

Privacy decisions, smartphone applications and pandemics

Bachelor Thesis International Bachelor Economics and Business Economics

Name student: Theo Post

Student ID number: 470907

Supervisor: Francesco Capozza

Second assessor: Benjamin Tereick

Date final version: 26-06-2020

The views stated in this thesis are those of the author and not necessarily those of the supervisor, second assessor, Erasmus School of Economics or Erasmus University Rotterdam.

Abstract

In the wake of the COVID-19 pandemic, the Dutch government has announced plans to launch a smartphone application for tracking and tracing, in order to combat the virus. Privacy economics demonstrates individual's privacy behaviour is rather malleable, most privacy-related decisions are made irrationally. If this applies to (in)-COVID-app related decisions that involve the sharing of sensitive private information, careful attention should be paid on framing these decisions. A positive information treatment is held to test this. Furthermore, online privacy behaviour and Corona-measure adherence are measured. Results indicate people with higher incomes and those working in IT were more inclined to download such an app. Older individuals were less inclined to download a COVID-app. Individuals that are more responsible in online privacy behaviour were more inclined to download an app. No significant results were found for the effect of the treatment nor the effect of seriousness in Corona-measure adherence on the chance of downloading the application.

Table of contents

Abstract.....	2
1. Introduction.....	4
2. Theoretical framework	
2.1 Literature review.....	7
2.2 Research question and hypotheses.....	10
3. Data & Methodology	
3.1 Data collection.....	12
3.2 Summary statistics.....	14
3.3 Methodology.....	16
4. Results.....	18
5. Conclusion and Discussion.....	24
6. Bibliography.....	28
7. Appendix	
7.1 Tables and figures.....	31
7.2 Survey questions.....	39
7.3 Stata do-file.....	44

Introduction

On January 11, official Chinese state media reported the first casualty due to COVID-19, a virus that had surfaced in a local market in the last quarter of 2019. This event was followed by the cut-off of the city of Wuhan, where numerous casualties were occurring. In the months after, the virus spread worldwide. This incentivised the World Health Organisation to declare the COVID-19 virus a global pandemic (Taylor, 2020). Whilst information was rapidly gathered, governments tried implementing different policies, with countries such as France and Spain implementing strict lockdowns, a situation in which all but necessary activities are suspended, implying people were bound to stay home for significant periods of time. Other countries such as Sweden implemented rather lenient policies (Holroyd, 2020). One of the methods deployed entailed a smartphone application, such as implemented in Singapore and South Korea. On the second of March, the south Korean government published an official playbook in which the methods of dealing with the virus during the first 90 days of the outbreak were elaborated upon. Although the countries policy was rather successful, the country refrained from large-scale lockdowns. South Korea implemented, among others, smartphone applications relatively quickly, whilst being amid the first countries severely hit. One application gave people the possibility to use telemedicine, they could fill in their symptoms, being redirected to a doctor when necessary. Another application gave real-time information on availability of facemasks in several stores.

Furthermore, two smartphone apps for patient-tracking were created. The first one to see anonymised movements of confirmed patients nearby. The second application launches a notification once an individual gets within 100 meters of a location where a confirmed patient has been recently (Woodward, 2020). These features were claimed to significantly contribute to the success in containing the COVID-19 virus. Subsequently, several others governments started considering the implementation of akin apps, as the situation severed in their respective countries. In the Netherlands, calls for a similar application increased, and near the beginning of March, the Minister for Health announced the government was considering the implementation of a COVID-19 app ("Corona-app laat", 2020).

Immediately, heavy criticism came up from the side of privacy experts (Klaassen, 2020). If a South Korea-like application would be enrolled, sensitive private data would have to be shared, such as location via GPS. Privacy experts stated that the app should minimally guarantee full anonymity for its users. Moreover, such an application should be extremely well protected, due to the immense

consequences a data breach could entail. Furthermore, a guarantee for all data to be erased after the pandemic was requested. A government-administered database with numerous location data for countless citizens seems to be in contrast with the right to privacy, as recorded in the eighth article of the European Convention on Human Rights, as well as the tenth article of the Dutch Constitution. Dozens of academics warned the government in a letter, stating that the implementation of a location-tracking app could open the path to other freedom-limiting tracking apps. According to Bill Marczak, surveillance technology expert at the University of California, even with anonymised data, governments might be able to estimate people's locations. This could e.g. form a threat to journalists ("Corona-app reëel", 2020).

Economics of privacy

Economics of privacy aims to analyse the interaction between firms, individuals and policymakers when personal data is involved (Cecere et al, 2017). With the information age emerging, the topic has become more relevant than ever. Numerous papers are being published, while quantities of privacy related articles in public press are soaring (Buchanan, Joinson, Paine & Reips, 2006). When confronted with questions about the importance of privacy, respondents tend to underline its importance in overwhelming numbers. On the other hand, when actual experiments are performed in which rather low sums of money are offered in exchange to transferring extensive private information, people tend to quickly give up their privacy for small monetary gains (Athey, Catalini & Tucker, 2017). Moreover, research in which psychology and economics of privacy are combined, indicates that the framing of a question is of paramount importance to the answer ultimately given (Acquisti, Brandimarte & Loewenstein, 2015). These kind of results indicate that privacy decisions are mostly not made rational, due to their malleable nature. This is an indication that governments should pay careful attention to the way questions about willingness to use COVID-19 apps are posed. Furthermore, in-app decisions should be carefully formulated and considered to enable objective, full-informed decisions. It could even be that nudging individuals towards privacy-protecting decision should be encouraged.

The governmental plans to launch a COVID-19 app therefore seriously contain the risk of inflicting damage to user's privacy. As shortly considered above, sundry research has been done on the economics of privacy, especially the psychological phenomena that come with it. With the COVID-19 situation being a 2020 phenomenon, privacy concerns about application have been outed widely. On the other hand, whether the malleability of people in the field of privacy decisions also applies to

COVID-19 app related decisions, has not yet been researched. This makes it both socially and scientifically relevant.

In the theoretic framework, several papers will be discussed, having in common their focus on the economics of privacy. This will provide an extensive insight in the relevant psychological phenomena that tend to occur commonly, and especially the risks that occur with privacy decisions. These findings will be combined with the ostensible outlook of a COVID-19 application. This brings up the research questions along with its hypotheses, whether a positive information treatment positively nudges people's inclination to download such an application. Thereafter the method of research will be elaborated on, an online survey. Summary statistics will be provided, after which results will be discussed, ending with a conclusion and discussion.

Theoretical framework

In this paper, 'privacy' refers to the concept of information privacy, defined as 'concerns with controlling whether and how personal data can be gathered, stored, processed and disseminated' (Posner, 1981). Moreover, privacy attitudes can be distinguished, defined by Acquisti et al., (2015) as 'concerns for the appraisal of specific privacy behaviours'. Furthermore, privacy preferences could be divided between intrinsic preference, a normative and relatively context-independent privacy valuation, as well as instrumental preference, a context-dependent utility change (economic loss) of disclosing one's private information (Lin, 2019). Revealed preferences are an indicator for actual privacy behaviour, 'the action to set control over data', as defined by Acquisti et al., (2015). The endowment effect is defined as individuals valuing an owned object higher than the market value, appointed to emotional attachment that people can feel to products (Ganti, 2020). Lastly, the notion of malleability in the field of privacy, describes the fact that people are easily influenced in what information and how much of it they decide to share (Acquisti et al., 2013).

With over 90% of all data ever collected being collected in the last two years, data collection is increasing at an unprecedented rate (Petrov, 2020). With it comes an increase in opportunities. One example of a clear benefit that comes with immeasurable amounts of search data, lies in the medical spectre. Yet unknown drugs interactions or effects that were not known to the medical world can be discovered by analysing search data. E.g., if numerous internet searches are performed on 'legs hurting after COVID-19 vaccination', this gives medics an indication there presumably exists a causation. On the other hand, this wealth of information entails risks. Some risks are clearer, such as leakage. Furthermore, blackmailing opportunities that arise especially after data leakage, are an increasing issue. More unknown to general public are the vast commercial options that are becoming available to corporations, which is the main reason companies appear to be offering free products or reductions in exchange for personal information. The more information, the more specific targeting can be. This could, hypothetically, lead to the situation where a company knows exactly with which words and at what moment to present their product online, to have the highest chance possible of their product being bought. These kind of commercial options could lead to social and economic discrimination, as those individuals with the lowest resistance to these forms of marketing will purchase substantially more products than they can afford (Acquisti, Taylor & Wagman, 2016). These results are some examples of the underlying framework of the discussion on privacy economics.

Individual's manipulability has been shown to be substantial. Acquisti, John & Loewenstein (2011) performed research in which respondents were displayed various websites: a rather professional,

government-like website, as well as a non-professional, clickbait-like site that was intentionally designed to appear less trustworthy. On both websites, participants were asked to provide privacy-related information. A survey was presented about their privacy behaviour, again in a rather professional and rather amateurish design. Contrasting the hypothesis, people admitted to having conducted risky privacy behaviour more often on the rather unprofessional website. This indicates rather irrational privacy behaviour. The default would indicate it is a much safer decision, data-wise, to share sensitive data with a professional party compared to an unbecoming party that might not have arranged trustworthy security. When considering privacy decisions that arise with a COVID-19 application, design could therefore be of paramount importance. These results are related to the framing effect: people's decisions are influenced by whether options are presented with positive or negative connotations (Kahneman & Tversky, 1974).

In 2010, GameStation included a clause in a license agreement for a videogame that stated the user's soul would become GameStation's property. In one day, over 7500 individuals agreed to this (Smith, 2010), providing another example of a massive public tendency to not read license agreements at all. Research by Cakebread (2017) looked into terms and conditions. Results were in line with the GameStation example, with over 90% of individuals not reading terms and conditions at all before agreeing to them. This would even be virtually impossible, given that on average one would need 76 working days to read all privacy policies one agrees to (Wagstaff, 2012). These results, considering license agreements, imply that merely including in-depth explanation of the exact conditions one agrees upon when downloading a COVID-19 app, could be insufficient when it comes to enabling individuals to be fully aware of their decisions.

As discussed in the introduction, when asked for their willingness to relinquish privacy, or merely the importance that is attributed to privacy, people tend to give answers that implicate serious concern and responsible behaviour. Revealed preferences, as defined earlier, are seen as more relevant. Revealed preferences entails retrieving behaviour by focussing on actual behaviour. Actual behaviour provides more relevant information than pure normative statements, that do not have consequences. In an experiment held by Acquisti, John & Loewenstein (2013), female shoppers were - after leaving a shop - offered a \$12 dollar gift card that allowed the researchers to trace their activities. The other option offered comprised of a \$10 anonymous gift card, which guaranteed no data tracking at all. Some subjects got to choose between the two to check for order effects. Other participants were provided with one of the two cards and were offered to switch cards to control for the endowment effect and willingness to accept & to pay. The choice between the two gift cards could also be interpreted as an implicit choice for those endowed with \$10 to sell their future purchase information

(privacy) to the researchers for \$2. Equally, those endowed with the \$12 gift card were given the opportunity to buy their singular transaction privacy for \$2. Participants tended to adhere to the gift card they were initially endowed with. These results underline earlier research that points to the strength of default settings. Furthermore, significant order effects were found, implying the order of choices has importance too. Moreover, a significantly higher proportion of participants decided to keep the \$12 gift card compared to the \$10 card. This implies that in supplement to order and endowment effects, an increase of only \$2 in card value was sufficient for half of the subjects to relinquish their transaction privacy in order to obtain the \$12 gift card. If small pecuniary amounts are sufficient for people to relinquish privacy, it seems plausible that people would share private information even easier with an app that could help them in preventing (serious) illness. This presumption is based on the idea that people would be willing to pay at least \$2 to not become ill.

The malleability of privacy decisions can even further be deducted from critical decisions. Johnson & Goldstein (2003) found results that critical life decisions, such as whether to be an organ donor or how much to save for retirement are heavily dependent on default settings. A retirement savings programme in which a standard percentage of income was diverted to a special retirement account failed to lure significant proportions of employees into participating as long as it stayed an opt-in choice. When the decision to relinquish monthly income, something that was clearly immediately felt by participants, was changed to an opt-out decision, a significant increase in participants occurred. These results were attributed to people's inherent laziness, only pushed into action with sufficient incentives, a requirement a monthly salary loss did not seem to meet. The manner in which choices are presented is highly influential on eventual outcomes. For this reason, the Dutch parliament voted in favour of a law that changed the default setting for organ donorship into donation, leading to an increase of 15% in donorship (Bouma, 2020). This again implies even extremely sensitive and critical decisions are influenced by simple default settings. Would this also apply to a COVID-19 app, which seems plausible, default-setting precautions should be made, based on the idea of libertarian paternalism (Sunstein, R., & Thaler, R., 2003). Although the choices (in e.g. default options) stay the same, a small adjustment to the presentation is made. This would be done by the design-team of the application. In this case, the default option would be the privacy-wise safer option. The fact another organ makes this decision makes it paternalistic. The libertarian part is explained by the fact the choice still exists.

Athey, Catalini & Tucker (2017) considered 'Notice and Choice' policies. In especially the United States, online privacy protection focus has been on providing users with information (e.g. terms and conditions). Subsequently, individuals were being asked to agree or disagree, based on the idea that

sufficient transparency and choices induce more responsible privacy decisions. According to the authors, this is not sufficient to guarantee online privacy and fully rational decisions. At first, small incentives, albeit pecuniary or not, tend to steer people into a certain direction. Second, minor navigation costs significantly influence privacy decisions. An example could be a simple hyperlink to a different page where certain terms are elaborated, this could be a major deterrent for people to fully read and therefore consider their privacy interests. Furthermore, little reassurances about privacy significantly enhance the chance of people agreeing to certain privacy policies, independent of the actual value of these kind of reassurances. Schudy & Utikal (2017) performed research to see whether the amount of data as well as the number of recipients influences willingness to share. The more recipients, the less willing individuals are to share data. Social proximity to the party with whom the data is shared does not influence willingness. Being less inclined to share data when more receiving parties are involved, is in contrast to people's general tendency to continuously agree with terms & conditions in which the right to transfer data to third parties is reserved. In accordance with the previously explained results concerning GameStation, relying on elaborate terms and conditions could be insufficient in COVID-19 app privacy, if people ultimately do not read them. Other ways of fully informing people should be considered, if simple terms & conditions do not suffice.

Research question and hypotheses

As previous papers considered in the paragraphs above confirm, people are easily influenced when it comes to privacy decisions. Privacy policies that only rely on informing the individual are unlikely to sufficiently protect individuals from their own malleability. The advantages of a possibly life-saving COVID-19 app appear to be perspicuous. Nevertheless, these potentially extensive benefits, should not crowd-out legitimate privacy considerations, based on codified legal principles. Whether considering in-app decisions about which information to share, as well as the decision to download a COVID-19 application at all, careful attention should be paid. The framing of questions, the presence of incentives and navigation costs individuals are confronted with influence privacy decisions significantly. Although research when it comes to government related issues such as organ donorship has been performed, implying significant malleability, government-developed application privacy decisions has not yet been researched. Furthermore, the situation of such an application being rolled out in times of a global pandemic, including potentially life-saving features comes in new too. This makes it relevant to research whether the privacy decision effects as described previously apply to this specific topic too, leading to the main research question:

Does an positive information treatment nudge people's willingness towards being inclined to download a COVID-19 smartphone application?

Sunstein & Thaler (2008) describe several heuristics that can be used to nudge people. One of the heuristics described is framing, in which information is presented in a certain way that alters the way it is perceived. By means of accentuating positive information about a COVID-19 app, it is tried to nudge people towards a more positive attitude towards the app. Due to the malleability of privacy decisions and lack of insight in possible consequences of a privacy loss, it seems plausible that an information treatment will have significant effects. This leads to the first hypothesis:

H1: A positive information treatment positively influences people's attitude towards a COVID-19 smartphone app.

Individual's Corona-concern will be measured (as explained in the methodology). The official reason any government would want to implement a COVID-19 app is mainly due to its potential support in fighting the virus. This implies that downloading such an app could actually increase physical safety, although privacy concerns arise. This makes it relevant to examine how the two measured statistics correlate, leading to the second hypothesis:

H2: People who adhere better to Corona measures show a higher willingness to download the app.

Furthermore, general privacy online behaviour will be measured. It seems plausible that those more concerned about privacy, therefore displaying more responsible behaviour, will be more worried about a COVID-19 app and therefore less inclined to download it, leading to the third hypothesis:

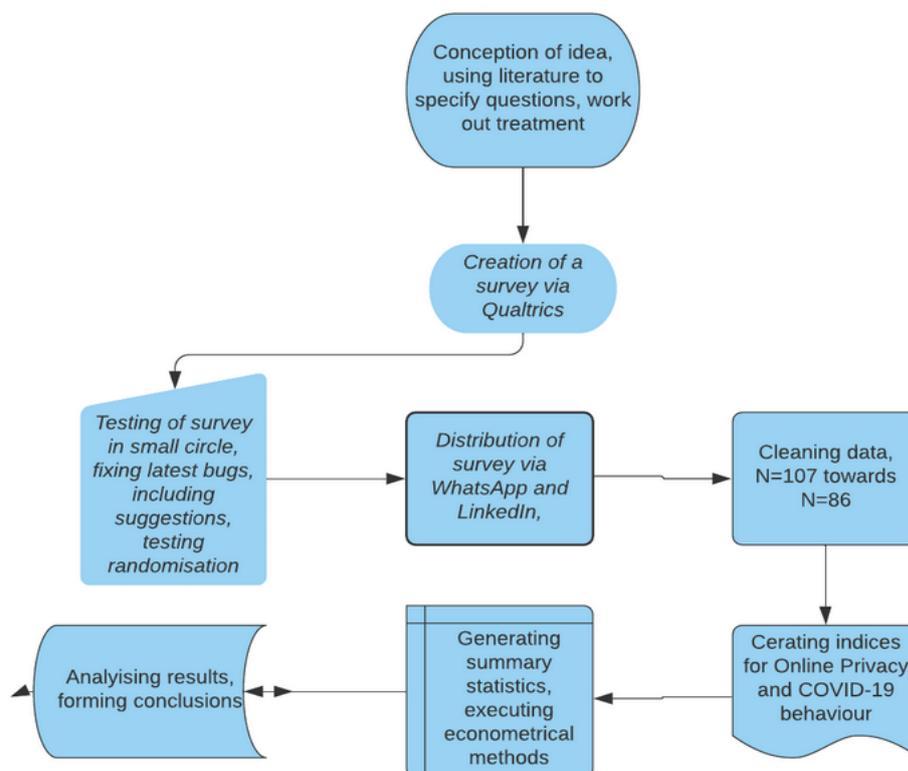
H3: Individuals that exhibit online privacy-protecting behaviour, will be less likely to download the app.

Data & Methodology

Data on privacy behaviour, COVID-19 behaviour and willingness to download a COVID-19 app were collected by means of a survey. The Dutch government originally intended to publish their final decision on whether to launch a COVID-19 app around halfway May. This meant that in order to prevent publicity around the final decision from influencing respondent's answers, the survey had to be held ahead of that decision, which succeeded. The experimental design as described in the following paragraphs, is summarised in Flow Diagram 1.1.

Flow diagram 1.1

Experimental design



Data collection

The survey was translated to Dutch, to ensure that respondents could pick their language of preference and understood the questions completely. Firstly, respondents were asked whether they agreed with cookies. Even though this was not actually real, implying there would be no real consequences in terms of information gathering to their responses, the question was made as real as possible. A hyperlink to the Qualtrics cookie policy was included. Due to Qualtrics being the

programme in which the survey was made, the question appeared more realistic. Thereafter, baseline questions about age, gender, sector, education and gross income. That last question entailed an answer option 'prefer not to say', an extra indication for ease of information disclosure by the individual. After these questions, Corona-behaviour questions were presented. Frequency of supermarket visits, number of times that individuals met non-household members in the week before, handwashing behaviour, news checking and self-reported chance of becoming seriously ill due to COVID-19 were asked. Furthermore, online privacy behaviour was demanded. General cookie web-browsing behaviour, license agreement conduct, unsecured network behaviour, number of different passwords used, usage of LinkedIn, Twitter, Facebook and Instagram and lastly online video-calling software used were asked. The question about number of passwords used also included an answer option 'prefer not to share'. Based on their level of (end-to-end)-encryption, all video calling software programmes were ranked. After these questions, the information treatment was presented. The questions were formulated as neutral as possible. Words such as privacy, security and data were avoided as much as possible. This is due to the cognitive biases that can result from the mere mention of concepts. An example of this was given in the theoretical framework, with the inclusion of the word privacy leading to more privacy-concern than a similar request without the word. For those receiving the positive information treatment, asking about online privacy and COVID-19 behaviour after the treatment could mean that their responses are influenced by being positively primed towards an application. For example, respondents would first read how South-Korea decreased its casualties rate. Afterwards, they would be asked for their level of responsibility concerning COVID-19 behaviour. This extra reminder could result in less trustworthy responses, as people tend to lie even in anonymous surveys (Schudy & Utikal, 2017). On the other hand, by the time they received the treatment, people had already answered several privacy and COVID-19-related questions. Therefore, eventual responses to the self-reported chance of downloading an application are likely to be biased either. This risk of influence would threaten all subsequent questions in case the treatment would be held before the baseline questions. Therefore, the order as chosen, with the treatment and subsequent smartphone app question towards the end, presumably led to the least influence as possible in this survey setting.

This treatment was automatically randomised by Qualtrics. Half of the respondents received a neutral information treatment, a short message about Zoom's increasing number of users. The other half received a short, positive text about how South Korea managed to flatten the curve (diminish the number of COVID-19 cases) by means of a smartphone app. This treatment was followed by a question about individual's estimated chance of downloading a COVID-19 app if such an app would be launched by the government. The last question was optional: respondents were informed that if they wanted to participate in a €10 gift card raffle, which was originally included to incentivise respondents, they

were required to fill in their email address. Another three text boxes were included for name, phone number and address. In smaller but clearly visible words, the page ingeminated twice that filling in these other three fields was not necessary to participate in the raffle, a measure that was included to measure both alertness and level of carefulness.

As said, priming by using words as privacy, personal data and consent was avoided as much as possible, in order to prevent the respondents from being influenced. Whereas people normatively express their concern for privacy, actual behaviour tends to lean towards patterns in which sensitive information is relinquished for low-hanging pecuniary amounts (Athey et al, 2017). Therefore, the survey questions were aimed to focus on measurable actual behaviour predominantly, to get more reliable information, such as actual number of passwords used or times meeting up with non-household individuals in the week prior to the survey. A paper by Buchanan et al (2006) on measuring internet privacy was partly used for the structure of the questions. For the main question, whether people would be inclined to download a COVID-19 app, revealed preferences could not be applied. Clearly, the application did not exist yet at the time the survey was spread. Due to the novelty of the concept, a comparable measure did not exist.

The survey was spread via LinkedIn and WhatsApp predominantly. Screenshots of the precise survey questions can be found in the attachment, page 37. Questions were asked in a convenient way for respondents. This meant that some questions generated numerical data immediately, such as percentage of the time that people wash their hands when coming home. Other questions were not provided with numerical answers. E.g., Facebook behaviour entailed the choices of not having a Facebook account, having a private account or having a public account. COVID-19 news behaviour included answers such as 'never', 'once per week', etc. These kind of answers had to be transferred to numerical values first. For this, a logic was followed in which the most 'responsible' behaviour, in terms of online privacy behaviour or COVID-19 behaviour was rated 100. In questions with in total six answers options, such as 'never, sometimes, regularly, often, quite often and always', the answers would be given respective values of '0, 20, 40, 60, 80, 100'.

Summary Statistics

The summary statistics as seen in Table 1.4, show that most participants had obtained university-level education, 51 individuals. Only one participant had started with a PhD. This education level of university can be seen back in the occupation categories, with 46 participants identifying their main occupation as student. Even though several categories were given, 15 participants stated they had non-mentioned, other occupations. 43 Participants indicated their gender to be female, with male

equally making up for 43 participants. A little more than a quarter of the participants denied cookies in Qualtrics, while only 5% refused to share their income data. Participants indicated that on average, one third of all internet cookies are denied by them, contrasting with the smaller percentage that denied cookies in the survey. This is in line with the ‘Privacy Paradox’, as described by Athey, Catalini and Tucker (2017). People estimate their (online) privacy behaviour to be responsible, but their actual behaviour shows different results. Participants indicated they generally read license agreements 8% of the time, in line with findings described in the literature review that state over 90% do not read license agreements at all. Of the social media networks, Twitter was the least popular. 47% Of participants tended to use video-calling software such as Zoom or Houseparty, leading to a low ‘safety-score’ in the field of video-calling software. 23% Of participants indicated they wanted to participate in the raffle, leading to a high privacy score for that question. Even though COVID-measures were fully prevailing by the time the survey was held, participants had met up with almost 2 non-household members in the week prior. Their supermarket visits were also little more than 2. One third of participants expected to become severely ill due to COVID-19 in the coming months, which is not in line with the percentage that actually suffers substantial consequences when infected. In line with the results for occupation and education level, the surveyed population had a low mean age of almost 27 years. Eventually, the average chance of downloading a COVID-19 was estimated to be 44%, a pooled average for both groups.

Table 1.4

*Summary statistics, all variables entail 86 observations. **

Variable name	Mean	Standard deviation	Minimum value	Maximum value
Cookies survey	17.44	38.17	0	100
Income privacy	4.65	21.18	0	100
Browsing cookie acceptance	34.64	29.13	0	100
License agreements approval	7.97	15.07	0	73
Free-Wi-Fi browsing	41.39	32.39	0	100
Number of passwords	50.70	30.09	10	100
LinkedIn account	34.42	45.67	0	100

Facebook Account	64.19	34.49	0	100
Twitter account	72.79	42.50	0	100
Instagram account	63.14	33.83	0	100
Videocall software	31.86	35.60	0	100
Gift card	77.03	39.34	0	100
<hr/>				
Washing hands	78.63	28.80	0	100
Self-reported COVID-19 risk	29.91	21.79	0	90
Supermarket visits	2.31	1.82	0	8
COVID-19 news	5.30	3.03	0	14
Meeting non-household individuals	1.78	1.38	0	4
<hr/>				
Age	26.81	14.08	12	86
<hr/>				
Download COVID-19 app	44.14	32.05	0	100

*All variables with range 0-100 imply 100 is the uttermost responsive COVID-19 / privacy behaviour. Supermarket is measured in supermarket visits per week, news checking in times of checking COVID-19 news per week, meeting non-household member in appointment per week and age in absolute age. Income privacy implies whether people kept their income private or shared it, whereas income means absolute income.

Methodology

Both online privacy behaviour and Corona-behaviour were asked via diverse questions. In order to be able to make calculations, an index score on COVID-19 behaviour and online privacy behaviour was required. The related questions had to be merged in an index, based on their respective weights. Firstly, the aim was to use factor analysis to group the variables based on their degree of correlation. Bartlett's test of sphericity was performed, leading to significant results. However, according to Kaiser

(1974), the Kaiser-Meyer-Olkin statistic should have a value of at least 0.7 to obtain useful results. The privacy score index resulted in a KMO measure of 0.49 for Corona-behaviour, and 0.55 for online privacy behaviour. Therefore, factor analysis could not be used. The sample simply did not contain sufficient data for factor analysis. Therefore, it was decided to standardise the variables, meaning that the mean answer for all observations was subtracted from each observation, after which the resulting number was divided by the standard deviation. With these standardised statistics, an average was created in order to obtain indices. Furthermore, a robustness check was conducted to evaluate the way the indices are constructed. Firstly, standardized regression coefficients were obtained by regressing the intention to download a COVID-19 app on the privacy behaviour components and, separately, COVID-19 behaviour. Secondly, the absolute values of the standardized coefficients were used to construct the weights for online privacy and Corona behaviour. Randomisation was automatically applied by Qualtrics, but in order to test whether this randomisation was successful, a nearest neighbour matching test was held. Furthermore, a summary of covariates was created, to get an insight in the ratio between the means of the covariates in both the treatment and control groups.

Results

As mentioned in the methodology, in order to test for the randomisation of the treatment and control groups, a nearest neighbour matching test was conducted. This resulted in a positive effect of treatment on downloading a COVID-19 application, but this result was insignificant. Furthermore, a t-test was performed to compare the differences in means between the two treatments for the chance of downloading a COVID-19 app. As both the control and treatment group included more than 30 respondents, a t-test could be performed. This resulted in a p-value of .82, meaning the null hypothesis of no statistically significant difference could not be rejected. Even though the assumptions of the absence of outliers and independence were met, a Shapiro Wilk test resulted in a p-value of .02, implying the Corona app data was not normally distributed. Another method that was applied to check for randomisation was the creation of a summary statistics table of covariates. In this Table 4.1, as seen in the appendix, the ratios between the means for all covariates of the treatment group are compared to the control group. This led to results close to 1, implying successful randomisation, for: age, gender, income categories €10.000-€20000 and €20.000-€30.000, working in the food industry, in transport, being a student and the category of other occupations. Furthermore, having university as highest education was also successfully randomised. The income categories of €30.000-€50.000, more than €70.000 and 'prefer not to say' were not successfully randomised, neither were occupations in health and IT, the education levels of high school and intermediate vocational education. The index-scores for online privacy and Corona-behaviour, implying individual's pre-treatment behaviour in terms of Corona-behaviour and online privacy behaviour were unsuccessfully randomised as well.

In order to test the first hypothesis, *A positive information treatment positively influences people's attitude towards a COVID-19 smartphone app*, an OLS regression of the self-reported chance of downloading a COVID-19 app on all control variables and the treatment was performed. Firstly, the assumptions of normality, homoscedasticity and multicollinearity had to be tested. For the normality assumption, a Jarque-Berg test was executed, leading to a p-value of .24, implying the null hypothesis of normally distributed residuals was not rejected. To test for the data being free of heteroskedasticity, a Breusch-Pagan test was performed, leading to a p-value of .38, implying the null hypothesis of constant variance in the data was not rejected. In order to check for multicollinearity, a Variance Inflation Factor-test was executed, leading to problematic amounts of collinearity (higher than 5, James (2017)) for occupation categories student (10.41) and other occupations (6.51). This is not a problem, as these are dummy indicator variables, of a categorical variable with eight categories. Due to the low proportion of occupations in the reference category, the other occupation indicator

dummies necessarily have high variance inflation factors. This implies all assumptions were met. The regression resulted in a treatment coefficient of 5.98, implying a substantial positive treatment effect. Nevertheless, this was insignificant at a 10% significance level. At a 5% significance level, age displayed a negative coefficient of -.61 on downloading a COVID-app, implying younger people are more inclined to download a COVID-app. At a significance level of 5%, working in IT showed a coefficient of 37.71 implying those confronted with technology on a daily basis are more inclined to download a COVID-app. At a significance level of 1%, having an income between €50.000 and €70.000 as well as more than €70.000 resulted in coefficients of respectively 68.27 and 57.08, implying those with higher incomes were more inclined to download a COVID-app. Furthermore, at a significance level of 1%, having finished a PhD demonstrated a negative coefficient of -31.89, implying those with highest education are less likely to download a COVID-app. As there was only one respondent indicating to have finished a PhD, this result cannot lead to any conclusions. The results of this regression can be seen in Table 3.1

Table 3.1

Regression of the chance of downloading a COVID-19 app on treatment, age, income, education and occupation, 86 observations

Variable	Regression Coefficient
Gender (female)	8.27 (8.61)
Age	-.61** (.30)
Education	
High school	-23.14 (15.72)
Intermediate vocational education	-11.45 (15.19)
University	-12.06 (13.54)
PhD	-31.89** (18.69)
Occupation	

Education	9.51 (20.78)
Food, Restaurants	2.75 (23.29)
Health	8.31 (18.08)
IT	37.71** (16.16)
Transport	-17.37 (18.88)
Student	18.00 (13.89)
Other occupations	14.20 (14.73)

Income	
€10.000-€20.000	-7.33 (10.38)
€20.000-€30.000	13.49 (17.58)
€30.000-€50.000	2.01 (13.56)
€50.000-€70.000	68.27*** (12.20)
More than €70.000	57.08*** (15.38)
Prefer not to say	4.55 (19.69)

Treatment (positive message)	5.98 (7.69)
Constant	47.34**

*p-value < 0.1, **p-value < 0.05, *** p-value < 0.01, (F(19, 65) = 1138.56, p = .0.00, R2 = .28)

For both the second and third hypotheses, the indices as mentioned in the methodology were used to run a regression with “Chance of downloading a COVID-19 app” as dependent variable, for both privacy behaviour and COVID-19 behaviour. In order to test the second hypothesis, *People who adhere better to Corona measures show a higher willingness to download the app*, an OLS regression of chance of downloading a COVID-19 app on all control variables and Corona-measure adherence was performed. Firstly, the assumptions of normality, homoscedasticity and multicollinearity had to be tested. For the normality assumption, a Jarque-Berg test was executed, leading to a p-value of .23, implying the null hypothesis of normally distributed residuals was not rejected. To test for the data being free of heteroskedasticity, a Breusch-Pagan test was performed, leading to a p-value of .41, implying the null hypothesis of constant variance in the data was not rejected. In order to check for multicollinearity, a Variance Inflation Factor-test was executed, leading to problematic amounts of collinearity (higher than 5, James (2017)) for occupation categories student (10.68) and other occupations (5.92). Again, as explained in the previous paragraph, this is not a problem. This implies all assumptions were met. The regression resulted in a coefficient of -2.59, implying a substantial negative correlation between seriousness in Corona-behaviour and downloading a COVID-19 app, in contrast to the hypothesis.. Nevertheless, this was insignificant. At a significance level of 5%, working in IT showed a coefficient of 34.48 implying those confronted with technology on a daily basis are more inclined to download a COVID-app. At a significance level of 10%, having an income between €50.000 and €70.000 as well as more than €70.000 resulted in coefficients of respectively 59.91 and 49.38, implying those with higher incomes were more inclined to download a COVID-app. The results of this regression can be seen in Table 3.2 in the appendix.

In order to test the third hypothesis, *Individuals that exhibit online privacy-protecting behaviour, will be less likely to download the app.*, an OLS regression of chance of downloading a COVID-19 app on all control variables and online privacy behaviour was performed. Firstly, the assumptions of normality, homoscedasticity and multicollinearity had to be tested. For the normality assumption, a Jarque-Berg test was executed, leading to a p-value of .43, implying the null hypothesis of normally distributed residuals was not rejected. To test for the data being free of heteroskedasticity, a Breusch-Pagan test was performed, leading to a p-value of .24, implying the null hypothesis of constant variance in the data was not rejected. The VIF-test led to high values for student (10.23) and other occupations (5.92). Referring to the paragraph concerning hypothesis 1, this is again not problematic. This implies, again,

that all assumptions were met. The regression resulted in a coefficient of 19.33, implying a substantial positive correlation between seriousness in online privacy behaviour and downloading a COVID-19 app. This was, again, in contrast to the hypothesis. This result was significant at a 1% significance level. At a significance level of 5%, working in IT showed a coefficient of 28.77 implying those confronted with technology on a daily basis are more inclined to download a COVID-app. Furthermore, at the same significance level of 5%, having finished a PhD resulted in a negative coefficient of -30.46, but this is only based on a singular individual. The results of this regression can be seen in Table 3.3 in the appendix.

For both indices, Corona-behaviour and online privacy behaviour, a robustness check was performed. For the normal indices, each standardised variable received a weight of $1/n$, creating an average score. As an alternative method, a regression was held to calculate the beta, providing standardised beta coefficients instead of confidence intervals, as discussed in the methodology. The absolute value of these betas was used to attribute weights to the standardised values. For COVID-19 behaviour, the results as seen in Table 3.2, are quite similar for both regression. This is despite of the fact that different indices were used. The similarity implicate a good sign in terms of robustness. For the other regression as seen in Table 3.3, only the question on license agreement approval showed noticeable differences, but this is presumably due to the small sample size, implying the creation of the index was very sensitive to the weights chosen for this variable.

Furthermore, on the ninth of June, three Dutch universities published results on people's willingness to download a COVID-19 application. The results indicated one third of the respondents were absolutely willing to download it if launched, one third were doubting and the last third would absolutely not download it. The survey used in this thesis asked for the chance of downloading such an application in percentage. Three categories were created with the COVID-19 app variable. Those indicating a percentage between 0% and 33.33% were considered as not willing to download the app. Those indicating 33.33% till 66.66% chance were considered as doubting about whether they would download an app. The highest ranking category, 66.67% and above were assumed to certainly download the app if launched. These three categories resulted in a ratio of 34/26/26 in absolute respondents. This implies a ratio of 40%, 30% and 30%. These results are partially in line with Delft University's results (Collewet et al, 2020).

Although a positive coefficient of almost 6% points resulted from the treatment, this result was not significant. It can therefore not be concluded from this experiment that an information treatment easily nudges people towards an answer when it comes to privacy decisions, contrasting the results described in the literature review implying small information provisions can heavily influence privacy

decisions made. The first hypothesis is therefore rejected. Furthermore, the theoretical framework lead towards the expectation that those more concerned with Corona, based on their revealed preferences, should be more inclined to download a COVID-19 application, as it is potentially life saving, given the results in South Korea. Contrary to this expectation and hypothesis, the normal regression found a small negative coefficient of -1 for more careful COVID-19 behaviour. This was countered by the beta-adjusted regression, that resulted in a positive coefficient of almost 32, implying those more concerned with Corona are indeed more inclined to download a COVID-19 app. Both results were insignificant, so nothing can be concluded from it. The second hypothesis is therefore rejected. When it comes down to the last hypothesis of online privacy behaviour, a positive coefficient of almost 19 percentage points resulted from the standardised regression, significant at a 1% significance level. This is in contrast with the hypothesis, as it was expected those with more responsible online privacy behaviour would be less inclined to download a possibly privacy-harming application, but the contrary appears to be true from the regression. Therefore, the third hypothesis is also rejected.

Conclusion & Discussion

A sudden pandemic that rampaged the world at a scale that had not been seen in decades brought dozens of governments to rapidly lay out options to combat the virus. Some political leaders demanded their citizens to indulge in a complete lock-down of normal life for months, whereas other countries implemented more lenient measures. One of the more astute and ground-breaking measures implemented entailed a smartphone application. Varying by country, it gave participating citizens the ability to avoid certain places or get in touch with medical personnel quickly. Some of the applications that were eventually applied in South Korea required extensive data, such as GPS-locations. When the Dutch government, among others, started considering the implementation of a similar application, privacy experts were quick on their feet to provide extensive criticism. Risks involved with governmental mass collection of sensitive data exist on several spectres.

Literature on privacy economics has been describing this combination of economics and psychology for years, resulting in numerous papers. Most results point to the fact that irrational behaviour prevails when considering the field of (online) privacy. People state they are unwilling to easily relinquish their privacy. On the other hand, their actual, revealed preferences indicate it is extremely easy for companies and governments to extract information and influence privacy-relinquishing decisions. Therefore, this thesis performed a survey to measure people's online privacy behaviour, Corona-measure adherence as well as their sensitivity to an information treatment. If people are also heavily malleable by the manner in which an application is presented, governments should pay careful attention to framing, to adhere to legal obligations on privacy, even in a possibly life-saving app. These measures were combined with people's self-reported chance of downloading a Corona smartphone application. The thesis hypothesised a positive information treatment would heavily influence people's willingness to download such an application, but this main expectation could not be concluded from the results, and was therefore rejected. Furthermore, the assumption that individuals that adhere better to Corona-measures would be more inclined to download a potentially lifesaving application could not be accepted either. Regression analysis resulted in a negative coefficient, but this was not significant. The third hypothesis indicated that those more concerned about online privacy, based on their (revealed) online privacy behaviour, would be less inclined to download a privacy-reducing application. In contrary, significant results implicated those people are more inclined to download such an application. This implies that being careful when it comes to social media, license agreements, providing information and cookies does not imply heavy scepticism on a smartphone app that has been criticised for its possibly significant privacy-harming consequences. It could be that although these people are generally more concerned about privacy, they assume that a government

service will take into account privacy sufficiently. Another explanation could be that the certainty they want in life given their behaviour in terms of online privacy, is something they pursue in other fields to, especially health-related choices. On the other hand, the chance of downloading a COVID-19 application was measured by self-reported chance. This has limitations, as people tend to self-report different results compared to their actual behaviour, as previously discussed in the theoretical framework. The combination of a relatively small sample and unrepresentative answers could therefore also be an explanation for the contra-intuitive result.

Furthermore, unexpected significant results indicate IT workers as well as younger people, and even those with higher salaries were more inclined to download a COVID-19 app. This could possibly be due to the fact that those working in IT are more familiar with the workings of applications and systems such as GPS, and experience greater trust towards these technologies. Moreover, it could be that younger people are more inclined to download an application in any situation, due to the fact that they tend to use their phones more often in general, creating an opener mind towards digital solutions. For those with higher incomes, a possible explanation could be their well-paying jobs (in the centre of society) enable them to frequently get a grasp of technology, increasing trust. Furthermore, it could be that their fear of becoming even slightly sick is higher due to increased cost when not being able to go to work.

This thesis has numerous limitations. At first, none of the hypotheses could be accepted, even though they were largely based on various literature. It seems implausible that diverse literature found results that are credibly questioned by a survey with less than 100 participants.. The small size of a mere 86 participants makes it hard to infer any trustworthy conclusions from the significant results that were found, even though the regression indicate some effects. Furthermore, when it comes to the treatment randomisation, quite some characteristics were not successfully randomised, given the summary statistics. In general, the quality of the randomisation can be doubted due to the results from the t-test as well as nearest neighbour matching. Furthermore, the creation of indices has its limitations. With answer options jumping from 'less than once per week' to 'more than 3 times per week', a choice had to be made between artificially assuming values in hindsight, or displaying dizzying numbers of answer options leading to significantly lower response rates. Eventually, answers such as 'more than 3 times per week' were interpreted as 4 times per week, even though it could as well be 5 e.g. Some questions were formulated in this imprecise way due to the benefits it brings for survey completion rates as well as comprehensibility of questions posed. On the other hand, the necessary randomness element that comes with it limits the trustworthiness of the results. Furthermore, some questions were asked in a hypothetical formulation, such as: how do you estimate the chance you

would download a COVID-19 application? Due to the difference between normative statements and actual, revealed preferences, this inevitably creates partly unrealistic statistics. Even the questions that were asked in terms of revealed preferences have their limits. Choices made can be subject to counter-preferential choices. This could be due to mistakes, such as inattentiveness or false beliefs. It could also be due to weakness of will. Furthermore, choices are influenced by beliefs and expectations. Furthermore, even though relatively balanced between treatment and control, the results were effectively largely filled in by young students. This creates a bias that makes it hard to generalise the significant results that could be found to the whole population. Moreover, except for the positive, steering treatment for half of the participants, no elaborate information was given on the potential outlook of a COVID-19 application. The way such an app eventually works, which functions are included and how they are protected, is rather important to provide the nuance on which individuals eventually base their decision to download it. This nuance lacked in the question as posed in the survey used. This implies that, mainly, the results found cannot be generalised to the topic and population in general due to their limitations.

Future research could minimally try to include a bigger sample pool. Furthermore, such a pool should be more diverse in all aspects. Measuring COVID-19 behaviour could be done more elaborate, by possibly including more factors as well as externally confirmed behaviour, instead of self-assessed. Online privacy could be assessed better when more measures such as the 'do you accept cookies for this website' would be included. On the one hand, a question such as 'how often did you visit the supermarket last week' is based on revealed preferences. On the other hand, the eventual number filled in is self-reported, creating possibilities for cognitive errors. Furthermore, underlying motives for willingness to download a COVID-19 app should be extensively researched. If other research underlines the results that e.g. IT workers are more inclined to download a COVID-19 app, it could well be that they consider arguments in favour that other individuals tend to neglect. In-depth knowledge on the substance of these arguments could therefore provide useful insights. Moreover, the results indicated that people with more responsible online privacy behaviour, are more inclined to download a COVID-19 app. This is a result that, as mentioned, seems rather contra-intuitive. The implementation of a government controlled smartphone app with a GPS-tracker, clearly infringes privacy. In-depth research could be performed to understand the underlying motives for this correlation. The question of whether people are interested in an application could also be asked more profoundly, by exemplifying several options in terms of functionality, settings and lifespan. The survey applied in this thesis tried to avoid words such as privacy or data in the questions. It would be interesting to use a potentially bigger population to test whether the usage of these kind of words, that induce contemplations on the topic of privacy, change willingness to download COVID-19 applications. The

prevarication and ambiguity of the results indicate a substantial part of the question on individual's malleability when it comes to COVID-19 app related decisions is still open. This essentially stresses the necessity as mentioned earlier, that governments should extensively research the interaction between the diverse concepts. The importance of the codified obligations towards citizens privacy already guarantee for sufficient justification for such research to be performed. Concluding, the vital importance of including a smartphone application in the worldwide battle against the Corona virus seems rather clear. On the other hand, the sensitivity and malleability of the topic of online privacy (decisions) must not be neglected. Finding the adequate balance is for each government to decide, thoughtfully and considerate.

References

- Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behaviour in the age of information. *Science*, 347(6221), 509-514. doi:10.1126/science.aaa1465
- Acquisti, A., John, L. K., & Loewenstein, G. (2011). Strangers on a Plane: Context-Dependent Willingness to Divulge Sensitive Information. *Journal of Consumer Research*, 37(5), 858-873. doi.org/10.1086/656423
- Acquisti, A., John, L. K., & Loewenstein, G. (2013). What is privacy worth?. *The Journal of Legal Studies*, 42(2), 249-274. doi.org/10.1086/671754
- Acquisti, A., Taylor, C., & Wagman, L. (2016). The economics of privacy. *Journal of economic Literature*, 54(2), 442-92. doi: 10.1257/jel.54.2.442
- Athey, S., Catalini, C., & Tucker, C. (2017). The Digital Privacy Paradox: Small Money, Small Costs, Small Talk. *National Bureau of Economic Research Working Paper*, doi:10.3386/w23488
- Bouma, F. (2020, January 14). Aantal orgaantransplantaties met 15 procent toegenomen. *NRC*. Retrieved from: <https://www.nrc.nl/nieuws/2019/01/14/aantal-orgaantransplantaties-met-15-procent-toegenomen-a3629035>
- Buchanan, T. B., Joinson, A. N., Paine, C. B., & Reips, U. R. (2006) Measuring Internet Privacy Attitudes and Behaviour: A multi-dimensional approach. *Journal of Information Science*, 32(4), 334–343.
- Cakebread, C. (2017, November 16). You're not alone, no one reads terms of service agreements. *Business insider*, Retrieved from <https://www.businessinsider.com/deloitte-study-91-percent-agree-terms-of-service-without-reading-2017-11?international=true&r=US&IR=T>
- Cecere, G., Le Guel, F., & Manant, M. (2017). The Economics of Privacy. *The New Palgrave Dictionary of Economics*, 1–11. https://doi.org/10.1057/978-1-349-95121-5_3058-2
- Collewet, M., Kessels, R., Mouter, N. (2020, June 8). Nederlanders zijn het niet eens over de wenselijkheid van de corona app. *TU Delft*
- Corona-app laat nog zeker 4 weken op zich wachten - De Jonge erkent dat meer tijd nodig is voor veilige toepassing (2020, April 22). *Business Insider*. Retrieved from: <https://www.businessinsider.nl/corona-app-hugo-de-jonge-coronavirus/>

- Corona-app reëel risico voor privacy (2020, April 14). *Duurzaam nieuws*. Retrieved from: <https://www.duurzaamnieuws.nl/corona-app-reeel-risico-voor-privacy/>
- Ganti, A. (2019, August 16). Endowment effect. *Investopedia*. Retrieved from: <https://www.investopedia.com/terms/e/endowment-effect.asp>
- Hsee, C., & Leclerc, F. (1998). Will Products Look More Attractive When Presented Separately Or Together? *Journal of Consumer Research*, 25(2), 175-186. doi:10.1086/209534
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2017). *An introduction to statistical learning: With applications in R*. Springer.
- Johnson, E. J., & Goldstein, D. (2003). Do Defaults Save Lives? *Science*, 302(5649), 1338-1339. doi:10.1126/science.1091721
- Holroyd, M. (2020, April 8). Coronavirus: Sweden stands firm over its controversial COVID-19 approach. *Euronews*. Retrieved from: <https://www.euronews.com/2020/04/06/coronavirus-sweden-stands-firm-over-its-controversial-covid-19-approach>
- Kahneman, D., & Tversky, A. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131. <https://doi.org/10.1126/science.185.4157.1124>
- Kaiser, Henry F. "An Index of Factorial Simplicity." *Psychometrika*, vol. 39, no. 1, 1974, pp. 31–36., doi:10.1007/bf02291575.
- Klaassen, N. (2020, April 18). Zeven corona-apps in de race, privacy-experts kritisch: 'Dit is een janboel'. *Algemeen Dagblad*. Retrieved from: <https://www.ad.nl/tech/zeven-corona-apps-in-de-race-privacy-experts-kritisch-dit-is-een-janboel>
- Lin, T. (2019). Valuing Intrinsic and Instrumental Preferences for Privacy. *SSRN Electronic Journal*. doi:10.2139/ssrn.3406412
- Petrov, C. (2020, June 10). Big Data Statistics 2020. *Techjury*. Retrieved from: <https://techjury.net/blog/big-data-statistics#gref>
- Posner, R. A. (1981) The Economics of Privacy. *The American Economic Review*, 71(2), 405-409.

- Schudy, S., & Utikal, V. (2017). 'You must not know about me'—On the willingness to share personal data. *Journal of Economic Behavior & Organization*, 141, 1-13. doi.org/10.1016/j.jebo.2017.05.023
- Smith, C. (2010, June 17). 7500 Online shoppers accidentally sold their souls to Gamestation. *Huffpost*. Retrieved from: https://www.huffpost.com/entry/gamestation-grabs-souls-o_n
- Sunstein, C. R., & Thaler, R. H. (2009). *Nudge: Improving Decisions About Health, Wealth, and Happiness* (Revised & Expanded ed.). Chicago, Illinois: Penguin Books.
- Taylor, D. B. (2020, May 26). How the Coronavirus Pandemic Unfolded: A Timeline. *The New York Times*. Retrieved from <https://www.nytimes.com/article/coronavirus-timeline.html>
- Wagstaff, K. (2012, March 6). You'd need 76 work days to read all your privacy policies each year. *Time*. Retrieved from: <https://techland.time.com/2012/03/06/you-d-need-76-work-days-to-read-all-your-privacy-policies-each-year/>
- Woodward, A. (2020, April 18). South Korea controlled its coronavirus outbreak in just 20 days. Here are the highlights from its 90-page playbook for flattening the curve. *Business Insider*, Retrieved from <https://www.businessinsider.com/how-south-korea-controlled-its-coronavirus-outbreak-2020-4?international=true&r=US&IR=T>

Appendix

Tables and figures

Table 1.1

Participant's education level, summary

Level of education	Number of respondents
High school	14
Intermediate vocational education	6
College	14
University	51
PhD	1

Table 1.2

Participant's industry / sector, summary

Sector / Industry	Number of participants
Education	5
Transport	3
IT	5
Banking, law, insurance	3
Food, restaurants	4
Student	46
Health	5
Other	15

Table 1.3

Participant's gender, summary

Gender	Number of participants
Female	43
Male	43

Table 2.1

Standardised regression of COVID-19 behaviour variables on the chance of downloading a COVID-19 smartphone application for 86 observations

Variable	Regression Coefficient	Beta (standardised coefficient weights)
Washing hands	.20 (.13)	.18
COVID-19 news	.14 (.14)	.11
Supermarket visits	.21 (.17)	.15
Meeting non-household individuals	.02 (.11)	.02
Self-reported COVID-19 risk	.13 (.17)	.09
Constant	-.07 (15.71)	

*p-value < 0.1, **p-value < 0.05, *** p-value < 0.01, (F(5, 80) = 1.83, p = .12, R2 = .10)

Table 2.2

Standardised regression of online privacy behaviour variables on the chance of downloading a COVID-19 smartphone application for 86 observations

Variable	Regression Coefficient	Beta (standardised coefficient weights)
Cookies survey	.04 (.10)	.05
Income privacy	-.08 (.18)	-.05
Browsing cookie acceptance	-.09 (.14)	-.08
License agreements approval	.34 (.27)	-.16
Free-Wi-Fi browsing	.00 (.13)	.00

Number of passwords	.19 (.13)	.18
LinkedIn account	-.01 (.11)	-.02
Facebook Account	-.17 (.13)	-.18
Twitter account	.15 (.09)	.19
Instagram account	.05 (.12)	.05
Videocall software	.03 (.13)	-.03
Gift card	-.07 (.11)	-.09
Constant	43.16*** (12.92)	

*p-value < 0.1, **p-value < 0.05, *** p-value < 0.01, (F(12, 73) = 0.79, p = .066, R2 = .11)

Table 3.2

Regression of the chance of downloading a COVID-19 app on Corona-behaviour, age, income, education and occupation, 86 observations

Variable	Regression Coefficient	Beta-adjusted regression results
Gender (female)	8.27 (-.56)	9.15 (8.28)
Age	-.56* (.33)	-.56* (.29)
High school	-22.22 (15.51)	-22.55 (15.08)
Intermediate vocational education	-9.52 (15.80)	-10.50 (16.08)
University	-11.01 (13.22)	-9.03 (13.28)
PhD	-26.83	-29.17*

	(18.88)	(16.35)
<hr/>		
Occupation		
Education	4.78 (20.53)	5.18 (20.38)
Food, Restaurants	.58 (22.68)	1.31 (23.59)
Health	6.81 (18.82)	5.23 (17.39)
IT	35.16** (15.43)	28.01* (16.79)
Transport	-20.35 (20.24)	-23.10 (18.09)
Student	15.02 (14.70)	13.16 (13.05)
Other occupations	11.03 (14.62)	10.07 (12.92)
<hr/>		
Income		
€10.000-€20.000	-7.37 (10.60)	-4.59 (10.84)
€20.000-€30.000	12.76 (17.27)	14.04 (17.98)
€30.000-€50.000	1.00 (14.35)	2.70 (13.95)
€50.000-€70.000	64.75*** (12.49)	65.95*** (11.66)
More than €70.000	56.23*** (16.19)	48.15*** (16.63)
Prefer not to say	3.66 (21.47)	2.12 (20.26)
<hr/>		

COVID-19 behaviour	-1.00 (10.02)	32.50 (26.04)
Constant	51.40** (20.92)	51.21** (19.71)

*p-value < 0.1, **p-value < 0.05, *** p-value < 0.01, (Normal regression: F(19, 65) = 1087.65, p = .00, R2 = .27) (Beta-adjusted regression: F(19, 65) = 1181.76, p = .00, R2 = .29)

For the data including Corona-behaviour, a Jarque-Bera normality test was executed on the residuals, leading to a p-value of .229, implying the residuals were normally distributed. (Chi2 2.952), the null hypothesis of normally distributed residuals is not rejected. To test for constant variance in the residuals, a Breusch-Pagan test was performed, leading to a p-value of .405 (chi2 .769), implying there data is free of heteroskedasticity, the null hypothesis of constant variance in the residuals is not rejected. A variance inflation factor (VIF)-test was executed to check for multicollinearity. This resulted in problematic amounts of collinearity (higher than 5, James (2017)) for occupation categories student (10.68) and other occupations (5.92). This is not a problem, as these are dummy indicator variables, of a categorical variable with 8 categories. Due to the low proportion of occupations in the reference category, the other occupation indicator dummies necessarily have high variance inflations factors. When plotting the chance of downloading a COVID-19 app and Corona-behaviour in a scatterplot, a negative linear relationship is visible, implying the required linear relationship exists.

Table 3.3

Regression of the chance of downloading a COVID-19 app on online privacy behaviour, age, income, education and occupation, 86 observations

Variable	Regression Coefficient	Beta-adjusted regression results
Gender (female)	5.17 (8.79)	8.06 (8.78)
Age	-.55* (.28)	-.54* (.31)
High school	-18.49 (14.18)	-22.59 (15.39)
Intermediate vocational education	-8.74 (15.09)	-9.89 (15.97)
University	-9.79 (12.59)	-11.19 (13.24)
PhD	-35.38** (14.48)	-25.92 (18.91)

Occupation		
Education	3.54 (18.03)	4.18 (20.46)
Food, Restaurants	10.16 (19.35)	.92 (22.31)
Health	13.29 (17.66)	6.61 (18.91)
IT	29.04** (12.59)	34.60** (15.38)
Transport	-19.72 (17.42)	-21.86 (20.12)
Student	18.94 (12.17)	14.09 (14.29)
Other occupations	14.33 (12.58)	10.83 (14.58)

Income		
€10.000-€20.000	-9.36 (9.75)	-7.68 (10.52)
€20.000-€30.000	13.13 (17.25)	12.23 (17.05)
€30.000-€50.000	9.75 (12.35)	.42 (14.29)
€50.000-€70.000	59.21*** (11.31)	64.04*** (12.34)
More than €70.000	53.87*** (16.25)	55.59*** (16.11)
Prefer not to say	-1.82 (18.52)	4.61 (21.14)

Online privacy behaviour	18.72***	-38.84
--------------------------	----------	--------

	(7.01)	(97.69)
Constant	47.77** (18.73)	51.94** (21.08)

*p-value < 0.1, **p-value < 0.05, *** p-value < 0.01, (F(18, 66) = 1652.19, p = .000, R2 = .35) (Beta-adjusted regression: F(19, 65) = 1050.48, p = .00, R2 = .27)

For the data including online privacy behaviour, a Jarque-Bera normality test was executed on the residuals, leading to a p-value of .425, implying the residuals were normally distributed. (Chi2 1.714), the null hypothesis of normally distributed residuals is not rejected. To test for constant variance in the residuals, a Breusch-Pagan test was performed, leading to a p-value of .24 (chi2 1.38), implying there data is free of heteroskedasticity, the null hypothesis of constant variance in the residuals is not rejected. A variance inflation factor (VIF)-test was executed to check for multicollinearity. This resulted in problematic amounts of collinearity (higher than 5, James (2017)) for occupation categories student (10.23) and other occupations (5.92). This is not a problem, as these are dummy indicator variables, of a categorical variable with 8 categories. Due to the low proportion of occupations in the reference category, the other occupation indicator dummies necessarily have high variance inflation factors. When plotting the chance of downloading a COVID-19 app and online privacy behaviour in a scatterplot, a negative linear relationship is visible, implying the required linear relationship exists.

Table 4.1

Descriptive statistics per treatment

Variable	Treatment		Control		Mean control / mean treatment
	Mean	Stand. Deviation	Mean	Stand. Deviation	
Baseline characteristics					
Age	25.33	12.88	28.11	15.07	1.11
Gender (female)	.45	.50	.54	.50	1.20
Income					
Prefer not to say	.08	.27	.02	.15	0.25
€10.000- €20.000	.13	.33	.17	.38	1.31
€20.000- €30.000	.08	.27	.07	.25	.88
€30.000- €50.000	.13	.33	No Result	No result	-
More than €70.000	.03	.16	No Result	No result	-

Occupation					
Food, Restaurants	.05	.22	.04	.21	.80
Health	.03	.16	.09	.28	3
IT	.05	.22	.07	.25	1.40
Other	.20	.41	.15	.36	.75
Student	.55	.50	.52	.51	.95
Transport	.03	.16	.04	.21	1.33
Education					
High school	.20	.41	.13	.34	.65
Intermediate vocational education	.03	.16	.11	.31	3.67
University	.55	.50	.63	.49	1.15
PhD	No result	No result	.02	.15	-
Main variables					
Corona behaviour	-.03	.44	.03	.46	-1
Online privacy behaviour	-.02	.60	.01	.47	-.50
Download COVID-19 app	43.33	31.75	44.85	Not applicable	Not applicable
Number of observations	46		40		



English ▾

Cookies are important for the proper functioning of the site. To improve your user experience, Qualtrics uses cookies to remember login details, provide a secure login, collect statistics to optimize the functionality of the site and provide content tailored to your interests. Click Agree and Continue to accept cookies and go directly to the site, or click More Information for a detailed description of the types of cookies and choosing whether to accept these cookies when visiting this site.

[More information](#)

- Agree and continue
- Reject all and continue



English ▾

Thanks for taking the time to complete this survey. Please start with filling out the baseline questions on the next page. The survey will take you approximately six minutes.

What is your age?

What is your gender?

Male

Female

What is your sector / industry?

- Education
- Transport
- Food, Restaurants
- Agriculture
- IT
- Banking, Law, Insurance
- Health
- Student
- Other

How much gross income did you earn last year?

- Prefer not to say
- 0-€10.000
- €10.000-€20.000
- €20.000-€30.000
- €30.000-€50.000
- €50.000-€70.000
- More than €70.000

What is the level of the highest degree you started with? (not necessarily finished)

- Elementary school (basisschool)
- High school (middelbare school)
- Intermediate vocational education (MBO)
- College (HBO)
- University
- PhD

How often did you go to the supermarket last week?

- 0
- 1
- 2
- 3
- 4
- 5
- 6
- 7
- More than 7

How often did you meet up with friends / non-household family members last week?

- 0
- 1
- 2
- 3
- More than 3

What percent of the time did you wash your hands when coming home in the last 7 days?

0 10 20 30 40 50 60 70 80 90 100

In %



How likely do you estimate the risk of becoming substantially sicker than a simple cold due to Corona in the next three months?

0 10 20 30 40 50 60 70 80 90 100

In %



How often do you check the news on the Corona situation?

- Never
- Less than once per week
- More than 3 times per week
- More than 5 times per week
- Everyday
- Multiple times a day

What percent of the time do you deny cookies when browsing the web?

0 10 20 30 40 50 60 70 80 90 100

In %



What percent of the time do you read license agreements / general conditions / privacy statements completely before agreeing?

0 10 20 30 40 50 60 70 80 90 100

In %



Do you adjust your browsing behaviour when connected with a non-secured network? (free wifi in an airport, restaurant)

- Never
- Sometimes
- Most of the time
- Always
- I never connect to non-secured networks

How many different passwords do you use in total?

I do not want to answer this question

- 1 2 3 4 5 6 7 8 More than 8

Is your LinkedIn profile public or private?

- I don't have a LinkedIn profile
 It is public
 It is private

Is your Twitter page public or private?

- I don't have a Twitter page
 It is public
 It is private

Is your Facebook page public or private?

- I don't have a Facebook page
 It is public
 It is private

Is your Instagram profile public or private?

- I don't have an Instagram profile
 It is public
 It is private

Which programme do you use for contact with coworkers, friends or classmates? (multiple answers possible)

- Skype
 Zoom
 Microsoft Teams
 Other, namely
 I don't use these kind of programmes

Zoom Surpasses 300M Daily Users, Announces Zoom 5.0 with AES 256-Bit GCM Encryption

In today's "Ask Eric Anything" webinar, Zoom founder and CEO Eric S. Yuan provided more progress updates on our 90-day security plan, including exciting announcements about Zoom 5.0 and surpassing 300 million daily Zoom meeting participants.

Source: <http://web.archive.org/web/20200423091751/https://blog.zoom.us/wordpress/2020/04/22/90-day-security-plan-progress-report-april-22/> (Zoom blog)

South Korea controlled its coronavirus outbreak in just 20 days. Here are the highlights from its 90-page playbook for flattening the curve.

Aylin Woodward Apr 18, 2020, 1:45 PM



Due to a Corona smartphone application, South Korea has managed to control the outbreak in just 20 days, with the death toll decreasing ever since. The app includes government instructions, notifications for increased cases in your area and warnings if you've been in proximity of an infected individual, and has allowed the country to avoid extreme Draconian measures as implemented in several Western countries.

Source: <https://www.businessinsider.com/how-south-korea-controlled-its-coronavirus-outbreak-2020-4?international=true&r=US&IR=T>

How do you estimate the chance that you would download a COVID-19 smartphone application?

(This app would provide anonymised information on whether you have crossed a location where an infected individual has been shortly before, or whether you are nearing a location where an infected person is located)

0 10 20 30 40 50 60 70 80 90 100

In %



If you would like to participate in the €10,00 Bol.com raffle, please fill in your email address*, name, address and phone number below.

**Required to participate in the raffle*

Email address *(required to participate in the €10,00 raffle)*

Name

Address

Phone number

Thank you for answering this survey, it helps me enormously with my thesis.

Stata do file

Obtain beta values to get weights for privacy measures in order to adjust index

```
Reg downloadcovidapp cookies income browsingcookies licenseagreements browsingfreewifi  
ofpasswords linkedin twitter facebook Instagram videocallsoft giftcard, beta
```

Obtain beta values to get weights for corona behaviour measures in order to adjust index

```
Reg downloadingapp newschecking supermarketseefriendsfam washinghand estimcorrisk, beta
```

Perform Bartlett's test of sphericity and Kaiser-Meyer-Olkin measure on covid-behaviour

```
Factortest q45news-q48_1risk
```

Perform Bartlett's test of sphericity and Kaiser-Meyer-Olkin measure on privacy behaviour

```
Factortest cookies-giftcarddata
```

regress corona-app downloading on controls and treatment

```
reg coronaapp female1 age i.education i.occupation i.income treatm1, robust
```

regress corona-app downloading on controls and corona behaviour

```
reg coronaapp female1 age i.education i.occupation i.income covidbeh, robust
```

regress corona-app downloading on controls and privacy behaviour

```
reg coronaapp female1 age i.education i.occupation i.income privacybeh, robust
```

get f-statistic because f-statistic lacking

```
testparm coronaapp female1 age i.education i.occupation i.income treatm1
```

create variable for residuals

```
predict resid, residuals
```

perform Jarque-Bera test for normality residuals

```
jb resid
```

Breusch-Pagan test for heteroscedasticity

```
estat hettest
```

variance inflation factor test for multicollinearity

```
vif
```

perform regression with beta-adjusted privacy score for privacy behaviour

```
reg coronaapp female1 age i.education i.occupation i.income2 privacyadj, robust
```

*Encoding string variables (occup as example) into numeric variable in order to use them for

```

regression*
encode occup, generate(occupation)
*perform regression with beta-adjusted privacy score for corona behaviour*
reg coronaapp female1 age i.education i.occupation i.income2 covidbehadj, robust
*obtain f statistic*
testparm coronaapp female1 age i.education i.occupation i.income2 covidbehadj
*standardise value (privacybehaviour)*
egen float zprivacy = std(privacy), mean(0) std(1)
*obtain summary statistics*
summarize (variables)
*nearest neighbour matching test*
teffects nnmatch (coronaapp privacybeh (other variables) (treatm1)
*check for normal distribution of coronaapp download variable by Shapiro Wilk test*
swilk coronaapp
*perform two-sample t test with equal variances*
ttest coronaapp, by(treatm1)

```