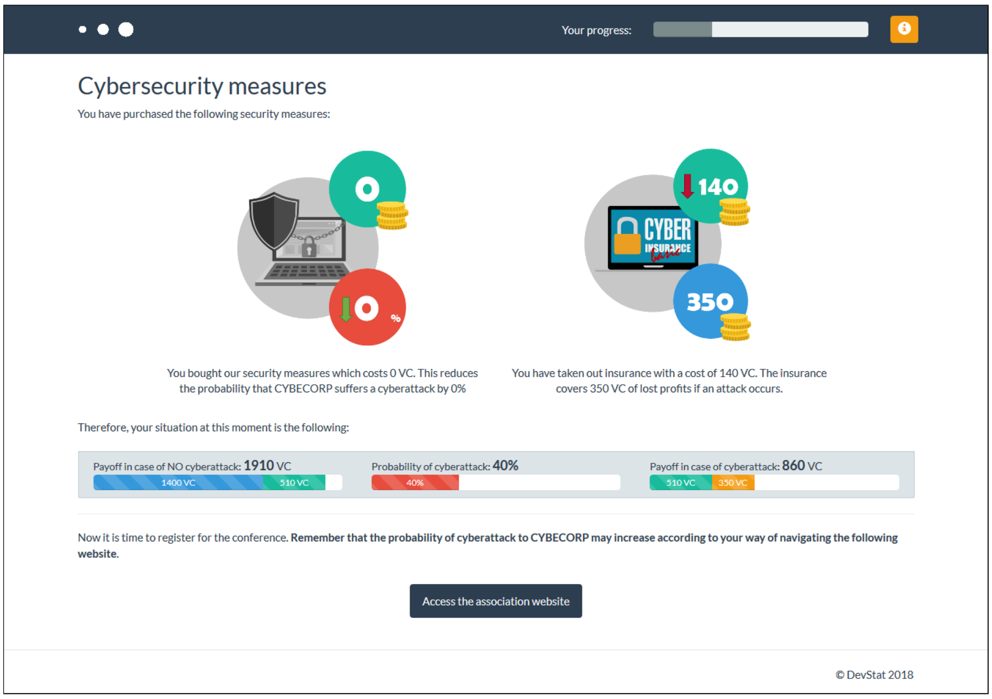
Premium insurance

The "Premium insurance" costs ~~280 VC~~ 140 VC and covers 700 VC of lost profits in case of attack

Cost	~~280 VC~~ 140 VC
Coverage	700 VC

Which one do you want to buy?

Figure 1.A.6: Cibersecurity shop when there are price dependency and the prices of insurance are medium, $P = 2$ and $I = 1$.



Cybersecurity measures

You have purchased the following security measures:

You bought our security measures which costs 0 VC. This reduces the probability that CYBECORP suffers a cyberattack by 0%

You have taken out insurance with a cost of 140 VC. The insurance covers 350 VC of lost profits if an attack occurs.

Therefore, your situation at this moment is the following:

Payoff in case of NO cyberattack: **1910** VC
1400 VC | 510 VC

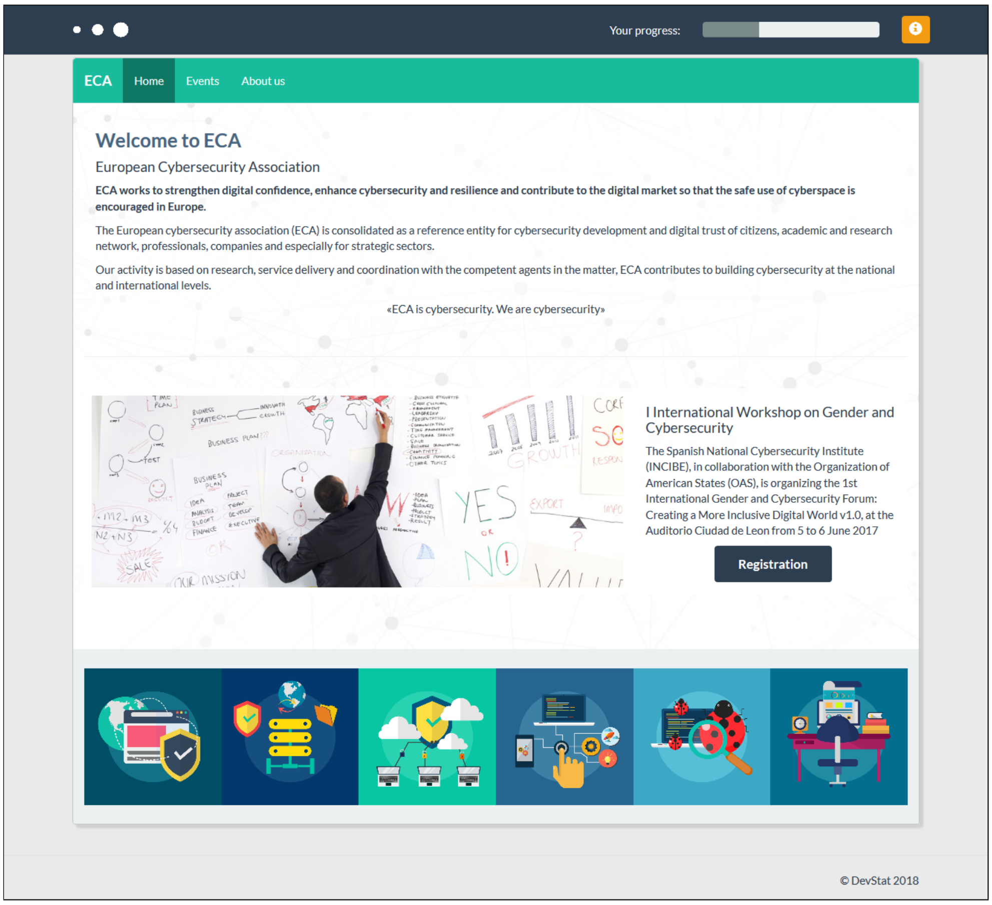
Probability of cyberattack: **40%**
40%

Payoff in case of cyberattack: **860** VC
510 VC | 350 VC

Now it is time to register for the conference. **Remember that the probability of cyberattack to CYBECORP may increase according to your way of navigating the following website.**

© DevStat 2018

Figure 1.A.7: Purchase summary.



Welcome to ECA

European Cybersecurity Association

ECA works to strengthen digital confidence, enhance cybersecurity and resilience and contribute to the digital market so that the safe use of cyberspace is encouraged in Europe.

The European cybersecurity association (ECA) is consolidated as a reference entity for cybersecurity development and digital trust of citizens, academic and research network, professionals, companies and especially for strategic sectors.

Our activity is based on research, service delivery and coordination with the competent agents in the matter, ECA contributes to building cybersecurity at the national and international levels.

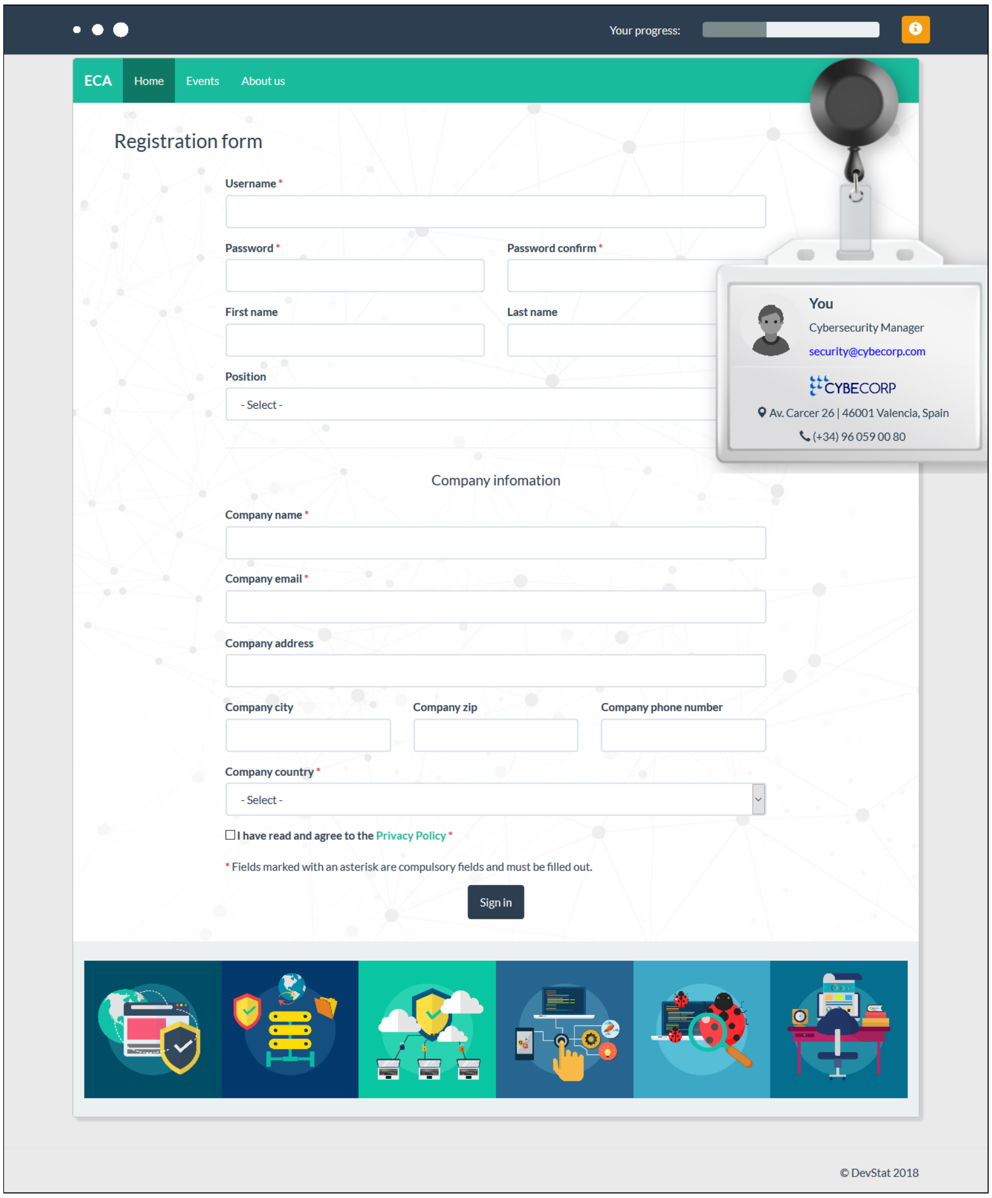
«ECA is cybersecurity. We are cybersecurity»

I International Workshop on Gender and Cybersecurity

The Spanish National Cybersecurity Institute (INCIBE), in collaboration with the Organization of American States (OAS), is organizing the 1st International Gender and Cybersecurity Forum: Creating a More Inclusive Digital World v1.0, at the Auditorio Ciudad de Leon from 5 to 6 June 2017

© DevStat 2018

Figure 1.A.8: Event website.



Registration form

Username *

Password *

Password confirm *

First name

Last name

Position
- Select -

You
Cybersecurity Manager
security@cybecorp.com

CYBECORP
Av. Carcer 26 | 46001 Valencia, Spain
(+34) 96 059 00 80

Company infomation

Company name *

Company email *

Company address

Company city

Company zip

Company phone number

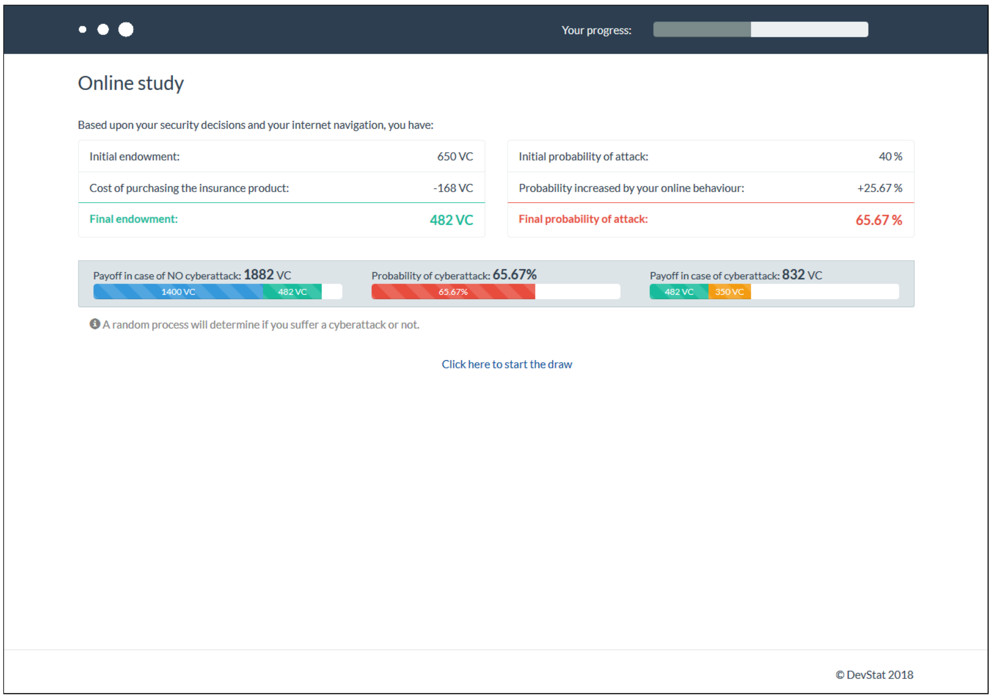
Company country *
- Select -

□ I have read and agree to the Privacy Policy *

* Fields marked with an asterisk are compulsory fields and must be filled out.

© DevStat 2018

Figure 1.A.9: Event registration.



Online study

Based upon your security decisions and your internet navigation, you have:

Initial endowment:	650 VC
Cost of purchasing the insurance product:	-168 VC
Final endowment:	**482 VC**

Initial probability of attack:	40 %
Probability increased by your online behaviour:	+25.67 %
Final probability of attack:	**65.67 %**

Payoff in case of NO cyberattack: **1882** VC

Probability of cyberattack: **65.67%**

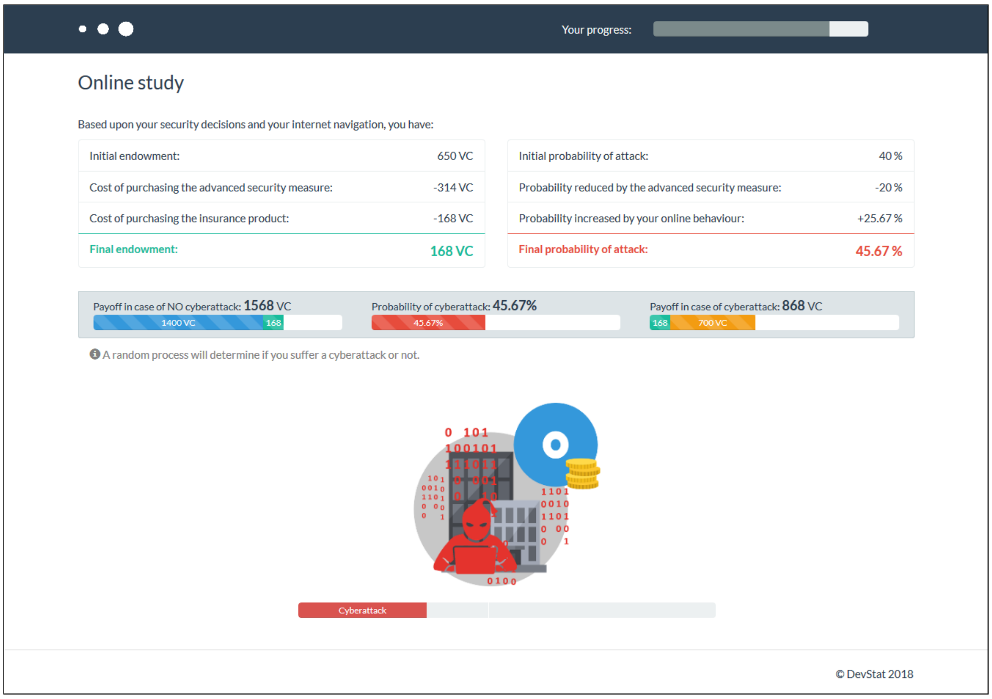
Payoff in case of cyberattack: **832** VC

ℹ A random process will determine if you suffer a cyberattack or not.

Click here to start the draw

© DevStat 2018

Figure 1.A.10: Event website - Logout.



Online study

Based upon your security decisions and your internet navigation, you have:

Initial endowment:	650 VC
Cost of purchasing the advanced security measure:	-314 VC
Cost of purchasing the insurance product:	-168 VC
Final endowment:	**168 VC**

Initial probability of attack:	40 %
Probability reduced by the advanced security measure:	-20 %
Probability increased by your online behaviour:	+25.67 %
Final probability of attack:	**45.67 %**

Payoff in case of NO cyberattack: **1568** VC

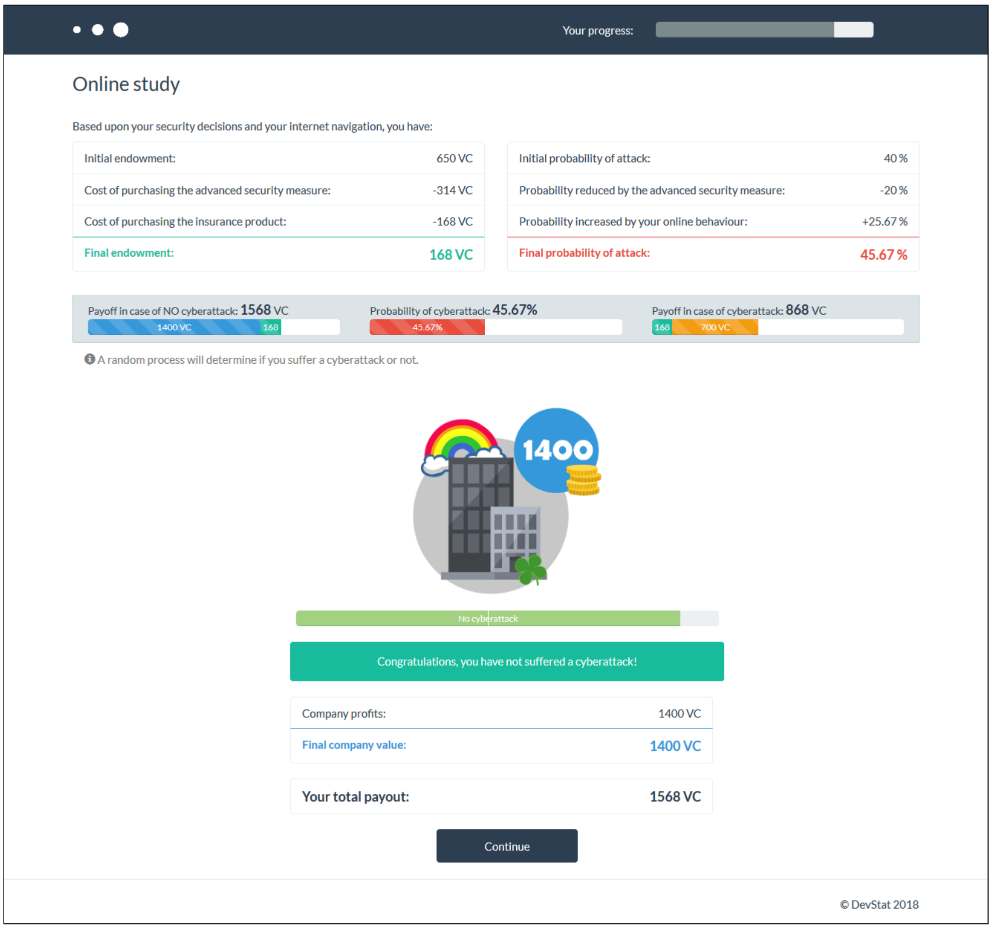
Probability of cyberattack: **45.67%**

Payoff in case of cyberattack: **868** VC

ℹ A random process will determine if you suffer a cyberattack or not.

© DevStat 2018

Figure 1.A.11: Cyberattack simulation - Cyberattack.



Online study

Based upon your security decisions and your internet navigation, you have:

Initial endowment:	650 VC
Cost of purchasing the advanced security measure:	-314 VC
Cost of purchasing the insurance product:	-168 VC
Final endowment:	**168 VC**

Initial probability of attack:	40 %
Probability reduced by the advanced security measure:	-20 %
Probability increased by your online behaviour:	+25.67 %
Final probability of attack:	**45.67 %**

Payoff in case of NO cyberattack: **1568** VC

Probability of cyberattack: **45.67%**

Payoff in case of cyberattack: **868** VC

ℹ A random process will determine if you suffer a cyberattack or not.

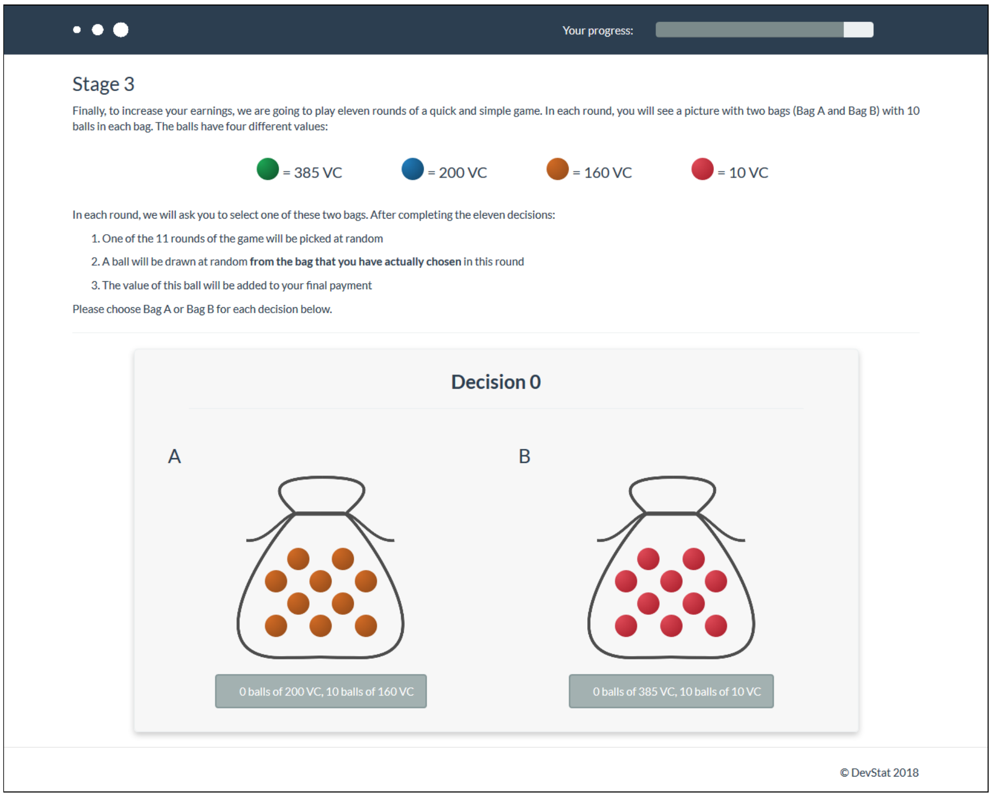
Congratulations, you have not suffered a cyberattack!

Company profits:	1400 VC
Final company value:	**1400 VC**

Your total payout:	**1568 VC**

© DevStat 2018

Figure 1.A.12: Cyberattack simulation - No Cyberattack.



Stage 3

Finally, to increase your earnings, we are going to play eleven rounds of a quick and simple game. In each round, you will see a picture with two bags (Bag A and Bag B) with 10 balls in each bag. The balls have four different values:

🟢 = 385 VC 🔵 = 200 VC 🟠 = 160 VC 🔴 = 10 VC

In each round, we will ask you to select one of these two bags. After completing the eleven decisions:

1. One of the 11 rounds of the game will be picked at random
2. A ball will be drawn at random **from the bag that you have actually chosen** in this round
3. The value of this ball will be added to your final payment

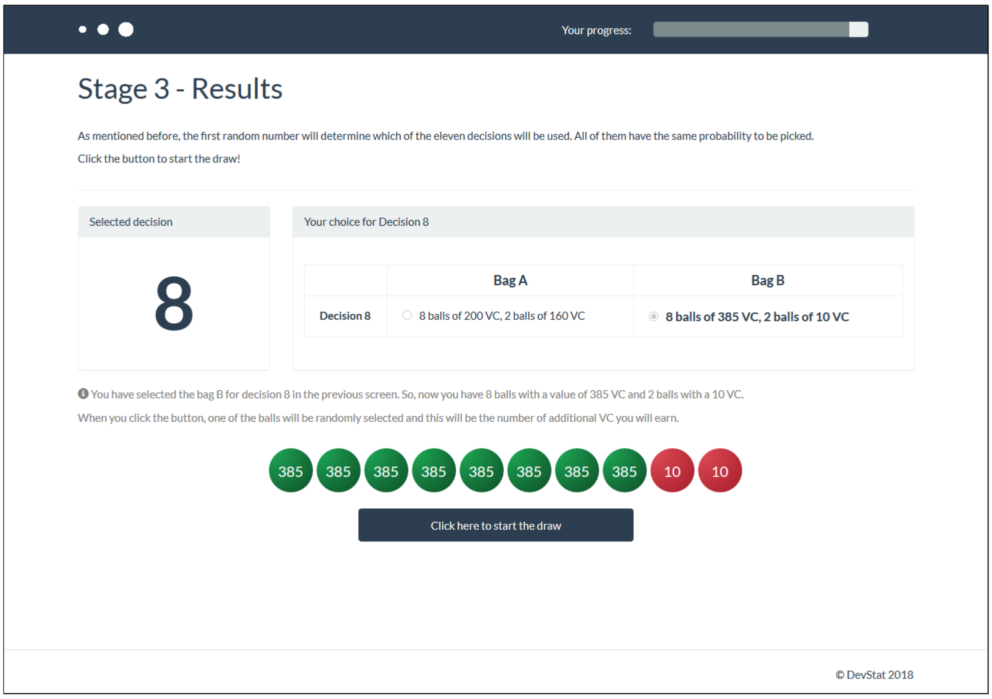
Please choose Bag A or Bag B for each decision below.

Decision 0

A
B
0 balls of 200 VC, 10 balls of 160 VC
0 balls of 385 VC, 10 balls of 10 VC

© DevStat 2018

Figure 1.A.13: Holt & Laury experiment.



Stage 3 - Results

As mentioned before, the first random number will determine which of the eleven decisions will be used. All of them have the same probability to be picked.

Click the button to start the draw!

Selected decision

8

Your choice for Decision 8

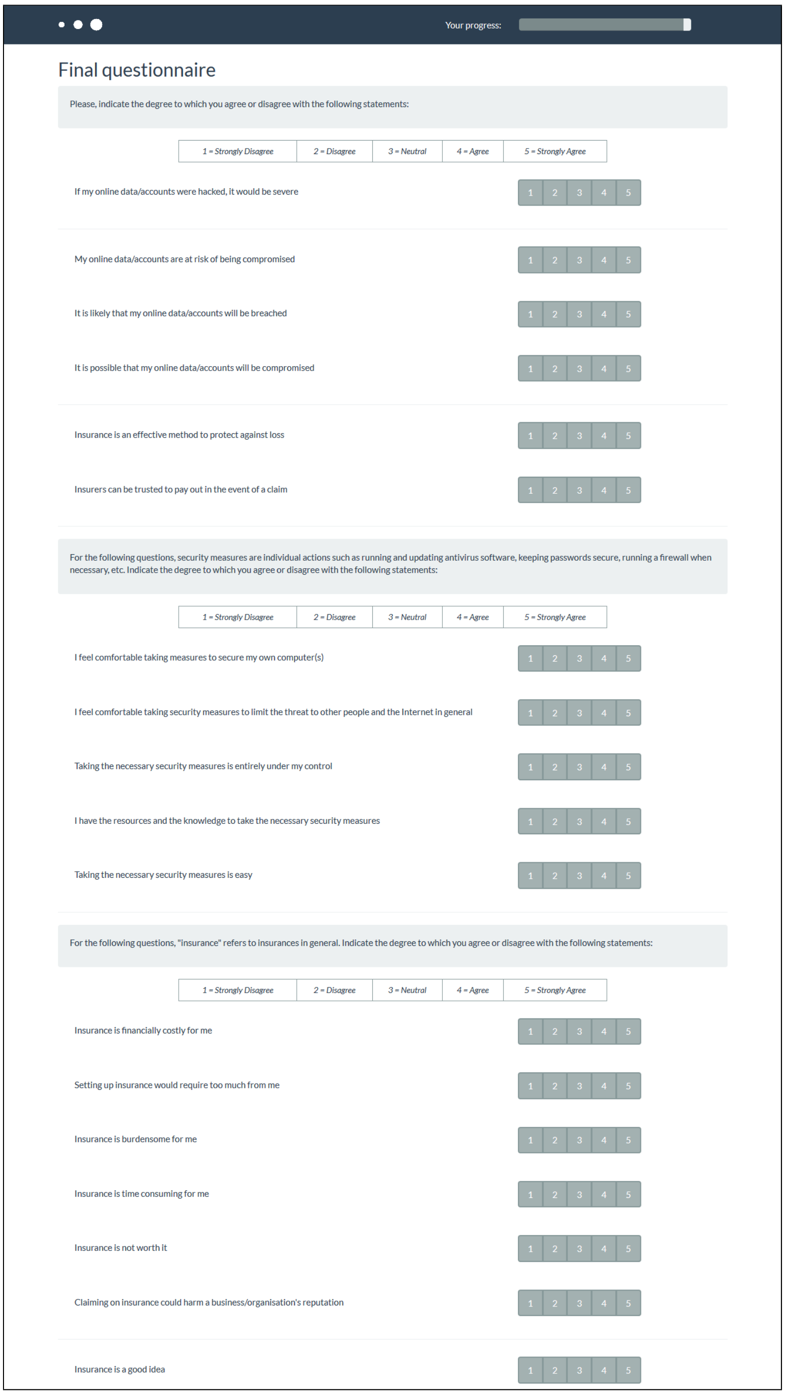
	Bag A	Bag B
Decision 8	○ 8 balls of 200 VC, 2 balls of 160 VC	● **8 balls of 385 VC, 2 balls of 10 VC**

You have selected the bag B for decision 8 in the previous screen. So, now you have 8 balls with a value of 385 VC and 2 balls with a 10 VC.

When you click the button, one of the balls will be randomly selected and this will be the number of additional VC you will earn.

© DevStat 2018

Figure 1.A.14: Holt & Laury experiment results.



Final questionnaire

Please, indicate the degree to which you agree or disagree with the following statements:

1 = Strongly Disagree	2 = Disagree	3 = Neutral	4 = Agree	5 = Strongly Agree

If my online data/accounts were hacked, it would be severe 1 2 3 4 5

My online data/accounts are at risk of being compromised 1 2 3 4 5

It is likely that my online data/accounts will be breached 1 2 3 4 5

It is possible that my online data/accounts will be compromised 1 2 3 4 5

Insurance is an effective method to protect against loss 1 2 3 4 5

Insurers can be trusted to pay out in the event of a claim 1 2 3 4 5

For the following questions, security measures are individual actions such as running and updating antivirus software, keeping passwords secure, running a firewall when necessary, etc. Indicate the degree to which you agree or disagree with the following statements:

1 = Strongly Disagree	2 = Disagree	3 = Neutral	4 = Agree	5 = Strongly Agree

I feel comfortable taking measures to secure my own computer(s) 1 2 3 4 5

I feel comfortable taking security measures to limit the threat to other people and the Internet in general 1 2 3 4 5

Taking the necessary security measures is entirely under my control 1 2 3 4 5

I have the resources and the knowledge to take the necessary security measures 1 2 3 4 5

Taking the necessary security measures is easy 1 2 3 4 5

For the following questions, "insurance" refers to insurances in general. Indicate the degree to which you agree or disagree with the following statements:

1 = Strongly Disagree	2 = Disagree	3 = Neutral	4 = Agree	5 = Strongly Agree

Insurance is financially costly for me 1 2 3 4 5

Setting up insurance would require too much from me 1 2 3 4 5

Insurance is burdensome for me 1 2 3 4 5

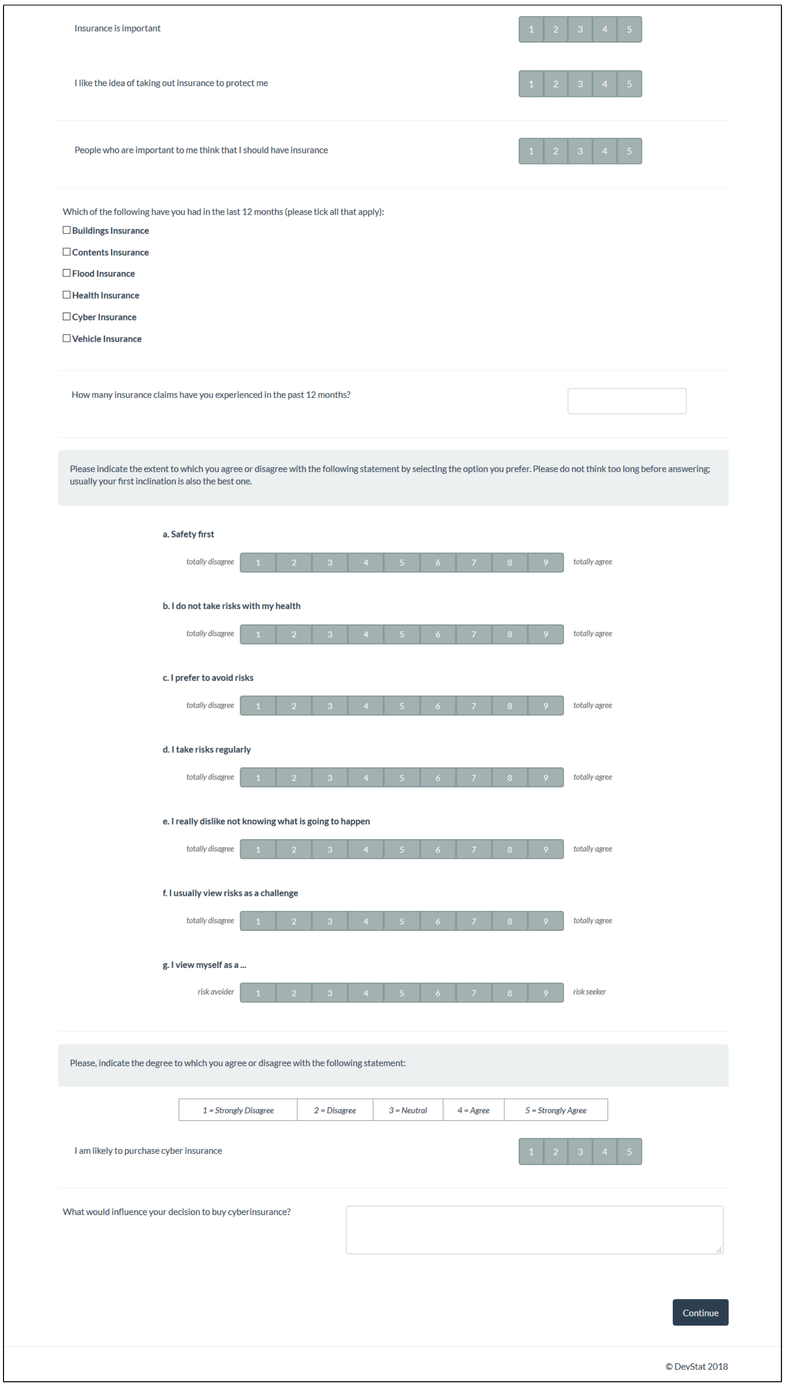
Insurance is time consuming for me 1 2 3 4 5

Insurance is not worth it 1 2 3 4 5

Claiming on insurance could harm a business/organisation's reputation 1 2 3 4 5

Insurance is a good idea 1 2 3 4 5

Figure 1.A.15: Final questionnaire (1 of 2).



Insurance is important 1 2 3 4 5

I like the idea of taking out insurance to protect me 1 2 3 4 5

People who are important to me think that I should have insurance 1 2 3 4 5

Which of the following have you had in the last 12 months (please tick all that apply):

□ **Buildings Insurance**
□ **Contents Insurance**
□ **Flood Insurance**
□ **Health Insurance**
□ **Cyber Insurance**
□ **Vehicle Insurance**

How many insurance claims have you experienced in the past 12 months?

Please indicate the extent to which you agree or disagree with the following statement by selecting the option you prefer. Please do not think too long before answering; usually your first inclination is also the best one.

a. Safety first

totally disagree 1 2 3 4 5 6 7 8 9 *totally agree*

b. I do not take risks with my health

totally disagree 1 2 3 4 5 6 7 8 9 *totally agree*

c. I prefer to avoid risks

totally disagree 1 2 3 4 5 6 7 8 9 *totally agree*

d. I take risks regularly

totally disagree 1 2 3 4 5 6 7 8 9 *totally agree*

e. I really dislike not knowing what is going to happen

totally disagree 1 2 3 4 5 6 7 8 9 *totally agree*

f. I usually view risks as a challenge

totally disagree 1 2 3 4 5 6 7 8 9 *totally agree*

g. I view myself as a ...

risk avoider 1 2 3 4 5 6 7 8 9 *risk seeker*

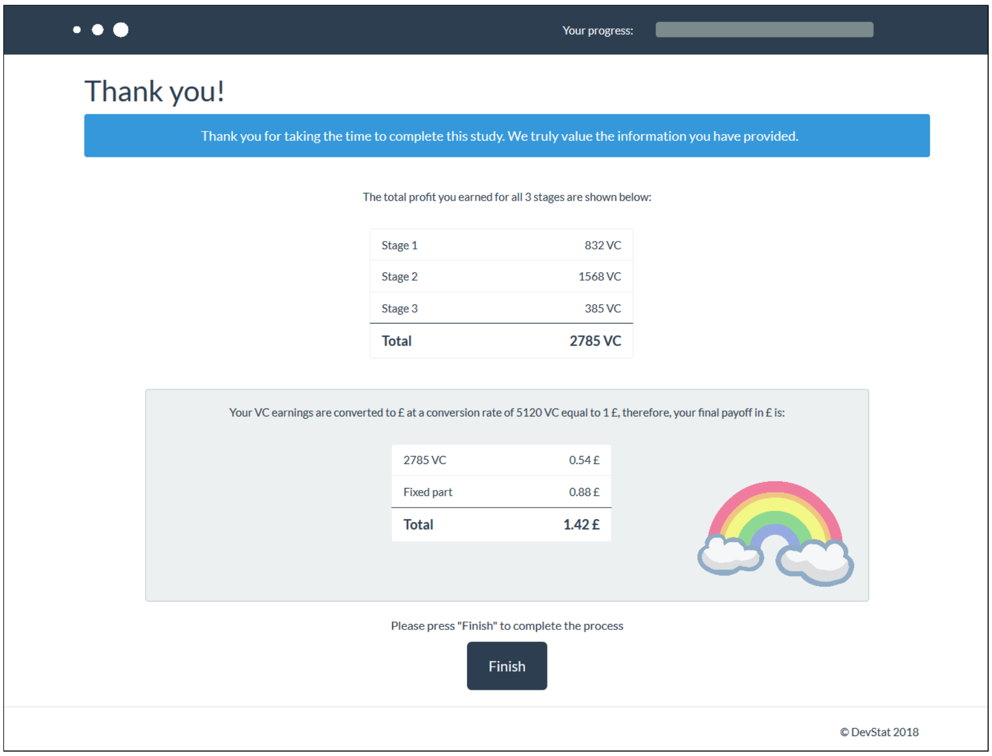
Please, indicate the degree to which you agree or disagree with the following statement:

1 = Strongly Disagree	*2 = Disagree*	*3 = Neutral*	*4 = Agree*	*5 = Strongly Agree*

I am likely to purchase cyber insurance 1 2 3 4 5

What would influence your decision to buy cyberinsurance?

Figure 1.A.16: Final questionnaire (2 of 2).



Thank you!

Thank you for taking the time to complete this study. We truly value the information you have provided.

The total profit you earned for all 3 stages are shown below:

Stage 1	832 VC
Stage 2	1568 VC
Stage 3	385 VC
Total	**2785 VC**

Your VC earnings are converted to £ at a conversion rate of 5120 VC equal to 1 £, therefore, your final payoff in £ is:

2785 VC	0.54 £
Fixed part	0.88 £
Total	**1.42 £**

Please press "Finish" to complete the process

Figure 1.A.17: Payouts page.

Chapter 2

Developing and validating a behavioural model of cyberinsurance adoption

Abstract

Business disruption from cyberattacks is a recognised and growing concern, yet the uptake of cyberinsurance has been relatively low. This study proposed and tested a predictive model of cyberinsurance adoption, incorporating elements of Protection Motivation Theory (PMT) and the Theory of Planned Behaviour (TPB) as well as factors in relation to risk propensity and price. Data was obtained from an online behavioural economics experiment with 4,800 participants across four EU countries. During the experiment, participants were given the opportunity to purchase different protection measures and cyberinsurance

products before performing an online task. Some participants then suffered a cyberattack in the experimental setup, the probability of which was dependent upon their adoption of protection measures and their behaviour during the online task. The consequences of this attack were in turn dependent upon their cyberinsurance purchase decisions (i.e., basic vs premium insurance purchase). Structural Equation Modelling (SEM) was applied and the model was further developed to include elements of the wider security ecosystem. The resulting model shows that all TPB factors, but only response efficacy from the PMT factors positively predicted adoption of premium cyberinsurance. Premium insurance adoption was also influenced by security measure adoption, individual propensity for risk, and the price differential between basic and premium products. Interestingly, adoption of cybersecurity measures was associated with safer behaviour online, contrary to concerns of 'moral hazard'. The findings highlight the need to consider the larger cybersecurity ecosystem when designing interventions to increase adoption of cyberinsurance and/or promote more secure online behaviour.

A conceptual map of this chapter is presented in Figure 7.

2.1 Introduction

Widespread cyberinsurance adoption has a number of potential benefits in a society facing increasing cybersecurity risk. It could lead to market-based management of that risk, acting as a mechanism for spreading the risk amongst multiple stakeholders. Also, since obtaining insurance requires that certain standards are met, it could act as an incentive towards organisational investments in information security; which would reduce risk for the investing organisation and for their wider network. Insurance investigators follow up on serious incidents to learn what went wrong, therefore uptake could also lead to data aggregation on best practices and better tools for assessing security – something that is currently lacking. In principle, a robust cyberinsurance offer could strengthen IT security for society as a whole (Baer and Parkinson, 2007; Kuru and Bayraktar, 2017). However, despite the growing risk of cyberattack, uptake of cyberinsurance as a mechanism to ameliorate risk (financial and otherwise) has not reached expectations, with some research reporting uptake rates as low as 10% in the UK (Low, 2017).

Recently, a number of studies have used psychological models that define the relationships between attitudes, intentions and behaviours, to understand more about insurance uptake. For example, Dittrich et al. (2016) used Protection Motivation Theory (PMT) to predict uptake of flood insurance, whilst Brahmana et al. (2018) applied a different model – the Theory of Planned Behaviour (TPB) to explore intention to purchase health insurance. These two models have also been widely used to assess cybersecurity vulnerability and explore behavioural intentions to engage in secure practices or comply with organisational Information Security Policy (ISP). Indeed, Lebek et al. (2014) have argued that TPB and PMT constitute two of the most significant behavioural models for understanding ISP compliance. We describe these two models below, outlining their putative relationship to cyberinsurance uptake and then go on to a more careful critique of the predictive power of their constitutent factors. We use this information to develop a hypothetical research model for cyberinsurance adoption.

2.2 Theoretical framework

2.2.1 Protection Motivation Theory

PMT was originally designed to explain engagement in protective actions in relation to health-related behaviours. However, as aforementioned, the theory has since been applied to the explanation of other protective actions, including uptake of insurance (Dittrich et al., 2016; Beck, 1984; Grahn and Jaldell, 2019) and the adoption of secure online behaviours (Tsai et al., 2016; van Bavel et al., 2019). PMT proposes that people protect themselves by making both a threat and a coping appraisal. The threat appraisal is dependent upon both the perceived severity of a threatening event (in this instance a cyberattack) and the perceived vulnerability to the event (i.e. the perceived probability of that event occurring). The coping appraisal reflects the perceived *efficacy* of the recommended protective behaviour (cyberinsurance adoption in this case) and the individual's perceived *self-efficacy*. Two other factors are present in the model: The threat of a particular behaviour is weighed up against the rewards of that behaviour (in health, for example, the health threats associated with smoking are traded against the perceived rewards of smoking) and the costs of the coping action are also a factor (in terms of time, effort or actual finances required to engage the protective action). Therefore, an individual(s) considering whether to invest in cyberinsurance may firstly weigh up the likelihood that they will receive a cyberattack of a particular severity against (a)

The cost of taking out cyberinsurance (finances, time, effort) and (b) How effective they believe that insurance will be (response efficacy) and/or how much confidence they have in their own ability to put insurance measures into place (self-efficacy) (see Figure 2.2.1).

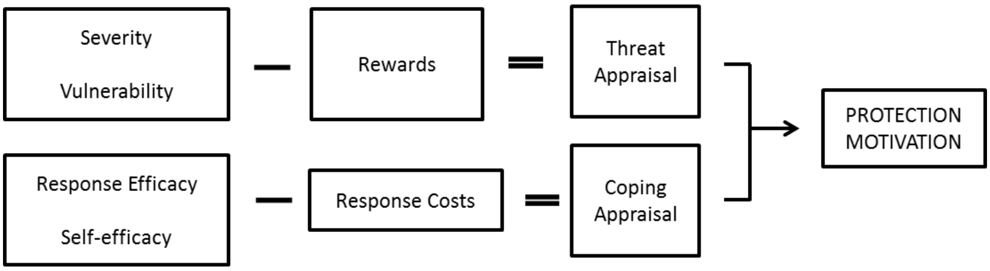


Figure 2.2.1: Protection Motivation Theory.

2.2.2 Theory of Planned Behaviour

TPB taps into one of the same constructs as PMT (as perceived self-efficacy and perceived behavioural control are thought to measure the same construct (Ifinedo, 2012)). However, TPB also highlights additional factors which may influence insurance purchase decisions. TPB states that intention to perform a behaviour is the most immediate and important determination of behaviour (Ajzen et al., 1991). Intention is influenced by the individual's attitude(s) towards the behaviour, subjective norm(s) and perceived control over the situation. This theory suggests that strengthening positive attitudes towards cyberinsurance (e.g. strengthening the belief that insurance companies would pay out in the event of a cyberincident) could increase cyberinsurance uptake. Likewise strengthening perceived social norms around cyberinsurance may help to increase uptake (e.g., strengthening the perception that others believe cyberinsurance to be a worthwhile product) (see Figure 2.2.2).

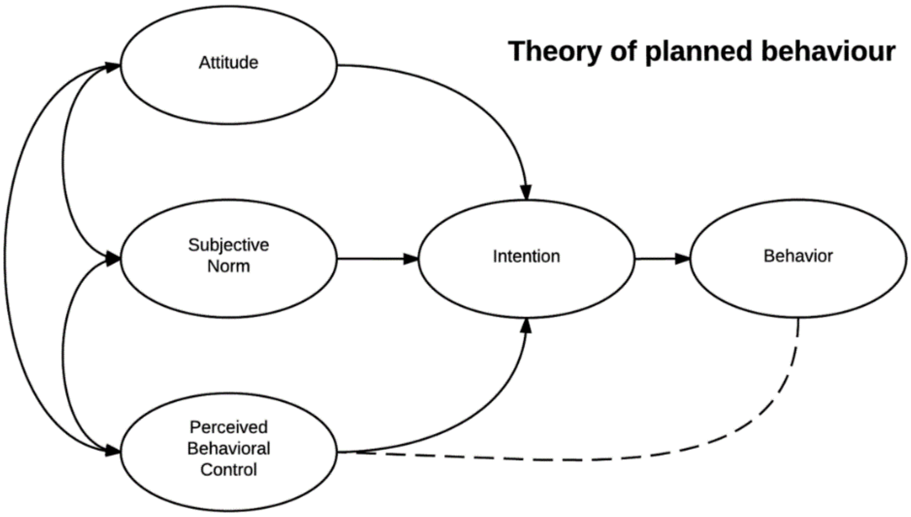


Figure 2.2.2: Theory of Planned Behaviour.

2.2.3 Combining PMT and TPB in our research model

The two models are often used together, sometimes in combination, with one of the most cited studies (Ifinedo, 2012) demonstrating that the inclusion of PMT constructs to the TPB model could improve the explained variance in ISP compliance from 0.60 to 0.70. Sommestad et al. (2015) also asked whether TPB was sufficient to account for cybersecurity policy compliance in employees of the Swedish Defence Agency, concluding that TPB alone accounted for 0.36 of the variance in intention to comply and 0.44 of the variance in reported current behaviour, but noting that the regression model could be improved by the addition of threat appraisal constructs from PMT.

Although widely used, these models are not without criticism. A systematic review by ENISA (2018)[pp. 11] found that the coping elements of PMT and TPB were useful, but questioned the predictive value of threat models, including the threat appraisal in PMT. This is curious when we consider that Sommestad et al. (2014) found added predictive value in the threat component. Subsequently, Sommestad et al. (2015) also point out that no single variable in the PMT is able to explain more than a small portion of the variance exhibited within the studied populations. This is well in line with the underlying idea of PMT, which describes how six variables together determine intentions through cognitive processes. This paper also points to inconsistency in the way that constructs are measured could lead to the discrepancy between studies. We should also bear in mind that for the ENISA (2018) report, mentioned earlier, the search terms included cybersecurity items but did not include 'Security', nor 'Information Security Policy' and so missed some of the studies cited above. That said, other recent work has also suggested that coping elements offer greater value than threat elements when trying to predict or improve online security behaviour (van Bavel et al., 2019).

In drawing up our hypothetical research model for the purchase of cyberinsurance (Figure 2.2.3), we have thus included all factors from PMT (Threat appraisal: perceived severity, perceived vulnerability, Coping appraisal: response efficacy, self-efficacy, response costs) and TPB (attitude and subjective norm) but have noted where the hypothesised links are weaker, drawing these as dashed lines in the model. Thus whilst we feel confident in the hypothesis that the coping appraisal factors should be influential, we are less certain about the predictive power of the threat appraisal components, based on the literature described above.

There are a number of other elements in our research model with a rationale as follows. Firstly, we include risk propensity (also referred to as risk preference or risk tolerance). Previous insurance research has shown that risk adverse individuals are more likely to purchase flood insurance (Botzen and van den Bergh, 2012; Petrolia et al., 2013), health insurance and life insurance (Barsky et al., 1997; Lammers et al., 2010). This makes sense, since insurance represents a means of risk mitigation. In our study we measure risk propensity using the seven item Risk Propensity Scale (Meertens and Lion, 2008) with the hypothesis that individuals who are more risk adverse would be more likely to purchase insurance, whereas individuals who are more willing to take risks may be less inclined to adopt insurance.

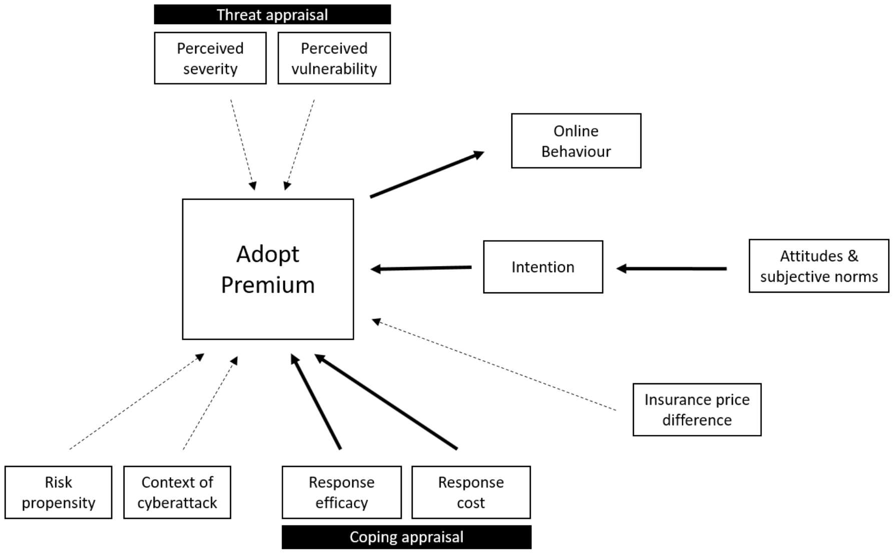


Figure 2.2.3: The Research Model. Strong hypothesised links are shown as solid lines. Weaker or less supported hypothesised links are showed as dashed lines.

We also include a factor that takes price into account. Cyberinsurance is likely to adopt a heterogeneous pricing model, in part because the measures company's take to mitigate threats will vary. Under such circumstances, consumer perceptions about what price is reasonable and fair will vary, but is likely to influence willingness to adopt both protective measures and a premium insurance product (Pal et al., 2017). Cyberinsurers will also want to reduce systemic risk (also known as correlated or aggregate risk) across their portfolio to avoid catastrophic losses that may arise from the interdependencies of networked organisations. Khalili et al. (2019) have suggested that this can be partly be achieved by setting the price to incentivise purchase of premium insurance products that themselves may be contingent upon the company's security posture. In terms of our own research model, we manipulated price to be either dependent or independent of the security measures in place (i.e., in the dependent category price of the insurance policy varied dependent upon the cybersecurity measures the individual had opted for). This allowed an empirical assessment of the extent to which making insurance premiums contingent upon a company's security posture would improve or limit cyberinsurance uptake.

Finally, we added one further factor, relating to the context of a theoretical cyberattack, relating to the intentionality of an attack, i.e., whether an attack is targeted or random (Rios Insua et al., 2019), something explicitly recommended in the ENISA (2018) report. This factor relates to a literature on cyber-risk communication (Blythe et al., 2011) which shows that users are more likely to be persuaded by messages which describe their particular vulnerability to an attack, but also relates to a literature on the efficacy of targeted risk communication in order to nudge cybersecurity behaviour (Egelman and Peer, 2015). Our hypothesis is that risk framed in terms of an intentional, targeted attack will be more likely to result in the adoption of premium cyberinsurance.

2.2.4 Dependent measures: The value of behavioural data

The majority of existing studies into cybersecurity behaviours have relied upon self-report measures rather than measuring actual behaviour (ENISA, 2018). Unfortunately, self-reporting does not always correlate with actual behaviour (Wash et al., 2017) and indeed has become something of a thorny problem for large scale survey studies of ISP (Lebek et al., 2014). Therefore, in this study we applied an economic experimental-behavioural approach to provide a scientific method to study how individuals interact in controlled settings and the collection of behavioural data. We thus addressed some of the concerns identified by Botzen and van den Bergh (2012)[pp. 152] who noted that "measuring risk attitudes and risk perception at the individual level and estimating their influence on insurance demand, [. . .] is rarely possible in actual insurance decisions and has hardly been addressed in empirical work". With this in mind, direct behavioural data, in combination with relevant attitude scales, were used as outcome variables in our test of a predictive model of cyberinsurance uptake. A similar approach to that applied by Mol et al. (2018) to investigate factors underlying uptake of insurance by home-owners in flood-risk areas. However, to the best of our knowledge this is the first study to apply an experimental, behavioural economics approach to understand decision-making in relation to cyberinsurance.

2.3 Materials and Methods

2.3.1 Sample

Participants (N = 4,800) were recruited across four EU countries (Germany, Spain, Poland and UK) in June 2018. Most participants (91.6%) were educated to high school level or above. Distribution by age and gender reflects Eurostat's data from the 2017 survey on ICT[1] that was used to create the quota.

	Germany	**Spain**	**Poland**	**UK**
Male	617	600	552	595
	51.4%	50.0%	46.0%	49.6%
Female	583	600	648	605
	48.6%	50.0%	54.0%	50.4%
18-34 years	932	842	713	844
	77.7%	70.2%	59.4%	70.3%
35-74 years	268	358	487	356
	22.3%	29.8%	40.6%	29.7%
Total	1,200	1,200	1,200	1,200
	100.0%	100.0%	100.0%	100.0%

Table 2.3.1: Distribution of the participants by gender, age and country.

[1] Data given in this domain are collected annually by the National Statistical Institutes and are based on Eurostat's annual model questionnaires on Information and Communication Technologies (ICT) usage in households and by individuals. `https://ec.europa.eu/eurostat/cache/metadata/en/isoc_i_esms.htm`

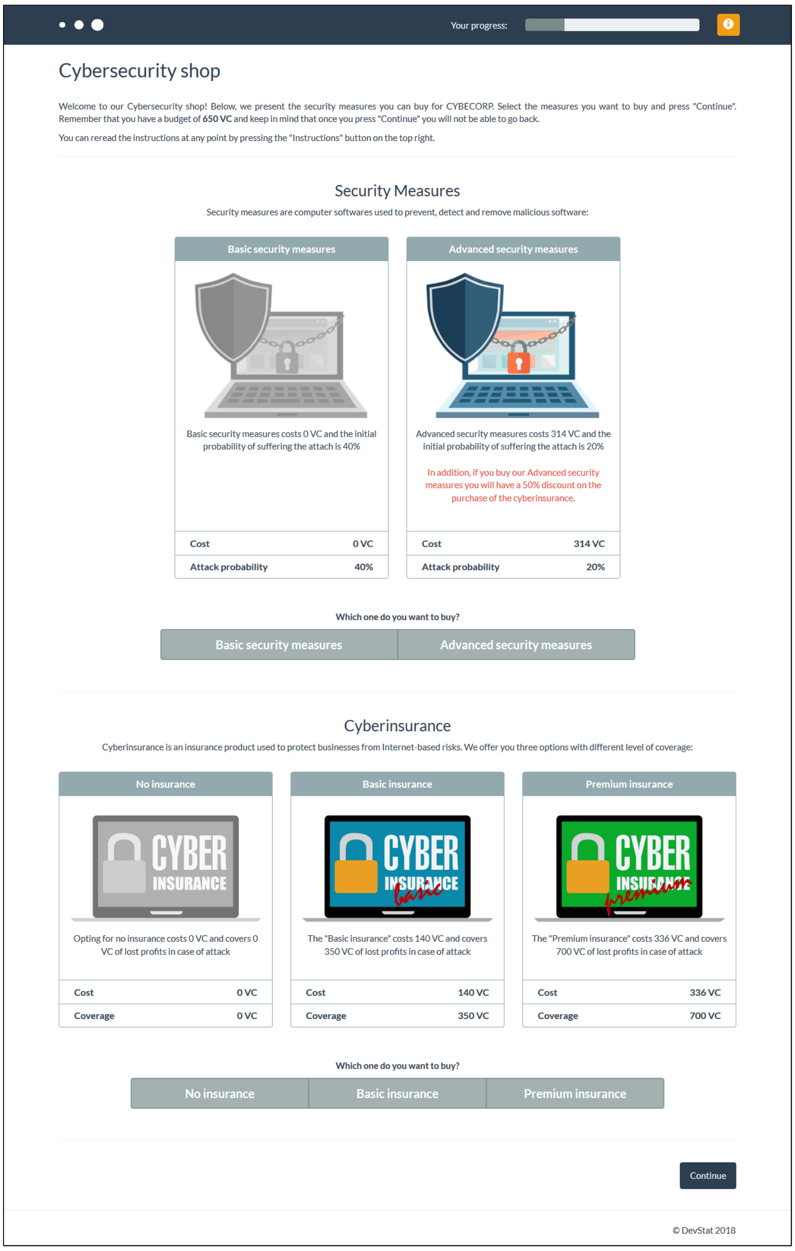
2.3.2 Experimental Design and Procedure

Each participant was informed that she or he may be at risk of a possible cyberattack affecting the value of their commercial data and therefore to the variable payment to be received at the end of the experiment. They were provided with an initial endowment in virtual coins (which could be exchanged for Euros at the end of the experiment). Participants were asked to make two decisions (i) whether to purchase security measures (basic or advanced) and (ii) what level of cyberinsurance to adopt (none, basic or premium). Participants visited the online "shop" (see Figure 2.3.1) to make their cybersecurity purchases. Note that they were given explicit information about the way that purchases were likely to either (i) reduce the likelihood of an attack (when buying security measures) or (ii) provide financial recompense (in the case of cyberinsurance).

Participants were then asked to complete an online task which involved registering for a conference. During the task, their possibility of suffering a cybersecurity attack was dependent upon the security of their online behaviour (i.e., the strength of their chosen password, whether they logged out after the task, whether they read the website terms and conditions, and whether they disclosed non-compulsory private information during the transaction).

At the end of the experiment, each participant received a variable payoff which depended upon their purchase decisions and whether a cyberattack did, or did not, occur.

Two experimental manipulations were applied: (i) The intentionality of the



Cybersecurity shop

Welcome to our Cybersecurity shop! Below, we present the security measures you can buy for CYBECORP. Select the measures you want to buy and press "Continue". Remember that you have a budget of **650 VC** and keep in mind that once you press "Continue" you will not be able to go back.

You can reread the instructions at any point by pressing the "Instructions" button on the top right.

Security Measures

Security measures are computer softwares used to prevent, detect and remove malicious software:

Basic security measures

Basic security measures costs 0 VC and the initial probability of suffering the attach is 40%

Cost	0 VC
Attack probability	40%

Advanced security measures

Advanced security measures costs 314 VC and the initial probability of suffering the attach is 20%

In addition, if you buy our Advanced security measures you will have a 50% discount on the purchase of the cyberinsurance.

Cost	314 VC
Attack probability	20%

Which one do you want to buy?

Cyberinsurance

Cyberinsurance is an insurance product used to protect businesses from Internet-based risks. We offer you three options with different level of coverage:

No insurance

Opting for no insurance costs 0 VC and covers 0 VC of lost profits in case of attack

Cost	0 VC
Coverage	0 VC

Basic insurance

The "Basic insurance" costs 140 VC and covers 350 VC of lost profits in case of attack

Cost	140 VC
Coverage	350 VC

Premium insurance

The "Premium insurance" costs 336 VC and covers 700 VC of lost profits in case of attack

Cost	336 VC
Coverage	700 VC

Which one do you want to buy?

© DevStat 2018

Figure 2.3.1: Mockup of the Online Shop.

attack and (ii) the pricing strategy applied to the protection and insurance products. The intentionality of the cyberattack (C), had two levels: intentional attack (participants are informed that an attacker may specifically target their company) and random attack (participants are informed that there is a virus in the Internet that may randomly affect any user).

Pricing strategy (P) had six levels obtained from the combination of three different prices (medium, asymmetric, and high) and two different relations between the prices of protection measures and insurance policies (dependent price - i.e., the price of the insurance policy decreases if advanced protection measures are chosen, and independent price, i.e., the price of the insurance policy remains the same regardless of whether the participant opted for none, basic or advanced security measures). This is shown in Figure 2.3.2.

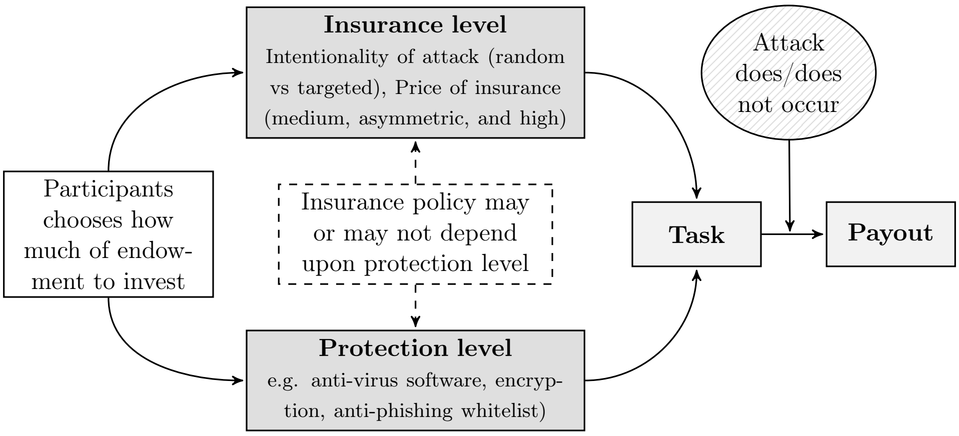

Insurance level
Intentionality of attack (random vs targeted), Price of insurance (medium, asymmetric, and high)
Attack does/does not occur
Participants chooses how much of endowment to invest
Insurance policy may or may not depend upon protection level
Task
Payout
Protection level
e.g. anti-virus software, encryption, anti-phishing whitelist)

Figure 2.3.2: Structure of Participant Task.

Participants were randomly allocated to each of the 12 experimental conditions. At the end of the experiment, participants received their pay-out and completed a short questionnaire of psychological measures.

2.3.3 Measures

The experiment included numerous behavioural measures (i.e., observation of participants' actual behaviour) and psychological measures (obtained through psychometric scales).

Behavioural Measures

Three behavioural measures were obtained from the purchasing decisions of the participant and their online behaviour during the task: (i) *Security measure adoption (basic or advanced)* and (ii) *insurance adoption (none, basic or premium).* The third measure, (iii) *risky online behaviour*, is calculated as a continuous variable between 0 (safest behaviour) and 1 (riskiest behaviour) as a linear combination of the proxy security variables included in the experiment: *security level of chosen password*, *disclosure of non-compulsory private information*, *viewing the terms and conditions* and *logging out* after completing the registration.

Psychological Measures

Following completion of the experiment, participants were presented with an online questionnaire measuring factors relating to PMT: *Perceived severity*, *perceived vulnerability*, *response efficacy*, *response cost* and *self-efficacy.* In addition to the PMT items, two other measures were included to fit with the Theory of Planned Behaviour: *Attitudes towards insurance* and *subjective norms.* The measure for attitudes towards insurance was based upon Anderson and Agarwal (2010) measure of attitudes toward security-related behaviour, amended to apply specifically to insurance. While subjective norms were measured using the single item “People who are important to me think that I should have insurance”. All PMT and TPB items were scored on a 5-point scale from strongly disagree (1) to strongly agree (5). The final two measures related to risk propensity (an individual’s natural tendency to take risks) and intention to purchase insurance. Risk propensity was measured using the Risk Propensity Scale (RPS) (Meertens and Lion, 2008). This 7-item scale has been used to measure risk propensity in relation to online behaviour (Branley and Covey, 2017) and requires considerably less space than the other commonly used, but lengthy, Domain-Specific Risk-Taking scale (Blais and Weber, 2006). The specific items used are shown in Table 2.3.2.

Items
Perceived Severity a. If my online data/accounts were hacked, it would be severe
Perceived Vulnerability a. My online data/accounts are at risk of being compromised b. It is likely that my online data/accounts will be breached c. It is possible that my online data/accounts will be compromised
Response Efficacy a. Insurance is an effective method to protect against loss b. Insurers can be trusted to pay out in the event of a claim
Self-efficacy/Perceived Behavioural Control a. I feel comfortable taking measures to secure my own computer(s) b. I feel comfortable taking security measures to limit the threat to other people and the Internet in general

Continued on next page

Table 2.3.2: Instrument items.

Items
c. Taking the necessary security measures is entirely under my control
d. I have the resources and the knowledge to take the necessary security measures
e. Taking the necessary security measures is easy
Response Cost & Rewards
a. Insurance is financially costly for me
b. Setting up insurance would require too much from me
c. Insurance is burdensome for me
d. Insurance is time consuming for me
e. Insurance is not worth it
f. Claiming on insurance could harm a business/organisations reputation
Attitudes
a. Insurance is a good idea
b. Insurance is important
c. I like the idea of taking out insurance to protect me
Subjective Norms
a. People who are important to me think that I should have insurance
Risk propensity
a. Safety first
b. I do not take risks with my health
c. I prefer to avoid risks
d. I take risks regularly
e. I really dislike knowing what is going to happen
f. I usually view risks as a challenge
g. I view myself as a.... [risk avoider vs. risk seeker]
Intention
a. I am likely to purchase cyberinsurance

Table 2.3.2: Instrument items (cont.).

2.3.4 Analysis

Preliminary analysis was carried out to confirm that the data was suitable for SEM. Correlation coefficients were used to examine relationships between all the variables. The structural model was tested using R (packages psych, semTools and lavaan). SEM is a method that combines and estimates two procedures simultaneously: Confirmatory Factor Analysis (CFA) and Path Analysis. CFA assesses the measurement component of the model, and path analysis assesses the relationship between latent variables (MacCallum and Austin, 2000). SEM allows us to include numerous endogenous variables and also to control for systematic and random measurement error (McDonald, 1990).

2.4 Results

2.4.1 Descriptive statistics

The majority of participants opted for a high level of protection in the experiment, i.e., 83.4% purchasing the advanced security measures, and 93% decided to purchase cyberinsurance (50.2% purchased premium insurance and 42.8% purchased basic insurance). Note, as only 7% of participants did not opt for any cyberinsurance, we collapsed the ‘no insurance and basic insurance’ categories and focused upon modelling the adoption of premium insurance in subsequent analyses.

2.4.2 Measurement model analysis

Using exploratory factor analysis, a test of reliability was conducted for each construct. During this analysis, items for attitudes and subjective norms loaded on the same factor and therefore were combined in the subsequent analyses. Some items of response cost (items a, e, f, Table 2.3.2) and risk propensity (item e) were eliminated to improve construct reliability. Means, standard deviations, and Cronbach's alpha scores for the remaining constructs are shown in Table 2.4.1. All Cronbach's alpha scores are greater than 0.7 indicating good reliability (Cortina, 1993; Nunnally, 1978).

Construct Items	**Mean**	**Std. dev.**	**Cronbach's α**
Perceived vulnerability	3.5	0.95	0.86
Response efficacy	3.5	1.00	0.74
Perceived behavioural control	3.7	0.78	0.84
Response cost	3.0	0.85	0.83
Attitudes & Subjective norms	3.8	0.84	0.87
Risk propensity	3.5	1.30	0.74

Table 2.4.1: Construct means, variances, and Cronbach's alpha scores.

2.4.3 Structural equation modelling

The SEM model for premium cyberinsurance adoption is shown in Figure 2.4.1. There are four significant pathways influencing premium insurance adoption: *Perceived response efficacy* and the TPB pathway (*social norms* and *attitudes through intention*) both positively influence premium insurance adoption. Whilst, *risk propensity* and *price difference* negatively influence adoption.

Adoption of premium insurance also shows a positive relationship with online behaviour. Attack intentionality does not appear to have any significant effect upon insurance adoption.

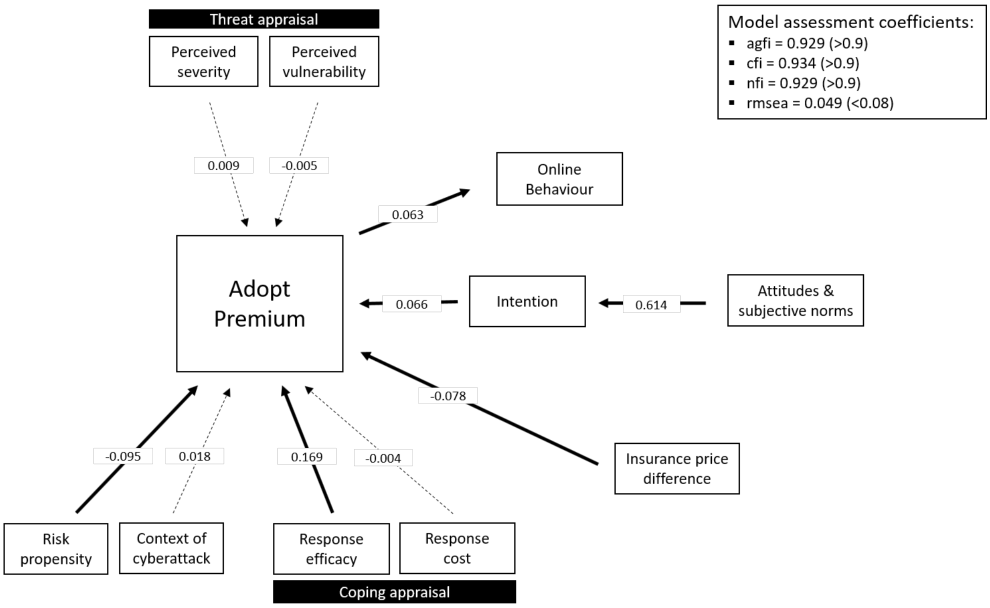


Figure 2.4.1: SEM Model of Cyberinsurance Adoption (standardised coefficients).

However, it is important to note that the decision to purchase cyberinsurance does not usually occur in isolation – it is likely to coincide with the decision to purchase additional security measures (e.g., antivirus, firewalls). This is further reinforced by the likelihood that insurance companies will require a minimum level of security before insurance will be granted. Therefore, a second model was created which includes the purchase of security measures - shown in Figure 2.4.2.

The second model shows a significant positive pathway that links the adoption of advanced security measures to the adoption of premium insurance. Thus, individuals who adopted advanced security measures were more likely to also adopt premium insurance. This is the strongest pathway in the model. The adoption of advanced security measures was also significantly positively related to security of online behaviour; those who adopted advanced security measures were also more likely to behave securely online. The pathway between insurance adoption and online behaviour, although positive, failed to reach significance once adoption of security measures was introduced into the model.

Response efficacy (part of PMT coping appraisal) and the TPB factors (attitudes & norms) were positively related to adoption of premium insurance; those who perceived insurance to be more effective, and those who had positive attitudes and positive subjective norms, were more likely to adopt premium cyberinsurance. Perceived self-efficacy and perceived threat severity (part of PMT threat appraisal) both positively fed into the adoption of advanced security measures rather than adoption of premium insurance directly. Those who had higher perceptions of their ability to put cybersecurity measures into place, and those who perceived the threat of the cyberattack as more severe, were more likely to adopt advanced security measures (which as aforementioned then subsequently fed into premium insurance adoption).

Risk propensity was negatively related to both adoption of security measures and adoption of insurance, i.e., a risk-seeking individual was less likely to adopt advanced security measures and premium insurance.

As found in the first model, context of the cyberattack (i.e., attack intention-

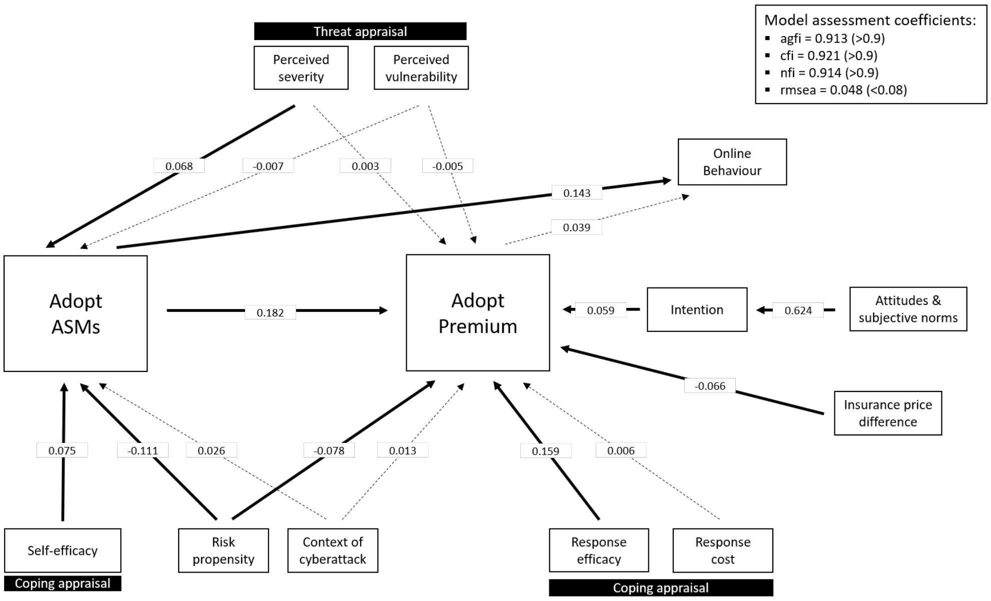


Figure 2.4.2: SEM Model including Security Measures Adoption (standardised coefficients).

ality: targeted or random) had no significant effect upon insurance adoption, nor upon purchase of security measures.

2.5 Discussion

The current study used SEM to test a model of cyberinsurance adoption and address a significant gap in the existing literature. Despite a fast-growing interest in, and industry around, cybersecurity — there is an overwhelming lack of knowledge in relation to understanding the mechanisms behind cybersecurity decision-making. The results support the model as a good fit to the data therefore providing important knowledge of the factors influencing cybersecurity decisions — including uptake of security measures and insurance.

Our findings highlight that cyberinsurance adoption is only one factor in a larger, more complex security ecosystem. The decision to adopt premium cyberinsurance was directly influenced by the adoption of other advanced security measures i.e., those who invested in advanced security measures were more likely to also purchase premium insurance (Interestingly, adoption of advanced security measures was also predictive of more secure online behaviour. Suggesting that concerns over moral hazard — i.e., that an individual may increase their exposure to risk if they do not bear the full costs of that risk — may be unfounded).

In turn, the adoption of advanced security measures was significantly influenced by perceived severity of an attack, risk propensity and perceived self-efficacy (i.e., confidence in one's own ability to implement security measures). Taken together, the findings suggest that, in order to adequately target insurance uptake, it is important to account for the wider portfolio of security measures available to an individual or organisation. It is vital that this is taken into account by any future research. This conclusion is entirely consistent with the ENISA (2018) recommendations for organisational contexts to be more carefully considered.

Outside of security measure adoption, uptake of premium insurance was negatively influenced by risk propensity and price difference [between basic and premium policies], and positively influenced by positive attitudes and social norms around insurance and perceived response efficacy. These results are relatively unsurprising: individuals with a higher propensity for risk were less likely to adopt premium insurance, and higher premium policy pricing (comparative to basic insurance) led to lower premium insurance uptake. In keeping with PMT, those who perceived insurance to be an effective method to protect against loss were more likely to adopt premium insurance; and in line with TPB, positive attitudes towards insurance (e.g., perceiving insurance as a good thing) and strong social norms (i.e., perceptions that other people think they should have cyberinsurance) were linked with intention to adopt premium insurance.

Perceived vulnerability (one element of PMT threat appraisal) was not a significant predictor of advanced security measure or premium insurance adoption. Nor was the intentionality of the attack a significant factor. These are troubling findings for the PMT model as applied to cyberthreat and beg the question as to why perceived severity of an attack (the other element making up PMT threat appraisal) may be influential, but perceived vulnerability less so. There may be lessons here for the design of studies in this space. Consider: severity is a relatively meaningful construct independent of the experimental set up as it relies upon knowledge of cyberthreats in the wild. However, vulnerability judgements require an assessment of the likelihood of a particular organisation will succumb to a threat. In an experimental set up, there is little offered in the way of organisational context to help make this assessment. True, we sought to manipulate context by describing the threat as either targeted or not, but we provided no background information as to the resilience of the organisation at the start of the study. Simply put, how could participants determine vulnerability? This is worth considering in future studies and may help to account for the ENISA (2018) report observation that threat information tends to be relatively ineffective in driving behaviour.

In this study we identified the key factors underlying decision-making around cybersecurity. The model presented here could be used to guide future interventions aimed at increasing cyberinsurance (and cybersecurity) uptake. Our findings show that it is vital to consider the larger cybersecurity ecosystem, rather than attempting to focus solely upon insurance adoption in isolation. This focus upon the wider ecosystem could help to improve societal cybersecurity, although we note that context rich studies of cybersecurity remain limited.

References

Ajzen, I. et al. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, 50(2):179–211.

Anderson, C. L. and Agarwal, R. (2010). Practicing safe computing: a multimedia empirical examination of home computer user security behavioral intentions. *MIS quarterly*, 34(3):613–643.

Baer, W. S. and Parkinson, A. (2007). Cyberinsurance in it security management. *IEEE Security & Privacy*, 5(3):50–56.

Barsky, R. B., Juster, F. T., Kimball, M. S., and Shapiro, M. D. (1997). Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study. *The Quarterly Journal of Economics*, 112(2):537–579.

Beck, K. H. (1984). The effects of risk probability, outcome severity, efficacy of protection and access to protection on decision making: A further test of protection motivation theory. *Social Behavior and Personality: an international journal*, 12(2):121–125.

Blais, A.-R. and Weber, E. U. (2006). A domain-specific risk-taking (dospert) scale for adult populations. *Judgment and Decision making*, 1(1).

Blythe, J., Camp, J., and Garg, V. (2011). Targeted risk communication for computer security. In *Proceedings of the 16th international conference on Intelligent user interfaces*, pages 295–298.

Botzen, W. W. and van den Bergh, J. C. (2012). Risk attitudes to low-probability climate change risks: Wtp for flood insurance. *Journal of Economic Behavior & Organization*, 82(1):151–166.

Brahmana, R., Brahmana, R. K., and Memarista, G. (2018). Planned behaviour in purchasing health insurance. *The South East Asian Journal of Management*.

Branley, D. B. and Covey, J. (2017). Is exposure to online content depicting risky behavior related to viewers' own risky behavior offline? *Computers in Human Behavior*, 75:283–287.

Cortina, J. M. (1993). What is coefficient alpha? an examination of theory and applications. *Journal of applied psychology*, 78(1):98.

Dittrich, R., Wreford, A., Butler, A., and Moran, D. (2016). The impact of flood action groups on the uptake of flood management measures. *Climatic Change*, 138(3-4):471–489.

Egelman, S. and Peer, E. (2015). The myth of the average user: Improving privacy and security systems through individualization. In *Proceedings of the 2015 New Security Paradigms Workshop*, pages 16–28.

ENISA (2018). Cybersecurity culture guidelines: Behavioural aspects of cybersecurity.

Grahn, T. and Jaldell, H. (2019). Households (un) willingness to perform private flood risk reduction–results from a swedish survey. *Safety science*, 116:127–136.

Ifinedo, P. (2012). Understanding information systems security policy compliance: An integration of the theory of planned behavior and the protection motivation theory. *Computers & Security*, 31(1):83–95.

Khalili, M. M., Liu, M., and Romanosky, S. (2019). Embracing and controlling risk dependency in cyber-insurance policy underwriting. *Journal of Cybersecurity*, 5(1):tyz010.

Kuru, D. and Bayraktar, S. (2017). The effect of cyber-risk insurance to social welfare. *Journal of Financial Crime*, 24(2):329–346.

Lammers, J., Warmerdam, S., and Ecorys, R. (2010). Adverse selection in voluntary micro health insurance in nigeria. *AIDS Research Series*, 6.

Lebek, B., Uffen, J., Neumann, M., Hohler, B., and Breitner, M. H. (2014). Information security awareness and behavior: a theory-based literature review. *Management Research Review*.

Low, P. (2017). Insuring against cyber-attacks. *Computer Fraud & Security*, 2017(4):18–20.

MacCallum, R. C. and Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual review of psychology*, 51(1):201–226.

McDonald, R. P. (1990). Structural equations with latent variables. *Journal of the American Statistical Association*, 85(412):1175–1177.

Meertens, R. M. and Lion, R. (2008). Measuring an individual's tendency to take risks: The risk propensity scale 1. *Journal of Applied Social Psychology*, 38(6):1506–1520.

Mol, J. M., Botzen, W. W., and Blasch, J. E. (2018). Behavioral motivations for self-insurance under different disaster risk insurance schemes. *Journal of Economic Behavior & Organization*.

Nunnally, J. C. (1978). Psychometric theory 2nd ed.

Pal, R., Golubchik, L., Psounis, K., and Hui, P. (2017). Security pricing as enabler of cyber-insurance a first look at differentiated pricing markets. *IEEE Transactions on Dependable and Secure Computing*, 16(2):358–372.

Petrolia, D. R., Landry, C. E., and Coble, K. H. (2013). Risk preferences, risk perceptions, and flood insurance. *Land Economics*, 89(2):227–245.

Rios Insua, D., Couce-Vieira, A., Rubio, J. A., Pieters, W., Labunets, K., and G. Rasines, D. (2019). An adversarial risk analysis framework for cybersecurity. *Risk Analysis*.

Sommestad, T., Hallberg, J., Lundholm, K., and Bengtsson, J. (2014). Variables influencing information security policy compliance. *Information Management & Computer Security*.

Sommestad, T., Karlzén, H., and Hallberg, J. (2015). A meta-analysis of studies on protection motivation theory and information security behaviour. *International Journal of Information Security and Privacy (IJISP)*, 9(1):26–46.

Tsai, H.-y. S., Jiang, M., Alhabash, S., LaRose, R., Rifon, N. J., and Cotten, S. R. (2016). Understanding online safety behaviors: A protection motivation theory perspective. *Computers & Security*, 59:138–150.

van Bavel, R., Rodríguez-Priego, N., Vila, J., and Briggs, P. (2019). Using protection motivation theory in the design of nudges to improve online security behavior. *International Journal of Human-Computer Studies*, 123:29–39.

Wash, R., Rader, E., and Fennell, C. (2017). Can people self-report security accurately? agreement between self-report and behavioral measures. In *Proceedings of the 2017 CHI conference on human factors in computing systems*, pages 2228–2232.

Chapter 3

Behavioural recomposition for adversarial belief assessment

Abstract

In domains such as homeland security, cybersecurity and competitive marketing it is frequently the case that analysts need to forecast adversarial actions that impact our decisions. Standard structured expert judgement elicitation techniques fall short as they do not take into account intentionality. A decomposition technique based on adversarial risk analysis followed by recomposition rules based on discrete choice models enable such process facilitating such assessments.

A conceptual map of this chapter is presented in Figure 8.

3.1 Introduction

Chapter 1 shows that defenders tend to adopt a lower level of cyberprotection in front of intentional cyberattacks than under the menace of random unintentional ones, even knowing that the chances and the impact of suffering both types of attacks are the same. This finding suggests that the intentionality of an attack may influence agents' beliefs and vulnerability self-perception and resulting in agents assuming that intentional cyberattacks will not happen to them. This results motivates a deeper analysis of risk belief formation in adversarial situations.

Since beliefs on agent's cybervulnerability are not observable, a reliable elicitation method will be required for their analysis. In this context, Structured Expert Judgementy (SEJ) elicitation becomes a major ingredient within decision analysis in cybersecurity and in any other thematic domain.(Clemen and Reilly, 2013). A significant feature in the practice of this discipline, as acknowledged in Raiffa (1968)'s seminal book, is the emphasis in decomposing complex problems into smaller pieces that are easier to understand and recombining the piecewise solutions to tackle the global problem.

Under a standard rational choice approach, when applied to decision analysis, this methodology seeks to solve complex decision making problems through maximum expected utility. In doing so, one avoids direct comparison of alternatives, which may be cognitively intricate and prone to bias, specially in presence of uncertainty and multiple objectives. Instead, we structure the problem by identifying alternatives, uncertainties and objectives; assess the decision maker's beliefs and preferences; and, then, find the optimal alternative. The value of such decomposition is assessed in Watson and Brown (1978).

Preference assessment also uses decomposition. It will usually be difficult to compare consequences of alternatives, specially in presence of multiple conflicting attributes. A typical approach is to search for a decomposable functional form for a utility function (often additive, linear or multilinear, e.g. González-Ortega et al., 2018), and then assess the component utilities and weights to later recompose the global utility function whose expected value must be maximised. Ravinder and Kleinmuntz (1991) and Ravinder (1992) provide theory showing the advantages of undertaking such decompositions in preference elicitation.

Finally, belief assessment also benefits from decomposition through the *extending the conversation* argument. Tetlock and Gardner (2015) call it *Fermitisation* and consider it as a key strategy for the success of their super-forecasters. Rather than directly assessing the probability of an outcome, one finds a conditioning partition and assesses the probabilities of the outcome given the conditioning events. From these, and the probabilities of these events, the law of total probability enables calculation of the unconditional probability of the outcome. Andradottir and Bier (1997, 1998) provide a methodological framework to validate this approach, empirically evaluated in MacGregor (e.g. 2001).

Decompositions thus uncover the complexity underlying the formation and direct elicitation of beliefs when facing intentional menaces, eliminating the burden on experts to perform sophisticated modelling in their heads. This simplifies complex cognitive tasks and mitigates their reliance on heuristics that can introduce bias, promoting that they actually analyse the relevant problem (Montibeller and von Winterfeldt, 2015). Decompositions typically entail more assessments, albeit simpler and more meaningful, leading to improved judgements and decisions.

Here we focus on developing and validating experimentally a decomposition strategy to support SEJ when forecasting adversarial actions based on Adversarial Risk Analysis (ARA) (Banks et al., 2015). There are two main uses for this. First, in line with Kadane and Larkey (1982) and Raiffa (1982), we could use decision analysis to support a decision maker in dealing with game theoretic problems and this leads to trying to forecast adversarial actions. Second, in cybersecurity settings, as well as in many other contexts such as security, counterterrorism, or intelligence, experts will face problems in which they need to deal with probabilities referring to actions potentially carried out by opponents. As an example, an important percentage of the questions posed to experts in Chen et al. (2016) refer to adversaries (e.g. *Will Raja Pervez Ashraf resign or otherwise vacate the office of Prime Minister of Pakistan before 1 April 2013?* or *Will the Palestinian group Islamic Jihad significantly violate its cease-fire with Israel before 30 September 2012?*). We could think of using standard SEJ tools, as in Dias et al. (2018) and Hanea et al. (2021), to deal with such problems. However, as cogently argued in Keeney (2007), knowledge about the adversaries beliefs and preferences may not be that precise as it would require them to reveal their judgements, which is not feasible in cybersecurity domains. Alternatively, we study here whether ARA decompositions serve better for such purpose and determine the right questions to ask.

Insua et al. (2020) argues theoretically that ARA may be used as a decomposition-recomposition strategy for adversarial forecasting. Here we assess empirically this research hypothesis from a behavioural perspective, using appropriate economic experiments. To this end, we integrate ARA with random utility models (Thurstone, 1927; Marschak, 1959; McFadden, 1973) to provide a behavioral perspective on ARA methods, adopting an asymmetric prescriptive-descriptive approach as in Raiffa (1982). Indeed, we use a prescriptive view for the defender assessing and decomposing her uncertain view of the adversary's preferences and beliefs, later recomposing them with the aid of descriptive random utility models to obtain improved adversarial forecasts. In such a way we provide a novel behavioural perspective of ARA, which we named as *Behavioural Beliefs Recomposition.*

After sketching the ARA approach to decomposition (Section 3.2) and briefly recalling the theoretical arguments to justify it together with its integration with random utility models, we present the experiments undertaken in Section 3.3 and analyse the results drawing conclusions in Section 3.4. We end up with a discussion in Section 3.5.

3.2 ARA as a SEJ decomposition method

We start by outlining the role of ARA as a SEJ decomposition method according to the two proposed uses and then describe its integration with random utility models to provide a normative decomposition-behavioral recomposition to forecast adversarial actions.

3.2.1 Games from a decision analytic perspective

ARA was originally introduced to deal with game theoretic problems studied from a decision analytic perspective (Banks et al., 2015). Games are formulated in a Bayesian manner, as in Kadane and Larkey (1982) and Raiffa (1982), and operationalised through the provision of procedures to forecast the actions of the adversary with the aim of mitigating common knowledge assumptions standard in game theory.

To conceptualise the discussion, consider an example. Imagine a company D (she, defender) which deploys cybersecurity controls $d \in \mathcal{D}$. Then, having observed d, a cybercriminal A (he, attacker) decides whether to launch a cyber attack $a \in \mathcal{A}$. The outcome s would be a random variable S whose distribution depends upon both d and a (the controls deployed and the attack launched). To solve her problem through decision analysis, the company requires $p_D(s \,|\, d, a)$, which reflects her beliefs on s given both agents' actions, and her utility function $u_D(d, s)$, modelling her preferences and risk attitudes over the consequences. Besides, she needs the distribution $p_D(a \,|\, d)$, her assessment of the probability that A will choose action a after having observed her choice d. Once D has available these judgements, she computes the expected utility of control d through $\psi_D(d) = \int \left[\int u_D(d, s)\, p_D(s \,|\, d, a) \, \mathrm{d}s \right] p_D(a \,|\, d) \, \mathrm{d}a \,$, and seeks for the optimal $d^* = \operatorname{argmax}_{d \in \mathcal{D}} \, \psi_D(d)$. This is a standard decision analysis problem under expected utility, except for the elicitation of $p_D(a \,|\, d)$ which entails strategic aspects.

D could try to assess $p_D(a \,|\, d)$ from a standard belief elicitation perspective, as in Cooke et al. (1991) or O'Hagan et al. (2006), but ARA usefully suggests a decomposition approach to such assessment that requires her to analyse the problem from A's perspective. Assume, for the moment, that the adversary is an expected utility maximiser. The Defender puts herself in A's shoes, using all the information she can to obtain about A's probabilities $p_A(s \mid d, a)$ and utilities $u_A(d, s)$. Instead of using point estimates for them to find A's optimal response for a given d, her uncertainty about A's decision would derive from her uncertainty about (p_A, u_A), through a distribution F on the space of probabilities and utilities. This weakens the standard, but unrealistic, common knowledge assumptions in game theoretic approaches (Hargreaves-Heap and Varoufakis, 2004), according to which the agents share information about their beliefs and preferences. In this case, not having common knowledge means that the defender has uncertainty about (p_A, u_A) modelled through F. This induces a distribution over A's expected utility which, for each d and a, is $\Psi_A(d, a) = \int U_A(a, s)\, P_A(s \mid d, a)\, \mathrm{d}s$, where $(P_A, U_A) \sim F$. Then, D finds the required

$$\widehat{p_D}(a \mid d) = \mathcal{P}_F \left[a = \mathrm{argmax}_{x \in \mathcal{A}}\ \Psi_A(d, x) \right]$$

in the discrete case (and, analogously, in the continuous one).

The standard ARA approach assumes rationality for both the Defender and the Attacker, i.e. A and D are expected utility maximisers. We could argue that, since we have not the attacker A available, unlike the Defender D, it is doubtful that he uses expected utility to aggregate his judgements. Therefore, we may consider that he actually uses

$$\Psi_A(d, a) = V(U_A(a, s), P_A(s \mid d, a)),$$

for some general functional V representing A's decision rule, from the viewpoint of the Defender. This conforms to the asymmetric prescriptive-descriptive approach to games in Raiffa (1982) and the opponent modeling processes described in Rios Insua et al. (2016).

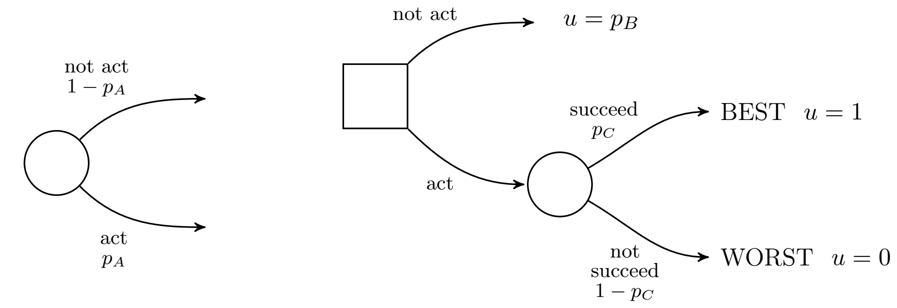
3.2.2 Predicting adversarial actions

Figure 3.2.1 represents the context in which we aim at predicting the action of the adversary and serves also as an sketch of the experimental design described in Section 3.3. The left panel represents the uncertainty of interest: we aim at forecasting whether an adversary will act, which will happen with probability p_A. This amounts to its direct assessment. The right panel reflects the decision problem faced by the adversary. If he does not act, the status quo remains and he attains utility u, which we designate p_B; if he acts and succeeds, which happens with probability p_C, he gets the best result with utility 1; however, if he fails, he would get the worst result with utility 0.

Assume for the moment that the adversary is an expected utility maximiser. Should we know p_C and u, as the expected utility of both actions (act and not act) are, respectively, u and p_C, we would predict that the adversary will act if $u < p_C$ and will not act, otherwise. However, since there is uncertainty about such elements, which we model through random distributions U and

P_C, we have uncertainty about the adversary's decision and we would estimate $\widehat{p_A} = W(U, P_C) = Pr(V(U, 1) < V((1, 0), (P_C, 1 - P_C))$.

However, we could argue again that as the adversary is not actually available, the defender can consider that the attacker may apply other deterministic or random rule beyond expected utility maximisation to select an action. In a general way, attackers decision rule can be modeled as $\widehat{p_A} = W(U, P_C)$, where W is a generic functional, potentially different from that in the above paragraph and coming from maximisation of expected utility. Some examples of behavioural-based specific forms for W are presented in Section 2.3.



(a) Direct assessment (b) Assessment through ARA decomposition

Figure 3.2.1: Two views on adversarial forecasts.

3.2.3 Recomposing adversarial beliefs with discrete choice models

The introduction of V and W in Sections 2.1 and 2.2 allows us to consider alternative behavioural choice models to represent D's beliefs about the decision rules that her adversary is using. As a consequence, the generalisation (3.2.1) of $\Psi_A(d, a)$ allows for considering alternative decision rules, coming for instance from potential cognitive biases or heuristic decisions involved in the decision procedure of A, from D's perspective. Such behavioural insights in $\Psi_A(d, a)$ would be expected to improve the accuracy of the recomposed estimations of the adversary's probabilities to take each of feasible actions.

Beyond expected utility maximisation, we explore two alternatives to D's perception of A choice rules. Both stem from introducing a random element within utility maximisation, known as *random utility* modelling. This concept was proposed by Marschak (1959) and Block et al. (1959) and later developed in McFadden (1973). In these models, the decision maker's utility is subject to random shocks. In a behavioural economics setting, these shocks are typically interpreted as a result of the agents' cognitive biases, of heuristics applied for decision-making or of errors in the implementation of the decision process (Hess et al., 2018). Utility maximization together with empirically estimated distribution of shocks leads to a probabilistic criterion for a particular alternative to be selected over the others. Note that although both ARA and random utility modelling translate into a probabilistic rule for action selection, the underlying uncertainty is different in both approaches: in ARA models, uncertainty about the actions to be selected is a consequence of the agents beliefs on their adversaries probabilities and utility; meanwhile, in random utility models the utility of each agent is intrinsically random, this randomness becoming the source of uncertainty for the selected action.

The probability distribution on the action to be chosen by the adversary these models can be obtained from a constructive or an axiomatic approach. In the initial conceptualisations by Marschak (1959) and McFadden (1973), this distribution is obtained by construction. To this end, the model is specified by defining an a priori utility and randomness is introduced by making the utility random or by making behaviour a random function of a fixed utility (Busemeyer and Rieskamp, 2014). From this constructive approach, MacFadden obtains a functional form for the choice probability, known as Multinomial logit (MNL) model, given by

$$\widehat{p}_A^{MNL} = \frac{e^{u(i)}}{\sum_{j \in S} e^{u(j))}}, i \in S,$$

where $u(i)$ is the agent's utility for alternative i. The axiomatic treatment was undertaken by Luce (1959), who introduced a choice axiom (LCA) postulating how the probability of selecting an alternative from one set is related to the probability of selecting such alternative from a larger set. The Axiomatic Random Utility (ARU) approach leads to a probability of selecting the i-th alternative from a set S through

$$\widehat{p}_A^{ARU} = \frac{u(i)}{\sum_{j \in S} u(j)}, i \in S,$$

Under LCA, $\hat{p}_A^{ARU}$ verifies the principle of independence of irrelevant alternatives (Arrow, 1951), and, thus, when the latter is normative, it becomes a reasonable assumption. Detailed discussions on LCA validity are in Luce (1977) and Pleskac (2015).

The behavioural recomposition models to be analysed empirically here will integrate the axiomatic and multinomial random utility probabilistic choice rules $\widehat{p}_A^{ARU}$ and $\widehat{p}_A^{MNL}$ within ARA models. The accuracy of these behavioural recompositions will be compared and also tested with that of standard ARA considering maximization of expected utility. These comparison will provide valuable insights for developing effective SEJ methodologies.

3.3 Validating behavioural ARA for SEJ decomposition-recomposition: rationale

3.3.1 Experiment design

To validate the proposed SEJ decomposition-recomposition method, a group of participants was recruited to elicit their beliefs on uncertain adversarial events within the controlled setting of an economic experiment.

A challenge for the design was to build a user-friendly mechanism to help the participants reveal such beliefs in a reliable and consistent manner. To this end, the experiment focused on the analysis of three uncertain events involving strategic decision-making from three topics (politics, consumption products and sports), specifically: *Will Theresa May ask for elections in the next 30 days?*; *Will Coca Cola's CEO announce a new marihuana based drink in the next 30 days?*; *Will Rafael Nadal participate in a hard court competition in the next 30 days?*. The design paid attention to the wording of the questions and the framing of the answering mechanisms to be used for preference and belief disclosure. The experiment was piloted with 10 subjects, who also participated in individual face-to-face debriefing interviews to confirm their understanding of the procedure and identifying potential improvements.

For each of the uncertain events, the subjects were required to complete four experimental tasks (12 tasks in total). Such tasks were presented sequentially as follows, recall Figure 3.2.1:

- *Task 1.* Beliefs about a well-known person, the adversary in the terminology above, referred to as Decision Maker (DM), to launch a strategic action in a given period of time, denoted p_A. As showcased in Figure 3.A.2, the participant assesses directly the required probability in a scale from 0% to 100%, in steps of 10%. Effectively, we identify an interval of width 10% where the participant believes that such probability lies.
- *Task 2.* Beliefs about the maximum success probability under which the DM would launch the strategic action, from which we deduce p_B. As showcased in Figure 3.A.3, the protocol asks whether the DM will act or not for various probabilities of success from 0% to 100%, in steps of 10%. Again, we effectively identify an interval of width 10% in which the participant believes such probability lies. Through this task, the subject is actually revealing his beliefs about the DM's utility of not acting, since utilities for acting and succeeding or failing are normalised to 1 and 0, respectively.
- *Task 3.* Beliefs about the chances of the DM succeeding if the strategic action is launched, denoted p_C. As in Figure 3.A.4, the participant assesses directly the probability in a scale from 0% to 100%, in steps of 10%. Again, we effectively identify an interval of width 10% where the

participant believes such probability lies.

- *Task 4.* Repetition of *Task 1.* We ask again for the probability p_D of the DM launching the strategic action in the given period of time. This repetition allows for checking if a participant has changed her beliefs after the exercise of reflecting about and revealing p_B and p_C. It is implemented as the first one.

The core associated screens are displayed in Appendix 3.A.

3.3.2 Experiment implementation

The face-to-face experiment was carried out in an experimental laboratory during April 2019. It was run using an experimental software developed in PHP specifically for this project. It was tested before the sessions to guarantee its usability and understandability.

The experiment embraced four sessions, each with 24 participants, for a total of 96 subjects. At the beginning of each session, subjects were randomly located around semi-cubicles in a room. To mitigate presentation bias, the same facilitator led all sessions. He read the instructions and accompanied his speech with a slide projector to explain the decisions that the subjects would have to make through three examples, and the benefits that they would realise as a function of the performance of their forecasts. They then undertook the actual tasks. The economic experiment entailed a variable payment to each subject, depending on the probabilities assigned to a series of events and its actual realisation during the days following the experimental session. The median duration of the experiment was 30 minutes, leaving about 10 minutes per topic.

3.4 Validating behavioural ARA for SEJ decomposition-recomposition: results

Our analysis of the results covers three stages.

1. The first one is exploratory. We assess the coherence of the respondents in relation with the adherence of the adversary to the expected utility model, with negative conclusions. We also explore the feasibility of alternative behavioural recomposition methods from subjects' responses.
2. We next assess whether the enforced probability assessments improve the predictive capabilities via proper scoring rules after the events took place, with positive results.
3. Finally, we check whether, upon the reflection induced by tasks, there are substantial changes concerning the probability p_A as reflected by p_D, with negative results.

3.4.1 Checking and enforcing coherence

Our first analysis is exploratory and aims at reflecting upon the coherence of the participants' responses with respect to the adherence of the adversaries to the expected utility postulates.

Recalling our discussion in Section 3.2.2, we essentially relate the assessment of p_A with those of p_C and $u = p_B$. Note that if $u < p_c$, it should be $p_A = 1$. However, taking into account the actual uncertainty of the participant about the adversary's judgements u and p_C, we would expect at least that $p_A > 1/2$: if it is more likely that the expected utility of acting is bigger than that of not acting, it should actually be more likely that the adversary acts (and viceversa). Such cases will be called agreements.

Figure 3.4.1 explores agreements in our experiment under two views. The left panel represents the scatter plot for the 288 (3 × 96) responses of $(p_C - p_B, p_A)$[1]. Points in white areas represent agreements with the expected utility model; those in red areas suggest disagreements. Observe the large number of observations in the red area. The right panel shows the histograms and density plots of p_A, for the cases $p_C > p_B$ and $p_C < p_B$, showcasing again a non-negligible amount of disagreements, as both densities have support (0,1).

Table 3.4.1 summarises the results, essentially conveying 70.2% of agreements vs 29.9% of disagreements with the (adversary's) expected utility model.

	$p_C < p_B$	$p_C > p_B$
$p_A < 1/2$	41.0%	17.7%
$p_A > 1/2$	12.2%	29.2%

Table 3.4.1: Agreements and disagreements with adversary's expected utility.

This suggests that there is no full conformity with the expected utility of ad-

[1] Because of the discreteness of the responses there were coincidences and we have jittered the observations with noise (Cleveland, 1985).

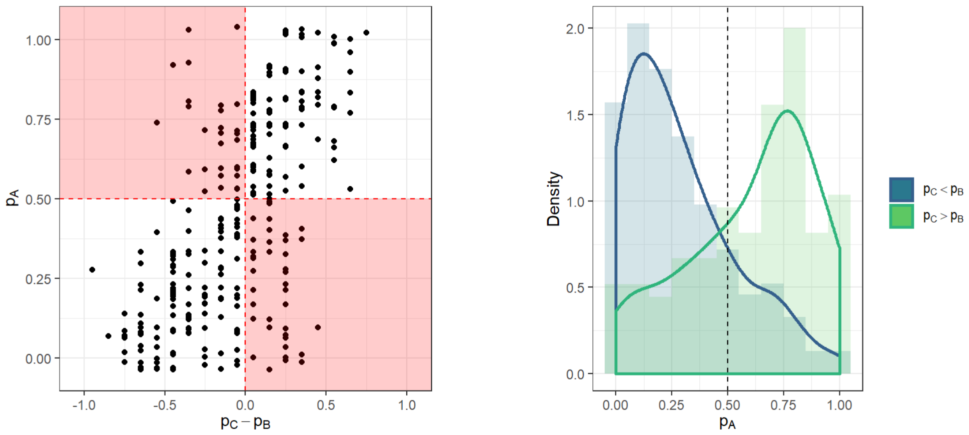


Figure 3.4.1: Exploring agreements with expected utility.

versary on behalf of the participants as discussed in Section 3.2.2. We finally display in Table 3.4.2 the percentage of participants that agreed with the adversary expected utility model in 0, 1, 2 or 3 of the proposed choices.

0	1	2	3
5.2%	15.6%	42.7%	36.5%

Table 3.4.2: Percentage of participants per agreement level.

Observe that nearly 64% of participants incur in at least one disagreement.

Given the high level of disagreement, in line with the suggestions in Section 3.2.3, we enforce coherent assessments integrating the use of three different forms for choice probability W, as discussed in Section 3.2.3.

- Under Expected Utility Maximisation (EUM), the probability distribution on selected actions $\hat{p}_A^{EUM}$ in Figure 3.2.1 is:

$$\widehat{p}_A^{EUM} = \begin{cases} 1 & \text{if } p_C > p_B \\ 1/2 & \text{if } p_C = p_B \\ 0 & \text{if } p_C < p_B \end{cases}$$

- Axiomatic Random Utility (ARU)

$$\widehat{p}_A^{ARU} = \frac{p_C}{p_C + p_B}$$

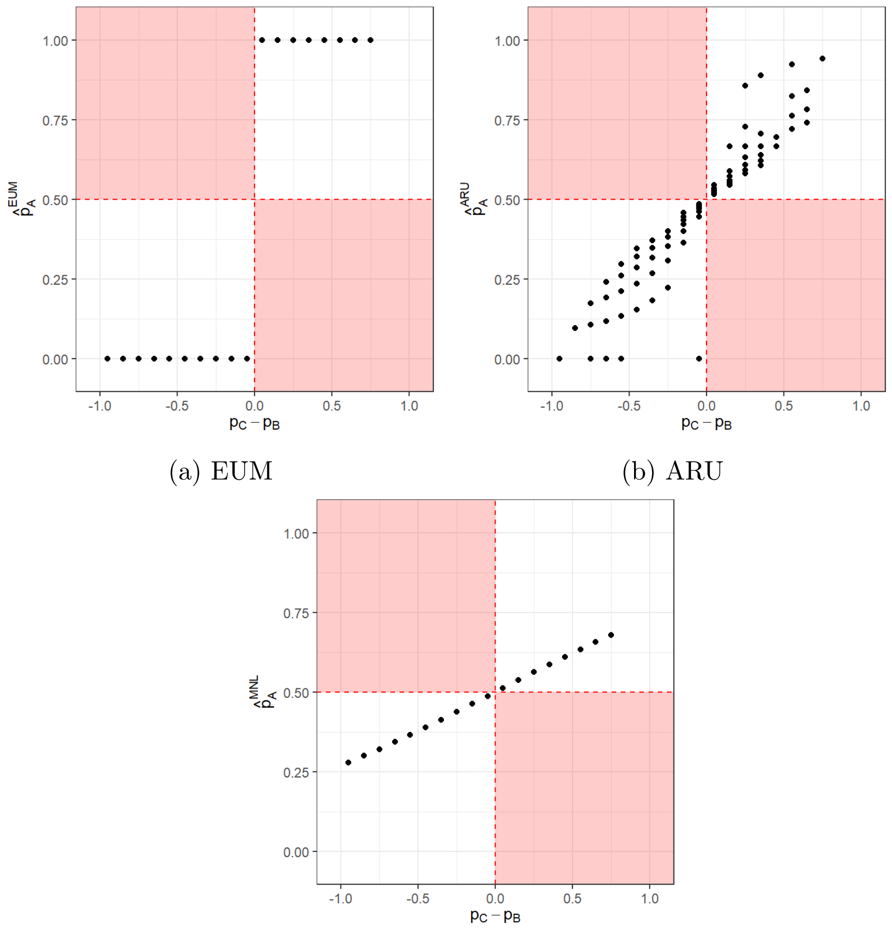
- Multinomial logit (MNL)

$$\widehat{p}_A^{MNL} = \frac{e^{p_C}}{e^{p_C} + e^{p_B}}$$

Figure 3.4.2 represents the scatterplots of $p_C - p_B$ and the three recompostions of the choice probability, which, as expected, show the enforced coherence through the basic constituents there being no observations in the disagreement areas.

3.4.2 Checking the predictive capacity of discrete choice ARA based recompositions

Given the above, we assess the previous recompositions in terms of predictive capacity. For this:



(a) EUM

(b) ARU

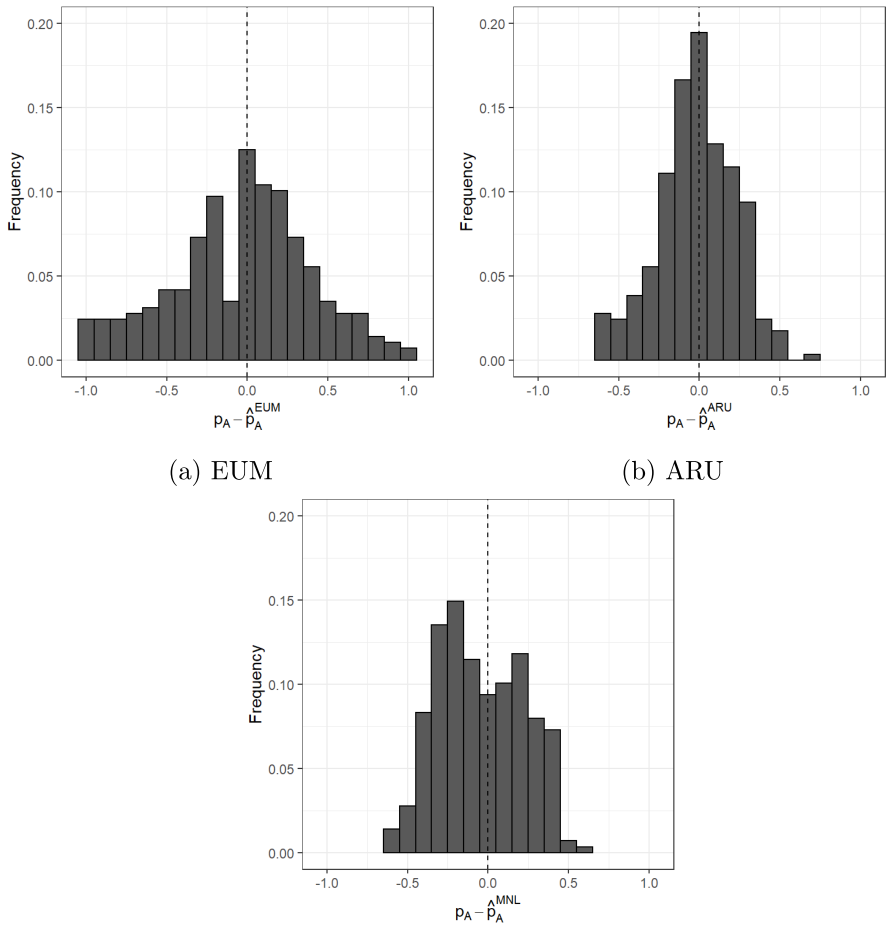
(c) MNL

Figure 3.4.2: Enforced coherence through behavioural recomposition of probabilities.

- We take into account the imprecision in the assessments. Recall, Section 3.3.1, that the three tasks identified the corresponding probabilities within an interval of width 10%. To acknowledge the uncertainty present, we use beta distributions with quantiles in the extremes of the interval having a probability of 0.9 and median in the midpoint of the interval. For example, participant 8 selected the option 30% in Task 1 of the first question; we assimilate it to the interval [0.25, 0.35]. By using routine `get.beta.par` from R, we obtain a $\beta eta(68.08, 158.56)$ distribution. We shall use such distributions to sample observations from the underlying distributions.

- We use the Brier (1950) score to assess the predictive capabilities of various models. Recall that, for a single question, the score is $(f - o)^2$ where f is the probability assigned to such event and o is 0 if it did not happen and 1 if it happened. We aim at minimising the score. For each participant and each question, using the actual answer of the question, we compute the expected Brier score, given the forecast uncertainty represented through the corresponding beta distribution. We then average the Brier scores over the three questions for each participant. In such a way, we obtain a sample from the Brier score distribution for a given predictive model.

We compare the forecasts derived from the direct assessment of p_A and our three recompositions based on p_B and p_C. Figure 3.4.3 represents the histograms of the differences between p_A and the three ARA coherent recompositions.



(a) EUM

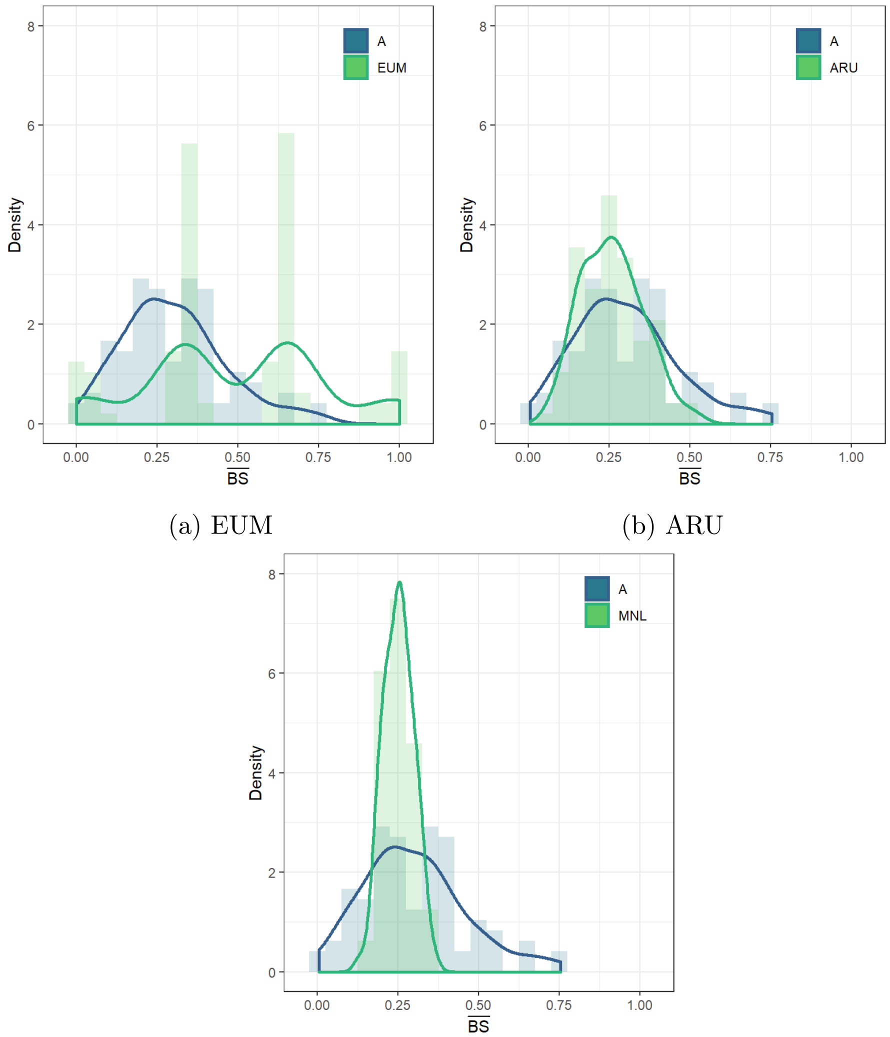
(b) ARU

(c) MNL

Figure 3.4.3: Histograms of the difference between direct elicitation and ARA recompositions.

Observe that there is bigger dispersion in panel (a) corresponding to the EUM recomposition (with several values -1 and 1) than in (c) and, specially (b), although the three histograms are located around 0. To further analyse the differences, Figure 3.4.4 compares the histograms of the Brier scores based on the direct probability assessments p_A and the three recompositions based p_B and p_C.

Recall that the smaller the score, whose range is [0,1], the better the predictive capacity of the model for such individual. Observe, thus, that the direct assessment seems better predictively that the recomposition based on EUM; the two modes in relation with the recomposition suggest that such model is too extreme. On the other hand, the other two recompositions seem to provide better predictive capabilities. To confirm these appreciations, we undertake hypothesis tests to check whether Brier scores attained with direct assessments are different that those attained with the behavioural recompositions. Assuming non-informative priors Figure 3.4.5 presents credible intervals with coverage probability 95% for the difference of the Brier score based on the direct assessment and the recomposed assessments. In the case of the EUM recomposition, 0 is above the interval and therefore we attain worse prediction capabilities, probably because of it being very extreme. On the other hand, for the other two recomposition methods, 0 is to the left of the intervals and, therefore, we obtain better, more accurate, predictions through both behavioural recompositions. This result suggests the behavioural recomposition is a promising way for analysing adversarial belief formation and designing effective SEJ methods in strategic situations.



(a) EUM

(b) ARU

(c) MNL

Figure 3.4.4: Brier score distribution for individuals based on direct assessment vs based on recompositions.

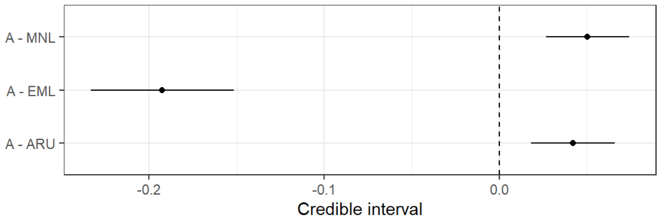


Figure 3.4.5: Credible intervals for differences of Brier scores with direct assessment and three behavioral recomposition methods.

It is thus natural to compare the performance of both recomposition methods via hypothesis tests. Figure 3.4.6 displays the 95% credible interval for the comparison between the ARU and MNL recompositions. The interval suggests no major difference between both methods, perhaps with some advantage for the MNL approach.

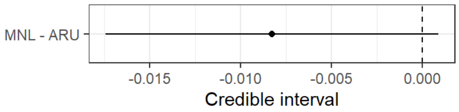
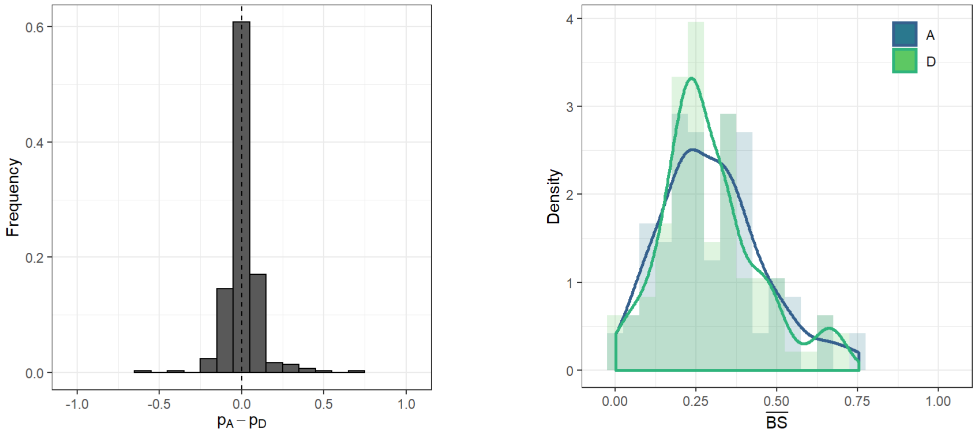


Figure 3.4.6: Credible intervals for differences of Brier scores with MNL and ARU recompositions.

3.4.3 Assessing the impact of reflection

As a final stage, we check whether upon the reflection induced by tasks, there are substantial changes concerning the probability p_A as reflected by p_D, with negative results. Figure 3.4.7a shows the histogram of the difference between both assessments.

The histograms suggest a similar behaviour of both estimates. We also check whether there is improvement in the predictive capabilities of both forecast



(a) Differences between probability estimates before and after the process.

(b) Brier scores for predictions based on p_D and p_A.

Figure 3.4.7: Impact of reflection.

approaches. Figure 3.4.7b provides the histograms of the Brier scores of the predictions based on the assessments before and after the tasks.

They suggest again a similar distribution. To confirm it, we use a hypothesis test much as before. Figure 3.4.8 provides the 95% credible interval for the average difference in Brier scores based on p_A and p_D. Observe that strictly speaking, 0 is not in the interval. However it is quite close and the average difference (0.009) is really small suggesting that the reflection induced by the proposed tasks had a minor effect on the the accuracy of the corresponding forecasts. The improvement was therefore minor.

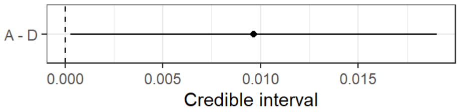


Figure 3.4.8: Credible interval - BS differences.

3.4.4 Conclusions

In summary, the following conclusions are suggested from the proposed experiments:

- We have identified relatively important disagreements between the probabilities assessed directly and those recomposed from the basic ingredients through the adversary expected utility model. The proposed recompositions allow us to enforce coherence.
- However, not all these recomposition rules seem equally relevant in predictive terms as based on Brier scores. In particular, we have that the best results seem to be derived from the ARU and MNL recomposition rules.
- The improvement in terms of predictability induced by the tasks has been minor.

3.5 Discussion

The need to forecast adversarial actions is frequent in cybersecurity and many other domains including national security, intelligence, and competitive business, to name but a few. This is a complex problem as we do not have directly available the adversaries to try to assess their preferences and beliefs and, consequently, predict their actions.

As a consequence we have introduced adversarial risk analysis as an approach to facilitate such strategic forecasts. ARA provides prescriptive support to one of the agents, based on a subjective expected utility model for a probability distribution of the decisions of the adversary. Such agent models the adversary's decision problem and, assuming that he is an expected utility maximiser (or has some other criterion, as in prospect theory), tries to assess his beliefs and preferences. If these were known, she could identify his optimal action. However, her uncertainty about the adversaries' beliefs and preferences is propagated to his decisions, leading to a probability distribution over his actions. ARA can thus be framed as a tool for SEJ elicitation when we need to deal with probabilities referring to actions by opponents.

In ARA models, the uncertainty behind the probability distribution of the adversary is a consequence of the uncertainty of the defender on the beliefs and preferences involved in the adversary's decision problem. However, we still need to take into account the uncertainty that the agent might have about her adversary's rationality, i.e. we need to model the adversary's behaviour. We have developed and validated three behavioural extensions of the ARA model used to recompose the strategic forecasts. Specifically, following a random utility approach, we consider that the adversary's utility includes a random component, whose realisation is not known by the agent when solving its optimisation problem. Such random component may be interpreted as a result of the agents' cognitive biases, decision making heuristics or errors in the implementation of the decision process and leads to a probability distribution over the adversary's actions. Two of the extensions provided recompositions which have shown particularly effective for adversarial forecasting purposes.

The proposed behavioural recomposition methods have been empirically validated with data from a face-to-face behavioural economics experiment. The results of this experiment suggest that behavioural recomposition based on ARU and MNL seem more accurate methods for SEJ than direct elicitation. Finally, the empirical evidence suggests that the reflection process required to make explicit the agents' beliefs on the probabilities and utilities to be considered by the adversary to make his mind is not capable to improve the accuracy of direct elicitation.

These results highlight the large cognitive burden and complexity level required to forecast adversarial actions in cybersecurity and many other relevant domains such as national security, intelligence, and competitive business. Therefore, adversarial beliefs formation seems to be a task for the conscious and analytic System 2 in Kahneman's dual thinking model (Kahneman, 2012). Since in most of the cases System 2 is overcome by the fast and automatic System 1, the result is that direct beliefs elicitation does not take into account the strategic aspects, providing then lousier estimations. An effective way to address this issue may be to rely on direct elicitation methods just for the belief on basic elements and perform externally the task that would have correspond to the lazy System 2. This is the approach after recomposition methodologies, which (i) focused on an accurate elicitation of the beliefs on the basic building blocks of the adversary's decision analysis; and (ii) recompose the beliefs on the complex final question of estimating the chances of an adversary's action to be selected by applying ARA models integrating a behavioural perspective.

The results of this paper also contribute to our understanding of the mechanism of beliefs formation in adversarial situations. We have shown that an agent's probability distribution on the actions to be chosen by an adversary, as obtained by direct elicitation, differs from that obtained from ARA and behavioural recompositions. Such a difference remains even the direct elicitation is performed after the reflection exercise of eliciting her beliefs on each individual elements of her adversary's decision problem. This fact suggests that strategic analysis does not play a relevant role in the mechanism of beliefs formation.

Relevant open research issues include the consideration of other behavioural models to be integrated in the ARA model and capable of increasing the accuracy of the estimates: cognitive biases related to uncertainty processing (such as probability insensitivity or formation of decision weights) or decision heuristics (such as anchoring). Further research is required to develop effective behavioural recomposition methods for specific fields, such as cybersecurity risk analysis, and validate them with larger size field and/or economics experiments.

References

Andradottir, S. and Bier, V. M. (1997). Choosing the number of conditioning events in judgemental forecasting. *Journal of Forecasting*, 16(4):255–286.

Andradottir, S. and Bier, V. M. (1998). An analysis of decomposition for subjective estimation in decision analysis. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 28(4):443–453.

Banks, D. L., Aliaga, J. M. R., and Insua, D. R. (2015). *Adversarial risk analysis*. CRC Press.

Block, H. D., Marschak, J., et al. (1959). Random orderings and stochastic theories of response. Technical report, Cowles Foundation for Research in Economics, Yale University.

Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly weather review*, 78(1):1–3.

Busemeyer, J. R. and Rieskamp, J. (2014). Psychological research and theories on preferential choice. In *Handbook of choice modelling*, pages 49–72. Edward Elgar Publishing.

Chen, E., Budescu, D. V., Lakshmikanth, S. K., Mellers, B. A., and Tetlock, P. E. (2016). Validating the contribution-weighted model: Robustness and cost-benefit analyses. *Decision Analysis*, 13(2):128–152.

Clemen, R. T. and Reilly, T. (2013). *Making hard decisions with DecisionTools*. Cengage Learning.

Cleveland, W. (1985). *The Elements of Graphing Data.* Addison-Wesley.

Cooke, R. et al. (1991). *Experts in uncertainty: opinion and subjective probability in science.* Oxford University Press on Demand.

Dias, L. C., Morton, A., and Quigley, J. (2018). *Elicitation: The science and art of structuring judgement.* Springer International Publishing.

González-Ortega, J., Radovic, V., and Insua, D. R. (2018). Utility elicitation. In *Elicitation: The science and art of structuring judgement*, pages 241–264. Springer.

Hanea, A., Nane, T., Bedford, T., and French, S. (2021). *Expert Judgement in Risk and Decision Analysis.* Springer, Chum, Switzerland.

Hargreaves-Heap, S. and Varoufakis, Y. (2004). *Game theory: A critical introduction.* Routledge.

Hess, S., Daly, A., and Batley, R. (2018). Revisiting consistency with random utility maximisation: theory and implications for practical work. *Theory and Decision*, 84(2):181–204.

Insua, D. R., Banks, D. L., and Aliaga, J. M. R. (2020). *Adversarial Risk Analysis as a Decomposition Method for Structured Expert Judgement Modelling.* Springer.

Kadane, J. B. and Larkey, P. D. (1982). Subjective probability and the theory of games. *Management Science*, 28(2):113–120.

Kahneman, D. (2012). *Think Fast and Slow.* Penguin. London.

Keeney, R. L. (2007). Modeling values for anti-terrorism analysis. *Risk Analysis: An International Journal*, 27(3):585–596.

Luce, R. D. (1977). The choice axiom after twenty years. *Journal of mathematical psychology*, 15(3):215–233.

MacGregor, D. G. (2001). Decomposition for judgmental forecasting and estimation. In *Principles of forecasting*, pages 107–123. Springer.

Marschak, J. (1959). Binary choice constraints and random utility indicators. In *K. Arrow (Ed.), Stanford symposium on mathematical models in the social sciences.* Stanford, CA: Stanford University Press.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In *P. Zarembka (Ed.), Frontiers in econometrics.*, pages 2373–2375. New York: Academic Press.

Montibeller, G. and von Winterfeldt, D. (2015). Biases and debiasing in multi-criteria decision analysis. In *2015 48th Hawaii International Conference on System Sciences*, pages 1218–1226. IEEE.

O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D. J., Oakley, J. E., and Rakow, T. (2006). *Uncertain judgements: eliciting experts' probabilities.* John Wiley & Sons.

Pleskac, T. J. (2015). Decision and choice: Luce's choice axiom. *International encyclopedia of the social & behavioral sciences*, 5:895–900.

Raiffa, H. (1968). *Decision analysis: Introductory lectures on choices under uncertainty.* Addison-Wesley.

Raiffa, H. (1982). *The art and science of negotiation.* Harvard University Press.

Ravinder, H. (1992). Random error in holistic evaluations and additive decompositions of multiattribute utility—an empirical comparison. *Journal of Behavioral Decision Making*, 5(3):155–167.

Ravinder, H. and Kleinmuntz, D. N. (1991). Random error in additive decompositions of multiattribute utility. *Journal of Behavioral Decision Making*, 4(2):83–97.

Rios Insua, D., Banks, D., and Rios, J. (2016). Modeling opponents in adversarial risk analysis. *Risk Analysis*, 36(4):742–755.

Tetlock, P. E. and Gardner, D. (2015). *Superforecasting: The art and science of prediction.* Broadway Books.

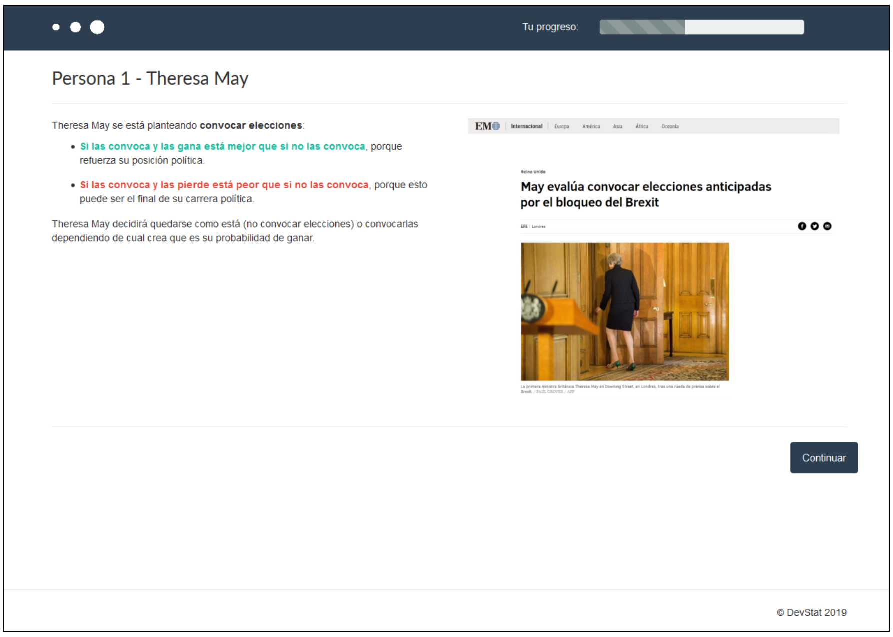
Thurstone, L. L. (1927). A law of comparative judgment. *Psychological review*, 34(4):273.

Watson, S. R. and Brown, R. V. (1978). The valuation of decision analysis. *Journal of the Royal Statistical Society: Series A (General)*, 141(1):69–78.

Appendix

3.A Appendix: Screenshots from experiment

This appendix shows the screenshots of the experiment for the first of the three decisions considered in the experiment (Theresa May's decision of calling for elections in 2018). The screenshots are in Spanish, since the experiment was run in Spain and no version of the experiment in English was developed.



Persona 1 - Theresa May

Theresa May se está planteando **convocar elecciones**:

- **Si las convoca y las gana está mejor que si no las convoca**, porque refuerza su posición política.
- **Si las convoca y las pierde está peor que si no las convoca**, porque esto puede ser el final de su carrera política.

Theresa May decidirá quedarse como está (no convocar elecciones) o convocarlas dependiendo de cual crea que es su probabilidad de ganar.

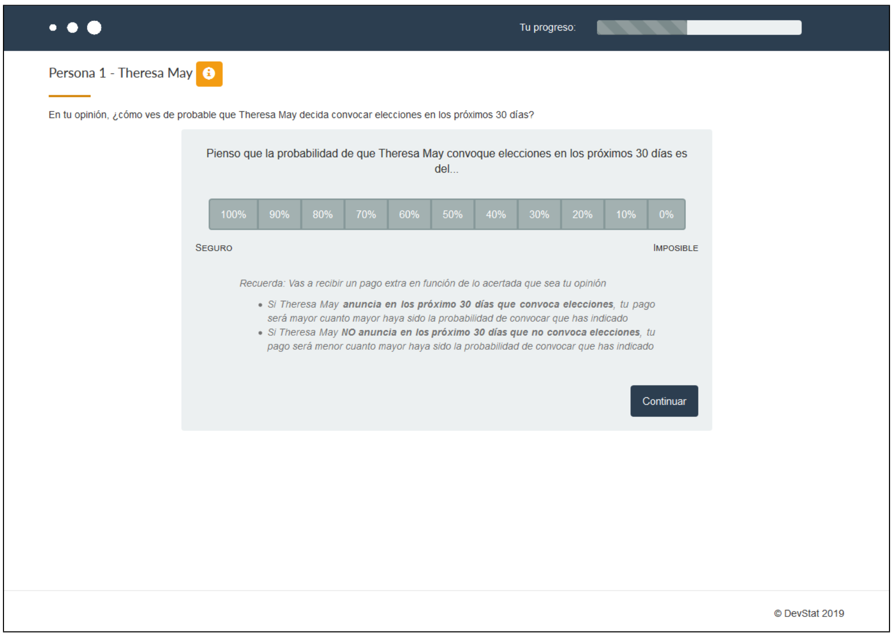
Reino Unido

May evalúa convocar elecciones anticipadas por el bloqueo del Brexit

EFE | Londres

La primera ministra británica Theresa May en Downing Street, en Londres, tras una rueda de prensa sobre el Brexit. / PAUL GROVER / AFP

Figure 3.A.1: First case considered in the experiment.



Persona 1 - Theresa May

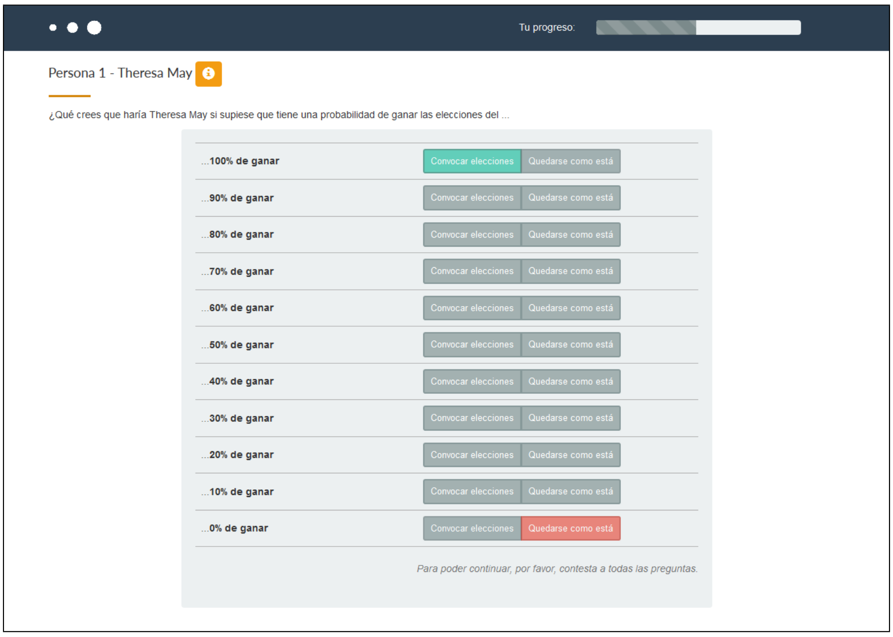
En tu opinión, ¿cómo ves de probable que Theresa May decida convocar elecciones en los próximos 30 días?

Pienso que la probabilidad de que Theresa May convoque elecciones en los próximos 30 días es del...

Recuerda: Vas a recibir un pago extra en función de lo acertada que sea tu opinión

- *Si Theresa May **anuncia en los próximo 30 días que convoca elecciones**, tu pago será mayor cuanto mayor haya sido la probabilidad de convocar que has indicado*
- *Si Theresa May **NO anuncia en los próximo 30 días que no convoca elecciones**, tu pago será menor cuanto mayor haya sido la probabilidad de convocar que has indicado*

Figure 3.A.2: First task: Disclosure of p_A.



Tu progreso:

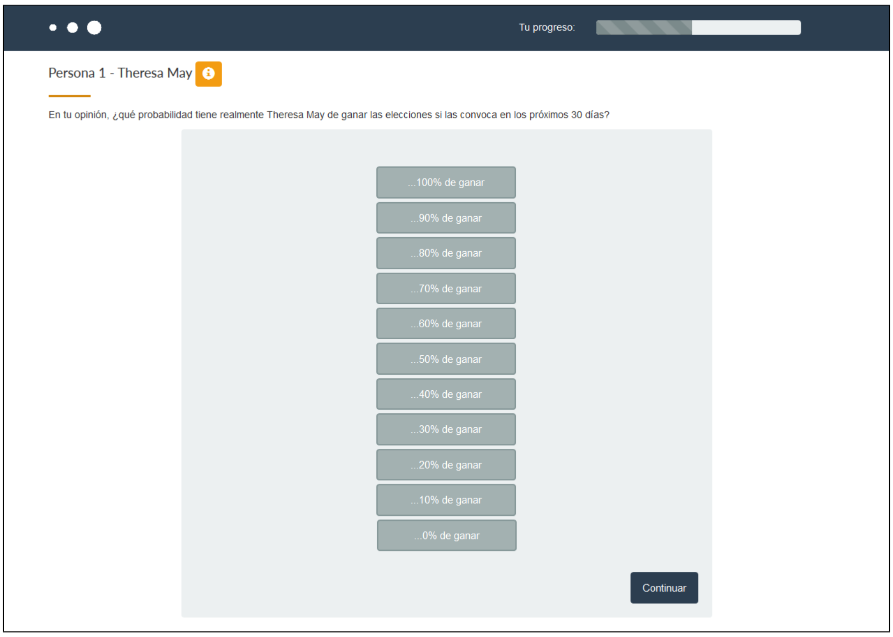
Persona 1 - Theresa May

¿Qué crees que haría Theresa May si supiese que tiene una probabilidad de ganar las elecciones del ...

...100% de ganar	[✓ Convocar elecciones]	Quedarse como está
...90% de ganar	Convocar elecciones	Quedarse como está
...80% de ganar	Convocar elecciones	Quedarse como está
...70% de ganar	Convocar elecciones	Quedarse como está
...60% de ganar	Convocar elecciones	Quedarse como está
...50% de ganar	Convocar elecciones	Quedarse como está
...40% de ganar	Convocar elecciones	Quedarse como está
...30% de ganar	Convocar elecciones	Quedarse como está
...20% de ganar	Convocar elecciones	Quedarse como está
...10% de ganar	Convocar elecciones	Quedarse como está
...0% de ganar	Convocar elecciones	[✓ Quedarse como está]

Para poder continuar, por favor, contesta a todas las preguntas.

Figure 3.A.3: Second task: Disclosure of p_B.



Tu progreso:

Persona 1 - Theresa May

En tu opinión, ¿qué probabilidad tiene realmente Theresa May de ganar las elecciones si las convoca en los próximos 30 días?

...100% de ganar

...90% de ganar

...80% de ganar

...70% de ganar

...60% de ganar

...50% de ganar

...40% de ganar

...30% de ganar

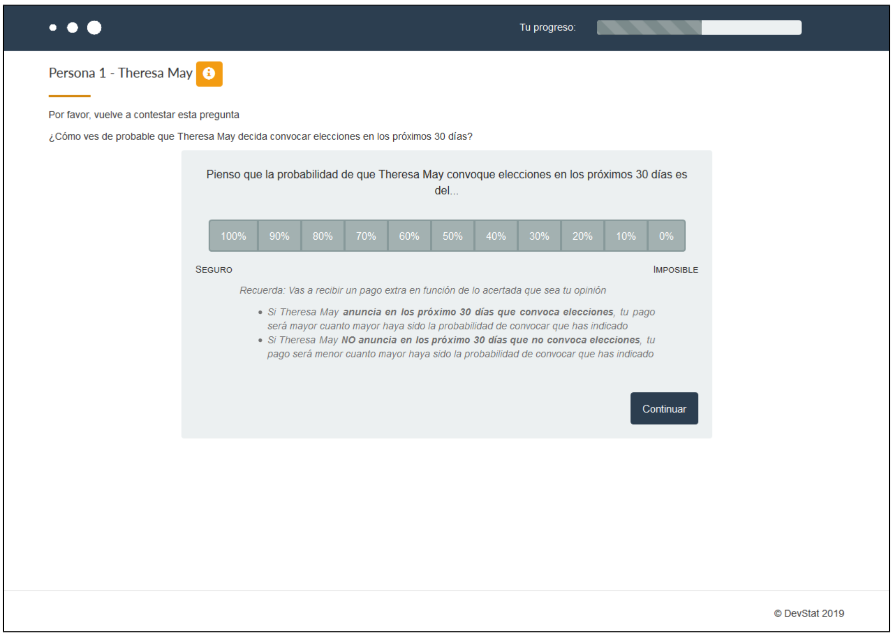
...20% de ganar

...10% de ganar

...0% de ganar

Continuar

Figure 3.A.4: Third task: Disclosure of p_C.



Persona 1 - Theresa May

Por favor, vuelve a contestar esta pregunta

¿Cómo ves de probable que Theresa May decida convocar elecciones en los próximos 30 días?

Pienso que la probabilidad de que Theresa May convoque elecciones en los próximos 30 días es del...

Recuerda: Vas a recibir un pago extra en función de lo acertada que sea tu opinión

- *Si Theresa May* ***anuncia en los próximo 30 días que convoca elecciones****, tu pago será mayor cuanto mayor haya sido la probabilidad de convocar que has indicado*
- *Si Theresa May* ***NO anuncia en los próximo 30 días que no convoca elecciones****, tu pago será menor cuanto mayor haya sido la probabilidad de convocar que has indicado*

Figure 3.A.5: Fourth task: Disclosure of p_D.

3.B Appendix: Distribution of features for respondents

The distribution by profile of the respondents is shown in Table 3.B.1. Three quarters of them were public/private workers and were almost equally distributed between both genders. Regarding the education of participants, most of them had either some years of university or a university degree.

Profile	**N**	**%**
Total	*96*	*100.0*
18-35	26	27.1
36-50	39	40.6
51-74	31	32.3
Male	50	52.1
Female	46	47.9
Primary education	19	19.8
Secondary education	28	29.2
Tertiary education	49	51.0
Freelance	8	8.3
Public/private worker	72	75.0
Unemployed	7	7.3
Other	9	9.4

Table 3.B.1: Distribution of participants by age, gender, education and employment.

Conclusions

Discussion

Aiming at filling an existing gap in the cyberinsurance literature, we have presented the design and the results of the empirical validation of a series of behavioural models helping to understand and predict human cybersecurity decision-making. Our findings provide relevant insights on the adoption of cyberinsurance, including its relations with other dimensions of cybersecurity (cyberprotection and online behaviour), as well as on belief formation on adversarial situations such as intentional cyberattacks.

As a departing point to motivate the behavioural approach to cybersecurity, we show in **Chapter 1** that the expected utility maximisation fails to predict decisions on cyberinsurance and cyberprotection adoption. Specifically, we found that subjects tend to overprotect and overinsure themselves with respect to the protection and coverage suggested by rational choice theory. Moreover, we establish that cyberinsurance and cyberprotection products are complementary, instead of substitutive. A reason that could explain this pattern is the existence of two segments of agents in terms of their awareness and concerns towards cybersecurity issues. Note that the complementarity of these two types of goods mitigates a critical hazard of the expected growth of the cyberinsurance market. We also show that higher coverage levels are not compensated by a reduction in the protection level of the agents, which may facilitate the expansion of cyberattacks and compromise the resilience of our increasingly digital markets and societies. Moreover, although our experimental design does not establish a causal relation from insurance to protection, the result suggests that there is no moral hazard for failures of protection in the cybersecurity market. A similar result can be obtained for the safety level during online navigation, since the empirical evidence presented in this dissertation rejects that subjects with a higher level of coverage behave in a less safe on the Internet. Again, and without being capable to test the existence of a causal relation, this fact suggests the absence of moral hazard effects coming from the adoption of cyberinsurance. These results, and the failure of predictive power of the rational choice model, suggests the convenience to develop new models with a sound behavioural foundation capable to explain insurance adoption, as well as identify and quantify the impact of potential irrational levers conducting to cyberinsurance purchase, as approached in Chapter 2.

Even knowing that the probability of suffering an intentional and a random unintentional attack is the same, we find that agents do not respond in the same way to both types of attacks. In fact, they adopt higher protection and insurance coverage levels in case of facing random cybermenaces. In the absence of previous experiences with adverse events, the dramatic communication impact of recent large-scales random cyberattacks may reinforce the perception of the risk of suffering a random attack (availability bias), as happened with the 2017 massive WannaCry attack. Moreover, Random cyberattacks and the protection against them are not something that can usually be contained within a single organisation and spread fast on the Internet, which may also increase the risk perception for such random attacks coming from multiple connections. On the other hand, risk beliefs on vulnerability from intentional attacks can be affected by cognitive biases such as optimism bias ("An attacker won't target my business"), which may result in some individuals assuming that intentional cyberattacks will not happen to them. A critical consequence of these misbeliefs on the risk of intentional cyberattacks is that individuals and organisations may not be investing time in understanding their vulnerabilities nor adopting as many protection measures and insurance coverage than in the case of intentional cyberattacks. This discussion motivates a deeper analysis of risk belief formation and elicitation methods in adversarial situations, as presented in Chapter 3.

The interpretation of the calibration of the predictive Structural Equation Modelling (SEM) of cyberinsurance adoption in **Chapter 2** provides valuable behavioural insights on cybersecurity related decision making. These insights, theoretically founded on Protection Motivation Theory (PMT) and the Theory of Planned Behaviour (TPB), help to fill the gaps in rational choice models identified in the previous chapter. Specifically, individuals who adopted advanced security measures are more likely to also adopt premium insurance, explaining the complementarity of protection and insurance products. Moreover, this is the strongest pathway in the model. The adoption of advanced security measures was also significantly positively related to security of online behaviour; those who adopted advanced security measures were also more likely to behave securely online. Response efficacy and the TPB factors (attitudes and norms) are also positively related to adoption of premium insurance. Perceived self-efficacy and perceived threat severity both positively fed into the adoption of advanced security measures rather than adoption of premium insurance directly. Those who had higher perceptions of their ability to put cybersecurity measures into place, and those who perceived the threat of the cyberattack as more severe, were more likely to adopt advanced security measures. As aforementioned the adoption of security measures then subsequently fed into premium insurance adoption. Risk propensity was negatively related to both adoption of security measures and adoption of insurance, i.e., a risk-seeking individual was less likely to adopt advanced security measures and premium insurance.

The findings presented in **Chapter 3** provides an empirical validation of the effectivity of behavioural recomposition as a belief elicitation method in strategic adversarial situations. Specifically, we show that this method, which integrates the random utility approach into ARA modelling, is more accurate for belief elicitation in strategic settings than just direct elicitation techniques and standard ARA models. Moreover, the experimental results show that the reflection process required to make explicit the agents' beliefs on the probabilities of the success of an attack and adversary's utilities does not improve the accuracy of direct elicitation. These results highlight the large cognitive burden and complexity level required to forecast adversarial actions in cybersecurity and many other relevant domains such as national security, intelligence, and competitive business. Therefore, adversarial beliefs formation seems to be a task for the conscious and analytic System 2 in Kahneman's dual thinking model (Kahneman, 2012). Since in most of the cases System 2 is overcome by the fast and automatic System 1, the result is that direct beliefs elicitation does not take into account the strategic aspects, providing then lousier estimations. An effective way to address this issue may be to rely on direct elicitation methods just for the belief on basic elements and perform externally the task that would have correspond to the lazy System 2. This is the approach after recomposition methodologies, which (i) focused on an accurate elicitation of the beliefs on the basic building blocks of the adversary's decision analysis; and (ii) recompose the beliefs on the complex final question of estimating the chances of an adversary's action to be selected by applying ARA models integrating a behavioural perspective.

The results of this last chapter also contribute to our understanding of the mechanism of beliefs formation in adversarial situations. We have shown that an agent's probability distribution on the actions to be chosen by an adversary, as obtained by direct elicitation, differs from that obtained from ARA and behavioural recompositions. Such a difference remains even the direct elicitation is performed after the reflection exercise of eliciting her beliefs on each individual elements of her adversary's decision problem. This fact suggests that strategic analysis does not play a relevant role in the mechanism of beliefs formation. As a consequence, the difference in protection and insurance level identified in Chapter 1 may be the results of other behavioural levers rather than of a reduction on the cybervulnerability perception. Anyway, this questions remains open and needs to be answered by future research.

Relevance

The findings presented in this dissertation have significant basic and applied scientific relevance, as well as practical applications in the industry and in policy making.

From a ***basic science*** viewpoint, our results contribute to both (i) the literature on the development of behavioural predictive models with sound psychological foundations which, going beyond the perfectly rational-decision approach, are capable to explain empirical data on decision-making and; (ii) the literature on belief formation in strategic interactions under uncertainty. As regards with the first contribution, and after showing in Chapter 1 that rational choice models are not able to explain the patterns found in our experimental data, we propose a new model integrating elements of two relevant psychologic theories of behavioural change, namely Protection Motivation Theory (PMT) and Theory of Planned Behaviour (TPB). The integration and calibration methodology proposed in this dissertation, Structural Equation Modelling (SEM), is disruptive and opens a new approach to model behavioural-experimental data under uncertainty with sound theoretical behavioural models. To the best of our knowledge, the SEM presented in Chapter 2 is the first attempt in the literature to use this method to validate the existence of causal relations from a series of psychological constructs (included in the model as latent variables, measured trough observable variables coming both from behavioural measures observed in an economic experiment and subject's self-assessment trough validated psychometric scales) and actual behavioural change (measured from actual responses of subject in a controlled and incentivised environment, instead of hypothetical answers to what-if items in a questionnaire). Our results contribute also to the understanding the process of belief formation beyond the common knowledge conditions assumed in rational choice models and game theory. As regards belief formation, combining Adversarial Risk Analysis (ARA) and random choice models to reconstruct the beliefs of an agent on the chances of a strategic adversary to select an option among a set of potential actions and validating experimentally the accuracy of such a combination is something that has not been previously done in the literature. The abilities of this approach to improve belief elicitation gives light on the mechanisms of belief formation and provides a usable alternative to the strategic common knowledge assumptions implicit in standard Game Theory and supporting the Nash Equilibrium solution to strategic interactions.

The results of this monography contribute also to the *applied research in the field of cybersecurity*, especially in the field of behavioural cyberinsurance. The SEM model of cyberinsurance adoption presented in Chapter 2 addresses a significant gap in the existing literature: despite a fast-growing interest in, and industry around, cybersecurity, there is an overwhelming lack of knowledge in relation to understanding the mechanisms behind cybersecurity decision-making. The results support the model as a good fit to the data therefore providing important knowledge of the factors influencing cybersecurity decisions, including uptake of security measures and insurance. In this area, our results contribute to a better understanding of the main features of cybersecurity adoption and its implications. On the other hand, the SEM adoption model shows that, although the perceived severity of cyberattacks affects the levels of cyberprotection or insurance cybercoverage of an agent, perceived vulnerability to cyberattacks is not a significant predictor of such levels. These are troubling findings for the PMT model as applied to cyberthreat and beg the question as to how the beliefs of vulnerability are built and how to develop new approaches to improve belief formation on the probability of suffering an attack. The results of Chapter 3 provides an answer to these question, contributing to fill such a critical gap. Finally, our research work does also contribute to collect reliable data on cyberinsurance adoption obtained in controlled situations and make them publicly available at `https://www.cybeco.eu/`. In a research fieldwork where data are so scarce, this dataset can also be used by other researchers to obtain additional behavioural insights of cyberinsurance adoption.

In this dissertation we identified key factors underlying decision-making around cybersecurity. The model presented here could be used to guide future interventions aimed at increasing cyberinsurance (and cybersecurity) uptake, supporting the **practical relevance** of our results for both policy-making and the setting of industrial strategy in a cyberinsurance sector. In a policy-making context to enhance cybersecurity and guarantee the resilience of digital systems is a key objective, as stated in the European Digital Strategy (European Commission, 2018). Although previously establish in the literature, this dissertation reinforces the empirical evidence on the fact that rational choice models are not capable to explain cyberinsurance adoption. This fact justifies the room for the implementation of behavioural based policy to nudge towards safer cybersecurity decisions. Moreover, the interpretation of the SEM models calibrated in Chapter 2 provide relevant behavioural insights for policy design. The findings in this monography are also relevant in the cyberinsurance industry in critical areas, such as cyber-risk assessment or the design of the policies portfolio. As shown in Chapter 3, unguided self-assessment risk protocols seem not be an effective way to obtain accurate estimates of the probabilities of strategic events to take place, for instance intentional cybercrime attacks. A more accurate risk estimation can be obtained from a guided assessment procedure, in which (i) adversarial strategic decisions are decomposed into a series of basic elements; (ii) beliefs on each basic element are independently elicited; and (iii) probabilities on the composed events are recovered by using ARA models with a random-choice approach. These guided procedures could be implemented in the cybersecurity audits and fine-tuning policy design in terms of vulnerability assessments, since, as shown in Chapter 1, this approach can increase the level of rationality of the level of coverage to be adopted trough cyberinsurance products.

Limitations

The results presented in this dissertation are not free of limitations. First of all, beyond cyberinsurance, our research focuses in just two other dimensions of the cybersecurity strategy, namely protection and safety of online behaviour. These dimensions have been selected for their relevance and the fact that they illustrate some of the most relevant issues in the relation of cyberinsurance adoption and the rest of cybersecurity elements. For instance, cyberprotection is observable (provided the corresponding audit mechanisms) and competes with cyberinsurance for the allocation of a limited budget for the purchase of commercial products (for instance, firewalls, antivirus versus insurance policies). Alternatively, decisions on online behaviour are made after the adoption of protection and insurance, have in general no budgetary implications and are hard to be observed by potential insurers. Despite of the diverse features of these two dimensions, other critical aspects of cyberinsurance, such as cyber-recovery or information sharing on cyberattacks, have been not covered by our research, which may translate in the loss of potential relevant behavioural insights.

A second limitation comes from the simplification of the protection and insurance products, as well as of the online navigation options, made available to experimental subjects. Specifically, we consider that protection measures only affect the vulnerability (probability) to receive an attack, but not its severity or the recovery options after the attack. On the other hand, we consider that the impact of an attack is mainly economic, with references to other critical effects such as the loss of customer's trust. Moreover, the cyberinsurance portfolio consists of just two different insurance policies, with a limited information which is easy to understand. This limitation takes out from our research relevant behavioural elements of cyberinsurance, such as the cognitive charge to understand the different products or decision heuristics to be applied by insurance takers in more realistic conditions.

Finally, and as for any other results in behavioural-experimental economics, the ecological validity of the results (i.e. the reliability of a translation of the results from the experimental data into real world behaviour) may be considered as a potential limitation of our results. However, the application of behavioural economics experiments to obtain behavioural insights in cybersecurity has become an established practice in research (van Bavel et al., 2019; Rodríguez-Priego et al., 2020; Insua et al., 2020) and has been successfully applied for evidence-supported policy-making, for instance by the European Commission (van Bavel et al., 2016; Monteleone et al., 2015; van Bavel et al., 2015).

Future research

The results presented in this dissertation motivate additional questions for further research. A first line in our research agenda will aim to enriching the behavioural models presented in Chapters 1 and 2 in several directions, such as (i) the inclusion of additional dimensions of cybersecurity (for instance, cyber-risk assessment or cyber-recovery), (ii) the consideration of more complex and realistic portfolios of protection measures and insurance products, and (iii) the application of quantitative methods for the analysis of experimental data beyond SEM and multinomial logit models. The foreseen enrichment of the models also includes the experimental calibration of the probability weighting function, in addition to the calibration of the utility functions as presented in Chapter 1. Although the achievement of this objective will require the design and implementation of additional behavioural economics experiments, it will made possible to analyse if Prospect Theory can predict observational data on cyberinsurance adoption in a more accurate way than Expected Utility. The new experiments are also planned to be implemented in an interactive way, with the participation of real subjects in both the role of defender and attacker.

The results of the behavioural models calibrated in Chapter 2, as well as other potentially coming from the enriched models and behavioural economics experiments to be developed in the future, will motivate another research line focused on the design and experimental validation of specific nudges devoted to promoting behavioural change in cybersecurity. These nudges, i.e. interventions addressing the automatic thinking System 1 as described in Kahneman (2012), will be designed to make cyberinsurance decision-making more according to the rational choice model solutions and to design other effective cybersecurity policies according to the objectives defined by the European Commission in its Digital Strategy (European Commission, 2018).

Several interesting questions for future research arise from the more basic research results regarding beliefs formation and elicitation methods. Deepening in the basic research implications of our work, we foresee to develop and validate (using large scale experiments) new models for behavioural recomposition. These new models will integrate alternative behavioural economics models for decision-making into the ARA modelling techniques. Aligned with the objectives of the previous paragraph, a relevant option will be the integration of the weighting function proposed by Prospect Theory in the decision problem faced by the attacker (always from the defender's perspective considered in ARA methods). We will also answer to a critical question that remains unanswered in this dissertation: Does make intentional attacks make the defender to feel less cybervulnerable than the menace of unintentional random attacks? To answer this question, we foresee the implementation of behavioural economics experiments of beliefs elicitation, based on the methodology in those in Chapter 3 and considering an additional treatment in terms of the internationality or randomness of the attack. The analysis of these experiments will provide empirical evidence to test if the intentionality is actually affecting the vulnerability beliefs or if the impact from intentionality to cybersecurity behaviour (as identified in Chapter 1) follows a different path.

Finally, an given that lack of effective cyber-risk assessment methods for a sound development of the cyberinsurance sector, our future research agenda will build from the learning in Chapter 3 to design and validate empirically a cyber-risk assessment tool for intentional attacks. This tool, planned as a web application, will be based in behavioural recomposition methods and will also include a behavioural approach to optimise user experience for information uploading and a nudge approach for the presentation of the results of the risk assessment.

References

European Commission (2018). European commission digital strategy: A digitally transformed, user-focused and data-driven commission. `https://ec.europa.eu/info/sites/info/files/strategy/decision-making_process/documents/ec_digitalstrategy_en.pdf`.

Insua, D. R., Baylon, C., and Vila, J. (2020). *Security Risk Models for Cyber Insurance*. CRC Press.

Kahneman, D. (2012). *Think Fast and Slow*. Penguin. London.

Monteleone, S., van Bavel, R., Rodríguez-Priego, N., and Esposito, G. (2015). Nudges to privacy behaviour: Exploring an alternative approach to privacy notices. *JRC Science and Policy Report. Luxembourg, Luxembourg: Publications Office of the European Union.*

Rodríguez-Priego, N., van Bavel, R., Vila, J., and Briggs, P. (2020). Framing effects on online security behavior. *Frontiers in Psychology*, 11:2833.

van Bavel, R., Rodríguez-Priego, N., et al. (2016). Testing the effect of the cookie banners on behaviour. *JRC Technical Reports.*

van Bavel, R., Rodríguez-Priego, N., Maghiros, I., et al. (2015). Seven points to remember when conducting behavioural studies in support of eu policy-making. *JRC Scientific and Policy Reports. Luxembourg: Publications Office of the European Union.*

van Bavel, R., Rodríguez-Priego, N., Vila, J., and Briggs, P. (2019). Using protection motivation theory in the design of nudges to improve online security behavior. *International Journal of Human-Computer Studies*, 123:29–39.

Conclusiones

Discusión

Con el objetivo de llenar un vacío en la literatura sobre ciberseguro, hemos presentado el diseño y los resultados de la validación empírica de una serie de modelos conductuales que ayudan a comprender y predecir la toma de decisiones en ciberseguridad. Nuestros resultados brindan información relevante sobre la adopción del ciberseguro y sus relaciones con otras dimensiones de la ciberseguridad (ciberprotección y comportamiento online), así como sobre la formación de creencias en contextos adversariales, tales como ciberataques intencionales.

Como punto de partida para motivar el enfoque conductual de la ciberseguridad, mostramos en el **Capítulo 1** que la maximización de la utilidad esperada no puede predecir las decisiones observadas de adopción de ciberseguros y ciberprotección. Específicamente, los sujetos tienden a sobreprotegerse y sobreasegurarse con respecto a la protección y cobertura sugeridas por la teoría de elección racional. Además, establecemos que los productos de ciberseguro y ciberprotección son complementarios en lugar de sustitutivos. Una razón que podría explicar este patrón es la existencia de dos tipos de agentes en cuanto a su concienciación e inquietud hacia los temas de ciberseguridad. También mostramos que los niveles de cobertura más altos no se compensan con una reducción en el nivel de protección de los agentes. De esta forma, aunque nuestro diseño experimental no establece una relación causal entre el seguro y la protección, el resultado sugiere que no existe riesgo moral asociado a la ciberseguridad. Un resultado similar se puede obtener para el nivel de seguridad durante la navegación online, ya que la evidencia empírica presentada en esta disertación rechaza que los sujetos con un mayor nivel de cobertura se comporten de una forma menos segura en Internet. Nuevamente, y sin poder contrastar la existencia de una relación causal, este hecho sugiere la ausencia de efectos de riesgo moral provenientes de la adopción del ciberseguro. Estos resultados, y la falta del poder predictivo del modelo de elección racional, sugieren la conveniencia de desarrollar nuevos modelos con una base de comportamiento sólida capaz de explicar la adopción de seguros, así como identificar y cuantificar el impacto de las posibles palancas irracionales que conducen a la compra de ciberseguros, tal y como se aborda en el Capítulo 2.

Aún sabiendo que la probabilidad de sufrir un ataque intencional y uno aleatorio no intencional es la misma, encontramos que los agentes no responden de la misma forma a ambos tipos de ataques. De hecho, adoptan niveles más altos de protección y cobertura de seguros en caso de enfrentarse a amenazas cibernéticas aleatorias. En ausencia de experiencias previas con eventos adversos, el impacto que tiene la comunicación de los recientes ciberataques aleatorios a gran escala, como el ataque masivo WannaCry de 2017, puede reforzar la percepción del riesgo de sufrir este tipo de ataques (sesgo de disponibilidad). Además, los ciberataques aleatorios no están contenido dentro de una sola organización y se propagan rápidamente en Internet, lo que también puede aumentar la percepción del riesgo de recibir ataques aleatorios provenientes de las múltiples conexiones que tiene el agente. Por otro lado, las creencias de riesgo sobre la vulnerabilidad ante ataques intencionales pueden verse afectadas por sesgos cognitivos como el sesgo de optimismo ("Un cibercriminal no va a prestar atención a mi negocio"), lo que puede hacer que algunas personas asuman que los ciberataques intencionales no les sucederán a ellos. Una consecuencia de este tipo de percepciones sesgadas es que personas y organizaciones no inviertan el tiempo necesario en comprender sus vulnerabilidades reales ni los recursos para adoptar medidas de protección y cobertura apropiadas frente a ciberataques intencionales. Esta discusión motiva el análisis más profundo de los métodos de formación y elicitación de creencias de riesgo en situaciones adversariales que se presenta en el Capítulo 3.

La interpretación de la calibración del Modelo de Ecuaciones Estructurales (SEM) del **Capítulo 2** proporciona información valiosa sobre el comportamiento en la toma de decisiones de ciberseguridad. Estos modelos, basados en la Teoría de la Motivación a la Protección (PMT) y en la Teoría de Acción Planeada (TPB), resuelven la falta de capacidad predictiva de los modelos de elección racional, identificada en el capítulo anterior. El modelo SEM predice que es más probable que las personas que adoptaron medidas de seguridad avanzadas también adopten seguros premium, lo que explica la complementariedad de los productos de protección y de los seguros. La adopción de medidas de seguridad avanzadas también se relaciona significativamente y de manera positiva con el nivel de seguridad del comportamiento online; aquellos sujetos que adoptaron medidas de seguridad avanzadas tienden a comportarse de forma más segura online. La eficacia de la respuesta y los factores de la TPB (actitudes y normas) también se relacionan positivamente con la adopción de seguros premium. La autoeficacia percibida y la gravedad de la amenaza percibida influyen positivamente en la compra de medidas de seguridad avanzadas y en la adopción de seguros premium. Los sujetos con una mayor percepción de su capacidad para implementar medidas de ciberseguridad y aquellos que perciben la amenaza del ciberataque como más severa, tienen también más posibilidades de adoptar medidas de seguridad avanzadas. Como se ha mencionado anteriormente, la adopción de medidas de seguridad se asocia también con la adquisición de mayores niveles de cobertura. La aversión al riesgo se relaciona positivamente tanto con la adopción de medidas de seguridad como con la de ciberseguros.

Los resultados presentados en el **Capítulo 3** proporcionan una validación empírica de la efectividad de los métodos de recomposición conductual propuestos para la elicitación de creencias en interacciones estratégicas. Concretamente, mostramos que un método que integre el enfoque de utilidad aleatoria con en el modelado ARA es más preciso para la revelación de creencias en entornos estratégicos que las técnicas directas y los modelos ARA estándar. Además, los resultados experimentales muestran que el proceso de reflexión requerido para hacer explícitas las creencias de los agentes sobre las probabilidades de éxito de un ataque y las utilidades del adversario no mejora la precisión de la elicitación directa. Estos resultados destacan la gran carga cognitiva y el nivel de complejidad necesarios para pronosticar acciones adversas en ciberseguridad y muchos otros dominios relevantes como seguridad nacional, inteligencia y negocios competitivos. Por lo tanto, la formación de creencias adversas parece ser una tarea para el consciente y analítico Sistema 2 en el modelo de pensamiento dual de Kahneman (Kahneman, 2012). Como en muchos casos el Sistema 2 es adelantado por el rápido y automático Sistema 1, el resultado es que la elicitación de creencias directas no toma en cuenta los aspectos estratégicos, proporcionando entonces peores estimaciones. Una alternativa mejor de abordar este problema resulta ser la elicitación directa de creencias sobre elementos básicos y la estimación a partir de ellas de las probabilidades adversariales finales utilizando métodos conductuales de recomposición, como los introducidos en este capítulo.

Los resultados del Capítulo 3 también contribuyen a nuestra comprensión del mecanismo de formación de creencias en situaciones adversariales. Hemos mostrado que la distribución de probabilidad de un agente sobre las acciones que elegirá un adversario, obtenida por elicitación directa, difiere de la obtenida por ARA y recomposiciones conductuales. Esa diferencia permanece incluso cuando la elicitación directa se realiza después del ejercicio de reflexión sobre cada elemento individual del problema de decisión de su adversario. Este hecho sugiere que el análisis estratégico no juega un papel relevante en el mecanismo de formación de creencias. Como consecuencia, la diferencia en el nivel de protección y seguro identificada en el Capítulo 1 puede ser el resultado de otras palancas de comportamiento más que de una reducción en la percepción de cibervulnerabilidad. De todos modos, esta pregunta permanece abierta y debe ser respondida por investigaciones futuras.

Relevancia

Los hallazgos presentados en esta disertación tienen relevancia científica, así como aplicaciones prácticas en la industria y en la formulación de políticas.

Desde el punto de vista de ***ciencia básica***, nuestros resultados contribuyen a (i) la literatura sobre el desarrollo de modelos predictivos conductuales con fundamentos psicológico más allá del enfoque de decisión perfectamente racional, con capacidad para explicar observaciones empíricas y; (ii) la literatura sobre la formación de creencias en interacciones estratégicas bajo incertidumbre. En cuanto a la primera contribución, y después de mostrar en el Capítulo 1 que los modelos de elección racional no son capaces de explicar los patrones encontrados en nuestros datos experimentales, proponemos un nuevo modelo que integra elementos de dos teorías psicológicas relevantes del cambio de comportamiento: Teoría de la Motivación a la Protección (PMT) y Teoría de Acción Planeada (TPB). La metodología de integración y calibración propuesta, Modelo de Ecuaciones Estructurales (SEM), es disruptiva y abre un nuevo enfoque para modelar datos conductuales-experimentales bajo incertidumbre con sólidos modelos teóricos de comportamiento. El SEM presentado en el capítulo 2 es el primer intento en la literatura de utilizar este método para validar la existencia de relaciones causales a partir de (i) inclusión conceptos psicológicos en el modelo como variables latentes, y (ii) medición de dichas variables latentes a través de variables observables provenientes tanto de respuestas conductuales en un experimento económico como de escalas psicométricas validadas.

Nuestros resultados contribuyen también a la comprensión del proceso de formación de creencias, más allá del supuesto conocimiento común aceptado en teoría de juegos. La combinación de un enfoque de Análisis de Riesgo Adversarial (ARA) con modelos de utilidad aleatoria para formación de creencias adversariales y la validación experimental de la precisión de dicha combinación es algo novedoso en la literatura. Este enfoque ayuda a entender los mecanismos de formación de creencias y proporcionan una alternativa práctica al supuesto de conocimiento común implícito en la teoría de juegos estándar y que respalda el concepto de Equilibrio de Nash como solución en interacciones estratégicas.

Los resultados de esta monografía contribuyen también a la ***investigación aplicada en el campo de la ciberseguridad***. El modelo SEM de adopción del ciberseguro presentado en el capítulo 2 contribuye a completar un vacío en la literatura: a pesar del creciente interés en la ciberseguridad y a la industria en torno a ella, existe una abrumadora falta de comprensión sobre los mecanismos conductuales detrás de la toma de decisiones en ciberseguridad. Nuestros modelos proporcionan conocimiento sobre los factores de comportamiento que influyen en las decisiones de ciberseguridad, incluida la adopción de medidas de protección y seguros. Por otro lado, el modelo de adopción SEM muestra que, aunque la gravedad percibida de los ciberataques afecta los niveles de ciberprotección o cobertura cibernética de un agente, la vulnerabilidad percibida a los ciberataques no es un predictor significativo de dichos niveles. Estos son hallazgos preocupantes y plantean la pregunta de cómo se construyen realmente las creencias sobre vulnerabilidad y de cómo desarrollar nuevos enfoques para mejorar las creencias sobre la probabilidad de sufrir un ciberataque intencional. Los resultados del Capítulo 3 proporcionan una respuesta a estas preguntas. Finalmente, nuestro trabajo de investigación también ha contribuido a la recopilación y difusión de datos fiables sobre la adopción del ciberseguro, obtenidos con experimentos económicos masivos en situaciones controladas (`https://www.cybeco.eu/`). En un campo de investigación donde los datos son tan escasos como es el del ciberseguro, otros investigadores pueden utilizar también estos datos en futuros proyectos.

En nuestra investigación, hemos identificado factores clave que subyacen de la toma de decisiones en torno a la ciberseguridad. El modelo que se presenta en el capítulo 2 puede utilizarse para orientar futuras intervenciones destinadas a aumentar la aceptación del ciberseguro y mejorar la ciberseguridad. Esto respalda la ***relevancia práctica*** de nuestros resultados, tanto para la formulación de políticas como para el diseño de estrategias empresariales en el sector del ciberseguro. En un contexto de formulación de políticas, mejorar la ciberseguridad y garantizar la resiliencia de los sistemas digitales es un objetivo clave, como se establece en la Estrategia Digital Europea (European Commission, 2018). Aunque previamente establecido en la literatura, nuestra investigación refuerza la evidencia empírica de que los modelos de elección racional no son capaces de explicar la adopción del ciberseguro. Este hecho motiva el desarrollo de políticas con un enfoque de economía conductual para mejorar el comportamiento en ciberseguridad. Además, la interpretación de los modelos SEM calibrados en el Capítulo 2 proporciona información concreta que puede ser utilizada para mejorar la efectividad de dichas de políticas. Las conclusiones de esta monografía también son relevantes en la industria del ciberseguro en áreas como la evaluación del ciber-riesgo o el diseño de la cartera de pólizas. Como se muestra en el Capítulo 3, los protocolos de autoevaluación de riesgos no guiados como se utilizan actualmente parecen no ser una forma efectiva de obtener estimaciones precisas de las probabilidades de que ocurran eventos estratégicos, por ejemplo, ataques intencionales de cibercriminales. Se puede obtener una estimación de riesgo más precisa a partir de un procedimiento de evaluación guiado, en el cual (i) las decisiones estratégicas adversariales se descomponen en una serie de elementos básicos; (ii) las creencias sobre cada elemento básico se obtienen de forma independiente; y (iii) las probabilidades finales se recuperan utilizando modelos ARA de recomposición conductual. Estos procedimientos podrían formar parte de las auditorías de ciberseguridad y ser utilizados en el diseño de pólizas de prima variable, ya que, como se muestra en el Capítulo 1, este enfoque puede aumentar el nivel de racionalidad del nivel de cobertura.

Limitaciones

Los resultados presentados en esta tesis no están exentos de limitaciones. En primer lugar, más allá del ciberseguro, nuestra investigación se centra en dos dimensiones de la ciberseguridad, concretamente la ciberprotección y seguridad del comportamiento online. Estas dimensiones han sido seleccionadas por su relevancia y el hecho de que permiten ilustrar características clave de la ciberseguridad. Así, la ciberprotección es observable (con los correspondientes mecanismos de auditoría) y compite con el ciberseguro por la asignación de un presupuesto limitado para la compra de productos comerciales (firewalls o antivirus frente a pólizas de seguro). Por otro lado, las decisiones sobre el comportamiento online se toman después de la adopción de la protección y el seguro, no tienen implicaciones presupuestarias y son difíciles de observar por las aseguradoras. Sin embargo, hay aspectos relevantes del ciberseguro, como la recuperación cibernética o el intercambio de información sobre ciberataques, que no han sido cubiertos en nuestra investigación.

Una segunda limitación proviene de la simplificación de los productos de protección y seguro, así como de las opciones de navegación online, efectuadas para el diseño de los experimentos económicos. Concretamente, hemos considerado que las medidas de protección solo afectan a la vulnerabilidad (probabilidad) de recibir un ataque, pero no a su gravedad ni las opciones de recuperación después del mismo. Además, consideramos que el impacto de un ataque es principalmente económico, sin tener en cuenta otros efectos como la pérdida de confianza del cliente. Además, la cartera de ciberseguros ofrecida a los sujetos constaba de solo dos pólizas diferentes, con una información sencilla y fácil de entender. Estas simplificaciones no han permitido considerar otros elementos conductuales relevantes del ciberseguro, como la carga cognitiva para comprender los diferentes productos o heurísticas de decisión que deben aplicar los potenciales compradores de seguros en condiciones más realistas.

Finalmente, y como para cualquier otro resultado validado con experimentos económicos conductuales, la validez ecológica de las conclusiones (es decir, la validez de la traducción de los resultados de los resultados experimentales al mundo real) podría considerarse como una limitación. Sin embargo, la aplicación de experimentos en economía del comportamiento en ciberseguridad es una práctica establecida en investigación (van Bavel et al., 2019; Rodríguez-Priego et al., 2020; Insua et al., 2020) y se ha utilizado con éxito para la formulación de políticas respaldadas por evidencia empírica, por ejemplo, por la Comisión Europea (van Bavel et al., 2016; Monteleone et al., 2015; van Bavel et al., 2015).

Futura investigación

Los resultados presentados en esta memoria plantean nuevas preguntas para futuras investigaciones. Un primer paso en esta dirección consitirá en el enriquecimiento de los modelos presentados en los Capítulos 1 y 2 con (i) la inclusión de dimensiones adicionales de ciberseguridad (por ejemplo, la evaluación del ciber-riesgo o la recuperación tras un ciberataque), (ii) la consideración de carteras más complejas y realistas de medidas de protección y seguros, y (iii) la aplicación de otros métodos cuantitativos para el análisis de los datos experimentales, más allá de SEM y modelos logit multinomiales. Esta línea de trabajo prevee también una calibración experimental de la función de ponderación de probabilidad. Si bien esto requerirá el diseño y la implementación de experimentos adicionales, la calibración de pesos de decisión permitirá analizar si la Teoría de la Prospectivas puede predecir los datos observacionales de adopción de ciberseguro de una manera más precisa que la Teoria de la Utilidad Esperada. Los nuevos experimentos serán implementados de forma interactiva, con la participación de sujetos reales tanto en el papel de defensor como en el de atacante.

Los resultados de los modelos conductuales calibrados en el Capítulo 2, así como los modelos enriquecidos y experimentos de economía conductual a desarrollar en el futuro, motivan otra línea de investigación enfocada al diseño y validación de *Nudges* para promover el cambio de comportamiento en ciberseguridad. Estos *Nudges*, o intervenciones dirigidos al Sistema 1 de pensamiento automático descrito en Kahneman (2012), se diseñarán con el fin de aproximar las decisiones de los agentes a las soluciones del modelo de elección racional y cumplir con los objetivos de políticas de ciberseguridad definidos por la Comisión Europea en su Estrategia Digital (European Commission, 2018).

Las conclusiones obtenidas sobre formación y elicitación de creencias plantean también nuevas cuestiones a investigar. Profundizando en los resultados de nuestro trabajo, prevemos desarrollar y validar, mediante experimentos a gran escala, nuevos modelos de recomposición conductual. Estos integrarán modelos alternativos de economía del comportamiento con técnicas ARA. En línea con los objetivos del párrafo anterior, una opción a considerar será la inclusión de la función de ponderación de probabilidades dentro del problema de decisión del atacante. También responderemos a una pregunta no resuelta en esta disertación: ¿Hacen los ataques intencionales que el defensor se sienta menos cibervulnerable que los ataques aleatorios no intencionales? Para responder a esta cuestión, prevemos la implementación de experimentos de economía conductual de elicitación de creencias, basados en la metodología del Capítulo 3 y considerando un tratamiento adicional en términos de la internacionalidad o aleatoriedad del ataque. El análisis de estos experimentos proporcionará evidencia empírica para probar si la intencionalidad realmente está afectando las creencias de vulnerabilidad o si el impacto de la intencionalidad al comportamiento de ciberseguridad (como se identifica en el Capítulo 1) sigue un camino diferente.

Finalmente, dada la falta de métodos efectivos de evaluación del riesgo cibernético en el sector del ciberseguro, nuestra futura agenda de investigación partirá de las conclusiones del Capítulo 3 para el diseño, validación y posible explotación comercial de una herramienta de evaluación del riesgo frente ataques intencionales. Esta herramienta, que se implementará como una aplicación web, se basará en métodos de recomposición conductual. La aplicación seguirá un enfoque conductual para optimizar la experiencia del usuario durante la carga de información, así como el uso de *Nudges* diseñados a medida para la presentación de los resultados de la evaluación de riesgos.

References

European Commission (2018). European commission digital strategy: A digitally transformed, user-focused and data-driven commission. `https://ec.europa.eu/info/sites/info/files/strategy/decision-making_process/documents/ec_digitalstrategy_en.pdf`.

Insua, D. R., Baylon, C., and Vila, J. (2020). *Security Risk Models for Cyber Insurance*. CRC Press.

Kahneman, D. (2012). *Think Fast and Slow*. Penguin. London.

Monteleone, S., van Bavel, R., Rodríguez-Priego, N., and Esposito, G. (2015). Nudges to privacy behaviour: Exploring an alternative approach to privacy notices. *JRC Science and Policy Report. Luxembourg, Luxembourg: Publications Office of the European Union.*

Rodríguez-Priego, N., van Bavel, R., Vila, J., and Briggs, P. (2020). Framing effects on online security behavior. *Frontiers in Psychology*, 11:2833.

van Bavel, R., Rodríguez-Priego, N., et al. (2016). Testing the effect of the cookie banners on behaviour. *JRC Technical Reports.*

van Bavel, R., Rodríguez-Priego, N., Maghiros, I., et al. (2015). Seven points to remember when conducting behavioural studies in support of eu policy-making. *JRC Scientific and Policy Reports. Luxembourg: Publications Office of the European Union.*

van Bavel, R., Rodríguez-Priego, N., Vila, J., and Briggs, P. (2019). Using protection motivation theory in the design of nudges to improve online security behavior. *International Journal of Human-Computer Studies*, 123:29–39.

La portada y contraportada han sido diseñadas usando imágenes de stories - Freepik.com