CITY, UNIVERSITY OF LONDON

DOCTORAL THESIS

Essays in Behavioural and Experimental Economics

Author:
Alina VELIAS

Supervisors:
Alice MESNARD,
Sotiris GEORGANAS

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

in the

Department of Economics
School of Arts and Social Sciences

January 11, 2022

Declaration of Authorship

I certify that the thesis I have presented for examination for the PhD degree of the City, University of London is solely my own work. The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This thesis may not be reproduced without my prior written consent. I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party. I declare that my thesis consists of approximately 40'000 words.

The research paper providing basis for Chapter 2, "On the Measurement of Disease Prevalence", was co-authored by: Sotiris Georganas (City, University of London), Sotiris Vandoros (King's College London and Harvard University), and Alina Velias. All authors contributed equally to the study. Foteini Angelidou conducted the Greek language interviews, and the Department of Economics, City University, provided funding for the experiments.

The research paper providing basis for Chapter 3, "The Best is Yet to Come: The Impact of Retirement on Prosocial Behaviour", was co-authored by: Sotiris Georganas (City, University of London), Ioanis Laliotis (City, University of London), and Alina Velias. All authors contributed equally to the study. Foteini Angelidou conducted the Greek language interviews in the experimental part of the work. Ioannis Laliotis acknowledges financial support from City, University of London (Research Pump-Priming Fund 90648AL).

The research paper providing basis for Chapter 4, "Who is miserable now? Identifying clusters of people with the lowest subjective wellbeing in the UK", was co-authored by: Paul Dolan, Kate Laffan, and Alina Velias. All authors contributed equally to the study. Dawn Snape and Eleanor Rees (Office for National Statistics) provided comments on earlier drafts of the paper. Salah Mehad and Vahe Nafilyan (Office for National Statistics) provided comments on methodology.

Alina Velias
London, January 2021

CITY, UNIVERSITY OF LONDON

Abstract

Department of Economics
School of Arts and Social Sciences

Doctor of Philosophy

Essays in Behavioural and Experimental Economics

by Alina VELIAS

This thesis comprises three studies in behavioural and experimental economics. The first Chapter is a methodological investigation into the effect of self-selection bias on measurement of disease prevalence. The main issue is that "random" testing is commonly used to estimate prevalence. However, as long as such testing is voluntary, field estimates suffer from *selection bias*. We conduct an empirical application of this insight to Covid-19 testing and prevalence. In an incentivised lab-in-field experiment we show that people feeling symptoms are up to 42 times likelier to seek testing. This leads to *prevalence bias*: test positivity can inflate true prevalence five-fold. We validate using external data and confirm the bias varies intertemporally, making comparisons misleading. We suggest sampling the population to bypass the bias, yielding more accurate estimates, real-time. Our results are relevant to any epidemic, besides Covid-19, when carrier status informs beliefs.

The following Chapter explores the effect of retirement on prosocial behaviour. We show that retirement leads to more altruistic behaviour, and this change is not just attributable to external factors, such as a lower need for virtue-signalling, but seems to be caused by a change in preferences. To measure the impact of retirement we use a novel combination of representative cross-sectional and longitudinal individual-level survey data from 22 European countries, and a complementary incentivised field experiment on a representative sample of individuals. The effect on volunteering is strong in the survey data, and using the field experiment we identify a change in preferences as a probable cause. Given the ageing of the population these are policy-relevant findings. The welfare gain driven by increased prosociality, through increased volunteering and transfers, should be considered in retirement age reforms.

The subsequent Chapter addresses the problem of identifying the worst-off members of society. We take various measures of subjective wellbeing (SWB) as indicators of the how well people are doing in life and employ Latent Class Analysis to identify those with greatest propensity to be among the worst-off in a nationally representative sample of over 215,000 people in the UK. Our results have important implications for how best to analyse data on SWB and who to target when looking to improve the lives of those with the lowest SWB.

Acknowledgements

I am extremely grateful for the invaluable support and advice of my supervisors, Alice Mesnard and Sotiris Georganas. Their guidance helped me engage with academic research with thoroughness, constructive criticism, and excitement for the topic that I have never imagined. Their mentorship motivated me through the thorny parts of the path and elevated my successes. Thank you very much (και τώρα, πού βαδίζουμε!). I would also like to thank Jan Christoph Schlegel, Sergiu Ungureanu, Spiros Galanis, Rosemarie Nagel for their support and fruitful conversations.

I thank the City, University of London and the Department of Economics for the financial support that allowed me to develop the ideas presented in this dissertation and for creating a thriving scientific community. Finally, I thank Andrey Novikov, Valeriya Akylbayeva, Najieh Gholami, Valera Simonchuk, Saniya Karakhanova, Maria Zhivitskaya, Vasilis Tsolis, for making the write-up months much less stressful, and my years in London -- so happy.

Contents

List of Figures

List of Tables

Моим родителям, которые всегда в меня верят

Chapter 1

Introduction

This thesis comprises three chapters on behavioural and experimental economics. The first one studies the effect of self-selection bias on the accuracy of disease prevalence measurement. The second chapter explores the effect of retirement on prosocial preferences. The third chapter addresses the heterogeneity in factors associated with the lowest subjective wellbeing.

The first two chapters, although applied in different domains, both develop new methodologies to study real life economic phenomena experimentally. For both questions, tight control of several factors is important, making the case for controlled experiments. However, disease symptoms or transition into retirement are difficult to reliably model in the lab, which explains why the experimental literature has not dealt with these issues conclusively yet.

Experimental, incentivized methods have been used on virus testing before, in a seminal paper measuring demand for HIV testing (Thornton, 2008). The first Chapter extends this application to disease prevalence estimation. In order to reliably estimate people's propensity to seek testing when they have disease symptoms, we offer people incentives to measure their demand for testing, and then apply these results to validate responses to hypothetical non incentivised scenaria. This allows us to address the challenge of modelling disease symptoms in the lab by combining hypothetical survey data with an incentive-compatible validation.These hypothetical questions can then be used in scalable surveys, to get a large enough sample for valid inference, especially of people without symptoms.

The first Chapter also contributes an experimental angle to the existing literature treating selection issues in pandemic measurement purely econometrically (Greene et al., 2021; Manski and Molinari, 2021). In addition, as compared to the aforementioned studies' focus on the supply side-driven bias (such as preferential testing eligibility), our Chapter deals with the selection bias on the demand side, driven by non-monetary costs to testing.

In the following Chapter, we combine a controlled incentivised experiment with econometric methods more often seen in labour economics, to identify the effect of retirement. The issue is that unlike most experiments, it is not possible to separate the effect of age from the effect of retirement, by experimental treatment alone. On the other hand, existing techniques using surveys suffer from ill defined incentives and potentially inaccurate self-reporting. By combining the benefits of the two approaches we identify the effect of retirement on well defined volunteering and charity choices. In doing so, this Chapter finds evidence of increased prosocial preference caused by retirement, which contributes to the debate about the direction of change in volunteering behaviour as well as the mechanism of this change (Menchik and Weisbrod, 1987; Mutchler, Burr, and Caro, 2003; Sherman and Shavit, 2012).

In broader sense, this study is relevant to the literature on the stability of social preferences. Evidence exists on how prosocial behaviour changes in response to traumatic shocks, such as war (Gilligan, Pasquale, and Samii, 2014). This Chapter, in contrast, investigates the effects of a non- violent, anticipated shock, such as retirement, finding an increase in prosocial preference.

Finally, this Chapter describes several details of experimental design we developed in order to make the study accessible to elder people without compromising on transparency and incentive compatibility. Given the scarce experimental literature involving elder people (Sutter and Kocher, 2007), we hope that these ideas may help reduce some of the barriers of involving this demographic group in incentivised experiments.

In the final Chapter, we apply an exploratory, hypothesis-free clustering technique from data science to identify sets of life circumstances associated with reporting the lowest subjective wellbeing. Dimensions of wellbeing, such as happiness, life satisfaction or feeling of purpose do not lend themselves naturally to controlled lab environments. Likewise, the likely heterogeneity in the factors associated with being worst-off place large demands on the sample size. To mitigate these issues, we use a large-scale nationally representative dataset and restrict our definition of the worst-off across all available dimensions to limit the noise from misreporting or experiencing short-term wellbeing fluctuations. This allows us to describe clusters of people united by scoring the lowest on all the available dimensions of subjective wellbeing – and their observable characteristics. Recent SWB studies bring up the difference in covariates of SWB between its multiple dimensions (Dolan and Kudrna, 2016; Knabe et al., 2010) and at the extremes of the distribution, compared to the average (Binder and Coad, 2011; Lamu and Olsen, 2016).This Chapter contributes to this literature by identifying heterogeneous groups of factors associated with self-reporting lowest SWB across the available dimensions. More broadly, it also adds to the normative debate on redistribution of resources within the society - by identifying the groups of individuals who, potentially, are in most need (Dolan and Tsuchiya, 2011).

Chapter 2

On the Measurement of Disease Prevalence

2.1 Introduction

How to measure prevalence for infectious diseases? In the Covid-19 pandemic, health agencies (ECDC, 2021) and lay citizens alike, closely watch two measures derived from daily testing, the absolute number of recorded cases and the percentage of positives in the tested population. These numbers influence individual decisions but also official measures against the pandemic, with a profound impact on public health and the economy.

In this Chapter we claim that such commonly used measures are fundamentally flawed, because they ignore the demand side for testing. In virtually all countries in the world, testing is voluntary, leading to *self-selection bias*. People are likelier to self-select into testing if they have reasons to believe they might be having Covid-19 (such as, e.g. if they have symptoms or if they are exposed to a high-risk environment). We experimentally show there is a substantial bias in testing, driven by self-selection and demonstrate how the testing bias translates into *biased prevalence estimates*. We then validate our results on how the accuracy of prevalence estimation is affected by the bias, using external data and indeed find that positivity is inflated by 3.8 to 23.6 times, depending on the time period. This means that we cannot use a constant adjustment to debias positivity figures. Finally, we propose a novel, fast and relatively economical method to estimate prevalence in real time, using a combination of polling methods and characteristics of endogenously done virus testing.

Let us start with a simple illustration of the problem for economic policy makers and health agencies, using a real example from a European country. During Christmas all shops and schools were closed. On January 18 2020 the government allowed elementary schools and the retail sector to open (for in-store buys). About a week later, recorded cases started to rise. On the January 29, 941 cases were recorded, almost double the cases a week before (506). Ignoring statistical issues of significance, two questions arise: is that rise in cases a clear sign of a worsening disease, and who is to blame? Due to the selection bias, even the first question is hard to answer. At the same time as cases rose, testing rose too. On 29 January there was twice the testing than on the 22nd. Actually, *test positivity* is similar between these dates. But the self-selection argument implies that the number of tests is endogenous. Higher disease prevalence leads to higher demand for testing. As we will demonstrate, the self-selection bias changes over time, making comparisons using test positivity data meaningless in many cases.

The self-selection bias varies with age, which complicates answering the second question too, how to tell whether schools or shops are to blame. One would think to (and indeed, health agencies *do*) compare test positivity among school pupils and middle aged people who went shopping, to see what channel of infection was more important. But our experiments show that demand for testing differs strongly by age, and also, virus symptoms affect this demand differentially. This means we cannot compare positivity across age groups either.

The use of standard test positivity or the number of recorded cases, to compare prevalence over time or across age groups, is rarely advisable. In the paper we use incentivised controlled experiments to estimate the size of the testing bias and calculate the corresponding prevalence bias. Interestingly the bias is estimated to be drastically different by age groups (as mentioned above) and to also rely greatly on two important characteristics of the testing procedure: waiting times and cost.

Our testing bias estimates can be used to calculate the prevalence bias, and debias the current test positivity estimates in the field. As long as the characteristics of the testing procedures are known by the health agencies (rare to date), we sketch the parameter estimations necessary for debiasing. Given that we have estimates by age, simulations can be done for countries with different demographic structures too. Of course accurately estimating all necessary parameters presents challenges of its own.

To fix all the problems with measurement, we suggest a novel method to *bypass* the self-selection bias altogether, with an estimation procedure that is at the same time faster, more accurate and more feasible than current methods. The idea is to poll a representative sample about their symptoms, and get the symptoms-to-virus conversion parameters from existing tests.[1]

Finally, we present an application of the testing bias to the much debated policy question of school openings. We show that the testing bias can explain why the young do not show up in simple case counts, while they are very likely getting infected (and possibly transmitting) more than older people.

To understand the relevance of these results, note that policy responses (e.g. social distancing) will inevitably be inefficient if we are not aware of real prevalence, by location and age. Mortality and hospitalisation rates are not real time measurements; they only provide an estimate of how many people caught Covid-19 *weeks earlier*(and estimating the fatality rate is also challenging, Atkeson, 2020). This time lag is very important when trying to evaluate interventions. Without real time data, measuring the effect of a vaccine will take months, added to the time necessary a medical effect. Understanding the full effect of other events on the disease, like the Christmas holidays (which led to more interaction and possibly higher transmission) similarly takes months (see Brauner et al., 2020 on the effectiveness of pharmaceutical interventions (NPIs), using death counts). On the other hand, knowing the current number of actual cases, allows the design of optimal policy response, and also provides a forward-looking estimate of hospitalisations and mortality. Health systems get warning several weeks ahead, gaining invaluable time for necessary adjustments.

The possibility that infection rates in the untested population can be different than in the tested subsample, has been raised (Manski and Molinari, 2021). The issue is treated as a purely econometric inference problem however, with no reference to self-selection. In a somewhat similar vein, (Greene et al., 2021) propose statistical

[1]Replacing mass testing with polling may sound unusual, but it is in line with suggestions of using statistical sampling to replace exhaustive counting, when the latter can be biased, as in a census. In the case of the pandemic, it has even been argued that symptoms-based diagnosis should be used instead of PCR testing (Cadegiani et al., 2021), because it is more informative.

nowcasting. The accuracy of both these methods is probably not as high as polling, and detection of trend reversals is not possible in real time.

Experimental methods with incentives have been used on virus testing before, in a seminal paper to measure demand for HIV testing (Thornton, 2008). However prevalence estimation was not the goal of that paper, and of course the diseases are different in several ways.

More generally, the existing literature does not offer much guidance on personal incentives to test. Should people be averse to learning they are infected, as information avoidance models suggest (Golman, Hagmann, and Loewenstein, 2017), prevalence figures would be deflated due to symptomatic people testing less that non-symptomatic ones. If, however, people do not test unless they experience symptoms, as is known to happen (Oster, Shoulson, and Dorsey, 2013), prevalence figures would be inflated due to non-symptomatic people testing less frequently than symptomatic ones.

Why care about test positivity rates? These are currently widely used to evaluate the effect of the mass testing *within* a country (Mahase, 2020; Hsiang et al., 2020), to compare the effect of government policies *between* countries (Haug et al., 2020; Brauner et al., 2020; Hsiang et al., 2020), to build arguments about which age or socio-demographic groups are most affected (Elimian et al., 2020), and generally as a "baseline against which the impact of subsequent relaxation of lockdown can be assessed" (p2, Riley et al., 2020). A biased prevalence estimate makes these comparisons at best uninformative (Middelburg and Rosendaal, 2020) - a problem to which we offer a solution. Our approach is also relevant for past research based on historical data. For example, major studies of policy measures to prevent spread of viral diseases rely on prevalence estimates affected by the same type of bias (Adda, 2016).

Some studies rely on death rates instead of test positivity to evaluate policies aimed to contain the pandemic (Dergiades, Milas, and Panagiotidis, 2020). This measure does not circumvent the problem of incomparability. Deaths are affected by harvesting and specifics of the health system, so do not fit as a perfect proxy of prevalence for cross country comparisons. Likewise, the infection fatality rates (IFR) are also subject to the testing bias. Whilst researchers already raise concerns about methodological and econometric issues affecting IFR (Shen et al., 2021), the bias we find cannot be addressed by the measures they propose.

The rest of this Chapter is organised as follows. Section 2 presents calculations of the self-selection testing bias. Section 3 describes the experimental procedures to measure this bias. Section 4 presents the experimental results and their implications regarding the prevalence bias. Section 5 compares our debiasing solutions, partly with parameters derived from the experiments, to field data. Section 6 presents an application to a common policy problem, the evaluation of school openings, while Section 7 concludes.

2.2 Bias calculations

The aim of the calculations is to infer the percentage of sick people in the population from the "random" testing in the field figures, as released by Health Agencies worldwide. The problem is that testing is voluntary, which leads to selection bias. How large is this bias?

To start, some people believe they have symptoms, some do not: call them $S(ymptomatic)$ and $H(ealthy)$. Note that the discussion below has to do with what people believe, not what they actually have. Also, we distinguish between people believing they have symptoms and those who do not, but the analysis readily extends to people having strong beliefs that they might be carrying the virus and those who do not.

Let the frequency of people who believe they have symptoms be p_s, or just p, with $(1-p)$ being the frequency of people who do not think they have symptoms.

Of each group, some percentage turns out having the virus. Let v_s be the virus prevalence for those who believe they have symptoms, v_h for those who do not.

Of each group, some percentage are willing to take the test (for a given waiting time to take the test). Assume this only depends on symptoms, but not on actually having the virus (this assumption is mostly innocuous, unless there is a very large number of people in hospital). Let then t_s be the percentage of people who believe they have symptoms who actually take the test, and t_h for those who do not.

True prevalence is then

$$\tau = p_s v_s + (1-p_s)v_h \tag{2.1}$$

The sample prevalence, also called test positivity throughout the paper (i.e. the virus frequency in the sample population) ϕ, however, is given by the positive rate in the sample (assuming that the test itself is perfect).

$$\pi = p_s t_s v_s + (1-p_s)t_h v_h \tag{2.2}$$

Divided by the total sampling rate

$$m = p_s t_s + (1-p_s)t_h \tag{2.3}$$

Note that if $t_s = t_h = t$, then $\pi = t(p_s v_s + (1-p_s)v_h)$ and $\phi = t(p_s v_s + (1-p_s)v_h)/t = p_s v_s + (1-p_s)v_h = \tau$ which makes sense; if testing propensities are equal, there is no bias.

If on the other hand the testing propensities t are not the same, then the sample is selected, leading to bias. Before we calculate the bias, express the propensities to test and be virus positive, for the people who believe they have symptoms, as a multiple of the propensities of those who do not: $v_s = av_h, t_s = bt_h$. Then, using these equations, rewrite (2.1), (2.2) and (2.3).

$$\tau = p_s v_s + (1-p_s)v_h = ap_s v_h + (1-p_s)v_h = v_h(ap_s + 1 - p_s)$$

$$\pi = p_s t_s v_s + (1-p_s)t_h v_h = abp_s t_h v_h + (1-p_s)t_h v_h = t_h v_h(abp_s + 1 - p_s)$$

$$m = p_s t_s + (1-p_s)t_h = bp_s t_h + (1-p_s)t_h = t_h(bp_s + 1 - p_s)$$

Simplify the notation by writing p for p_s and calculate

$$\phi = \frac{\pi}{m} = \frac{t_n v_h(abp + 1 - p)}{t_h(bp + 1 - p)} = \frac{v_h(abp + 1 - p)}{(bp + 1 - p)}$$

Now, divide $\frac{\phi}{\tau}$ which yields the bias in the estimates

$$\beta = \frac{abp + 1 - p}{(ap + 1 - p)(bp + 1 - p)}$$

For example, suppose the true symptoms prevalence is 10%, $p = 0.10$. Then $\beta = (0.1ab + 0.9)/(0.1a + 0.9)/(0.1b + 0.9)$. For instance, if $a = b = 20$, street testing is overestimating the virus prevalence by about 5 times.

In order to debias the test positivity in the field, on simply has to deflate the field figures by the estimated β, as long as p is known. If it is not, calculations are available upon demand to get p from the data.

2.3 Experiment Design

To find the testing propensity parameters t_h and t_s, we design an incentivised experiment where we

1. Elicit hypothetical willingness to wait (WTW) to take a rapid test for Covid-19, conditional on (i) feeling healthy, (ii) having flu-like symptoms, (iii) having Covid-19 like symptoms.

2. Elicit real WTW to gain a voucher for a free rapid test for Covid-19.

2.3.1 Experiment data

Data collection took place from 11 till 18 December 2020, mostly online. For representativeness, 94 responses (16%) from elder people (median age = 63) were collected using phone interviews. Out of 605 participants starting the online study, 24 (3.97%) dropped out mostly after the first few questions, 6 failed to answer all demographic, resulting in the final sample of 575 observations. Median age was 39 years (median for Greece 45.6), and the age distribution is shown in the appendix.

Subjects were recruited from the database *paignia.net* and invited to participate in a study, answering a few question on behavior. Upon signing up for the experiment, and signing a consent form, the participant was asked about general and Covid-19-related health (Figure 2.1 depicts thefull experimental flow). We then elicited hypothetical willingness to wait (WTW) to take a rapid test for Covid-19, conditional on (i) feeling healthy, (ii) having flu-like symptoms, (iii) having Covid-19 like symptoms. For all three scenarios, the test was being offered by the national health authority (EODY) while the participant was walking down the street (this is the actual procedure, much discussed on popular media). The hypothetical location was chosen to eliminate the (hypothetical) travel costs. After eliciting the hypothetical WTW, we asked the subjects control questions, including exposure to Covid-19 risky environments (e.g. taking public transport) and socio-demographics. After completing the compulsory part of the study, the participants were randomly allocated to one of two prizes.[2] In treatment *Test*, the participant was offered a 1/30 chance lottery for a voucher for a home-administered Covid-19 test, worth €80 at the time of the study. In the baseline treatment *Book*, the participant was offered the same 1/30 chance lottery for a voucher from a large bookshop chain, which we also set to €80 value, for comparability[3]. Crucially, the participant had to complete a real-effort task to enter the lottery, and we made it clear that the part was optional. Participants were also reminded that they could stop at any moment.

All 575 participants completed the hypothetical elicitation and the control questions. As expected, a substantial part of the sample (n=174) did not continue to the

[2]Delivery to both was guaranteed within 36 hours.

[3]Evidence shows that people tend to value a high stakes lottery much higher than a certainty equivalent of its expected value (Kachelmeier and Shehata, 1992)

optional task. A major part (n=78) was the elder people subsample. We are not very concerned that the inconvenience of the waiting task over the phone was the issue, since the participants came from the sample that had participated about a month ago in an unrelated study involving a real effort task over the phone.

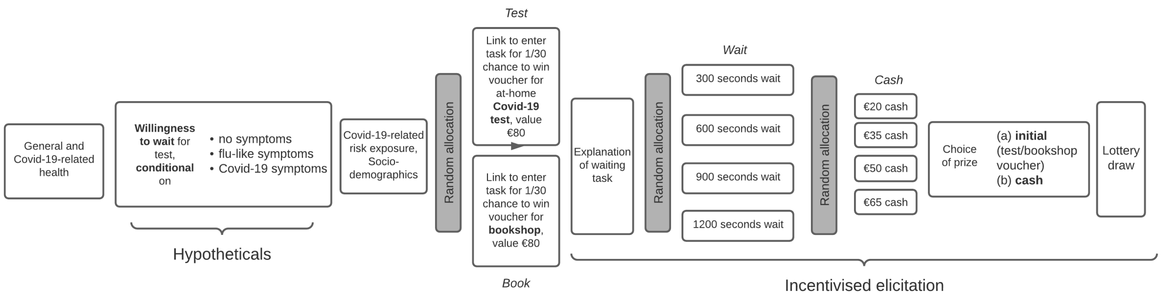


FIGURE 2.1: Experimental Flow.

The participants then read the description of the optional task. They learned that it involved waiting in front of their screen for some time (target) that would be revealed next, and the lottery draw for the prize would take place right after the wait. As attention check, a button would appear at random times and they had to press it within 4 seconds to avoid being disqualified. 300 of the 537 participants who read the description continued to the screen revealing the waiting target. They were randomly allocated to one of the four *Wait* target conditions $\{300, 600, 900, 1200\}$ seconds. Upon learning the *Wait* time, further 59 participants dropped out instantly (median target: 900 seconds). 69 dropped out before completing the target (median $Wait = 900$ seconds), while 172 participants did complete (median $Wait = 600$ seconds).

Upon completing the waiting task, each participant was randomly allocated to one of the four Cash conditions, {€20,€35,€50,€65}. The participant was offered a choice to enter the lottery for: (a) the original prize (*Book*, *Test*), or (b) the displayed Cash amount. Out of the 172 participants, 112 chose to swap the original prize for the cash amount, whilst 60 chose to stay with the original prize (median cash value €35 for both). Finally, 7 participants won the lottery.

2.4 Experiment Results and Prevalence Bias

2.4.1 Impact of self-selection on the bias in prevalence measurement

Hypothetical and incentivised waiting times

We find heterogeneity of waiting times between the age groups, driven by the self-assessed symptoms (Table 2.1). Younger people to behave similarly to elder people, while people between 30 and 50 are willing to wait less. This is reconcilable with the fact that this group has the highest employments rate and possibly family obligations, leading to less free time.

Table 2.1 shows the testing bias, as calculated by the ratio of willingness to test between people with symptoms and those without. The figure ranges between 1.5 and 42, depending on the age group and waiting times. People under 30 with symptoms are 1.5 times more likely to test when there is no waiting time, up to 42 with a 1-2 hour wait. The ratio for 30-50 year-olds ranges between 1.50 and 17.33. For over 50-year-olds, the ratio ranges between 1.66 and 9.4. The symptom-conditional difference is significant at $p < 0.01$, see Appendix A for details.

	Proportion WTW 30+ min, by symptoms			Bias by age and waiting time						
Age	**None**	**Flu-like**	**Covid-19**	**0′**	**5-15′**	**15-30′**	**30-60′**	**1-2h**	**2h+**	**N**
Under 30	0.156	0.391	0.641	1.50	2.74	4.10	11.67	42.00	42+	192
30-50	0.094	0.279	0.558	1.50	3.29	5.91	9.57	17.33	17.333+	222
50+	0.161	0.373	0.596	1.66	2.67	3.69	7.62	9.4	9.4+	161
Average				1.54	2.91	4.46	9.43	15.67	15.67+	575

TABLE 2.1: Bias by symptoms and age groups. LHS: Raw proportions of respondents reporting willingness to wait (WTW) for a Covid-19 PCR test for over half an hour. RHS: Bias (ratio of people with Covid-19 symptoms to people with no symptoms), by hypothetical waiting time for rapid test.

The propensity to test bias, translates to a biased virus prevalence estimate β, according to the calculations in Section 2. The prevalence bias is also time varying, even with no changes in testing strategies. It depends, crucially, on symptom prevalence, which can change drastically in a short period of time.

Apart from waiting times, self-selecting into testing also depends on the cost associated with it (if applicable – costs can vary from time to monetary value, travel etc). We test whether the hypothetical willingness to wait to take a Covid-19 test correlates with the incentivised real waiting time for the 1/30 lottery and find a significant positive relationship between the two (Table 2.2).

TABLE 2.2: Hypothetical vs incentivised waiting time for Covid-19 test

	Dependent variable:
	Real wait time (seconds)
Prizе:Test (Ref:Book)	-69.470^{**}
	(31.321)
Age	-7.643^{***}
	(1.014)
Hypothetical wait time (No symptoms)	15.175^{**}
	(7.703)
Constant	576.002^{***}
	(50.785)
Observations	537
R^2	0.112
Adjusted R^2	0.107
Residual Std. Error	362.864 (df = 533)
F Statistic	22.422^{***} (df = 3; 533)
Note:	*p<0.1; **p<0.05; ***p<0.01

Also, we measure willingness to pay for the test. Of those who won a test voucher, 83.8% swapped it for cash, as opposed to 48.9% of those who won the book voucher, indicating that the majority of subjects would not be willing to pay to receive a test.

Note that there were too few people reporting no symptoms to be able to compare the willingness to pay of people with symptoms, to those without. The scope of this study is to measure and correct the bias for free tests subject to different waiting times, and further experiments are needed to explore the effect of other monetary and non-monetary costs.

2.4.2 From Testing Bias to Prevalence Bias

To estimate the prevalence bias one needs estimates of symptoms prevalence and parameters a and b. We recommend polling to estimate symptoms prevalence, experiments for b, while a can be obtained by asking subjects at testing stations to self-report their symptoms before testing.

Suppose, for example, that community testing reveals 10% positivity, and 5% of the population reported symptoms. If those with symptoms are 5 times more likely to be positive than those without, and waiting time was 0, then the results of community testing exaggerate by 27.71%, and the true prevalence in the population is 7.83%. At a 30-60 minute waiting time, the bias increases to 106.95%, meaning that the true prevalence is 4.83%.

To illustrate our results, Figure 2.2 depicts our best estimate of the virus prevalence bias depending on symptoms prevalence and waiting time, by age.

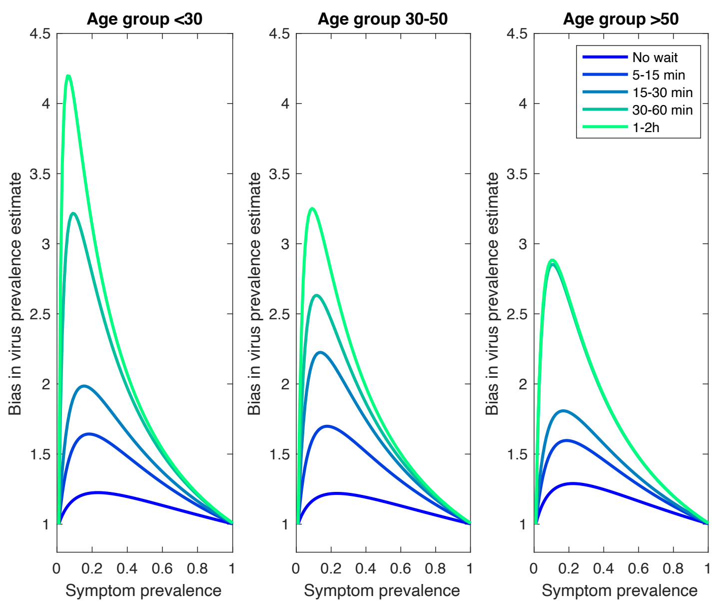


FIGURE 2.2: Best estimate of the virus prevalence bias: The ratio between reported prevalence and actual, depending on symptoms prevalence and waiting time, for the three age groups.

Based on these estimates, we can simulate how demography affects the prevalence bias. We use 3 million draws from the plausible parameter space (assuming symptoms prevalence of 5%, and allowing the testing bias parameter to vary uniformly within the 95% confidence interval gained from the experiments) applied to three countries, with different demographic structures: Nigeria (one of the youngest populations globally), Italy (heavily ageing population) and the USA (between the two extremes).[4]. The simulation shows that demography matters: Nigeria could have a substantially higher prevalence bias than Italy. However, waiting times are clearly more important than demographics. Lowering waiting times would result in a low bias for all countries, Figure A.1 .

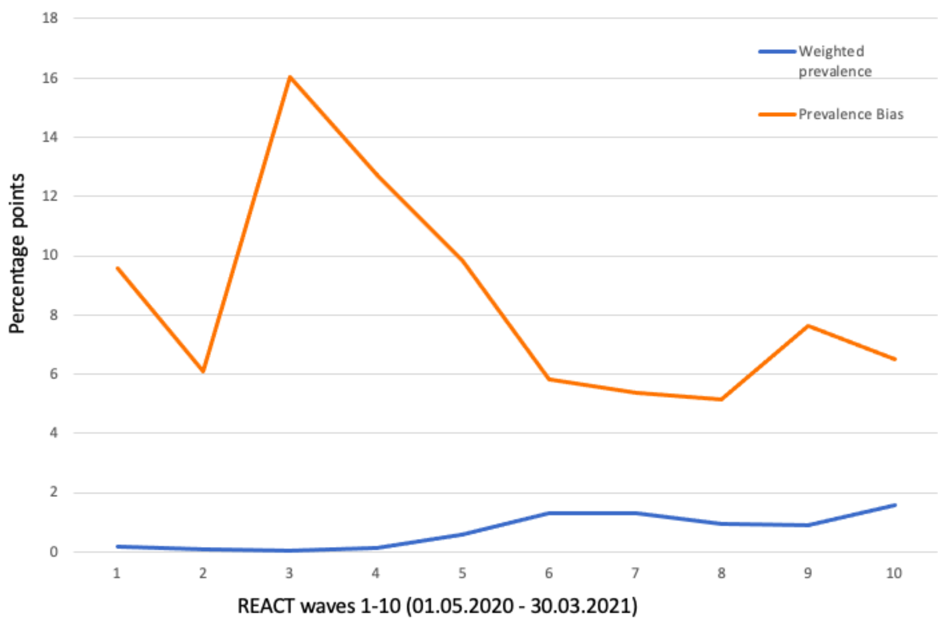


FIGURE 2.3: Estimate of the prevalence bias in field testing. Test positivity divided by the best prevalence estimate using REACT and ONS data.

2.5 Debiasing vs Polling for Prevalence Estimation: Validation using Existing Data

Debiasing the field prevalence numbers can be performed using our methodology, as long as there are good estimates for the relevant parameters, which can be hard. We suggest a novel, more economical and accurate alternative for prevalence estimation. The important parameter to estimate is the probability of having covid-19 conditional on having symptoms, and on not having symptoms, similar to parameter a above. This can be done by asking a simple question at existing testing sites (indeed we have ongoing parallel work underway to obtain these estimates in cooperation with testing centres in the field). These parameters could be country-specific and time-variant, but we do not expect changes to be too fast. Obtaining a few estimates

[4]Note that here we assume a fixed symptoms prevalence a of 5% across countries, and apply same values from the uniform variation of b to each country, to enable comparisons of the bias arising from aggregating within the country, across ages.

in each virus season could suffice, and this estimate could be used for many similar countries. The next step is unusual in the context of the pandemic: poll a representative sample regularly, *to obtain symptoms prevalence*. A common misunderstanding involves the argument that laymen cannot measure their symptoms properly. This is not a bug, but a feature of our procedure. Since the testing bias depends on self-reported symptoms, we need to condition on subjects *believing* they have symptoms, not on actually having them. Using both steps above can yield accurate prevalence estimates in real time at very low, comparatively, cost.

In the following we simulate the novel polling method and compare to data that are as accurate as possible. We use the REACT study in England (REACT, 2020) and the ONS Infection Survey as benchmarks, since these suffer, to our knowledge, from the lowest testing bias.[5]

REACT has been conducted in eight waves, including two sub-waves, yielding 10 different observations (we match the ONS data to these dates). For those registering to take part, a swab kit was sent with a request to provide a self-administered throat and nose swab, and a history of symptoms was also asked. The publicly available data includes the raw figures, as well as estimates weighted to be representative of the population of England as a whole.

We focus on weighted prevalence, as the most accurate and take the simple average of the two surveys to get our best prevalence estimates. The number of daily tests is publicly available, along with the number of tests being positive, yielding test positivity. We divide test positivity by the best prevalence estimate to obtain an estimate of the prevalence bias in field testing.

From our calculations and the experiment, three main hypotheses follow regarding the prevalence bias:

1. Test-positivity is always inflated due to self-selection, meaning the prevalence bias is large.

2. The prevalence bias is time-varying

3. As virus prevalence in the population increases, so does the bias in its measurement (for reasonable prevalence ranges in the Covid-19 pandemic)

In the 10 different sub-waves of the study, the estimated prevalence bias indeed is positive, substantial, but also highly variable, ranging from 3.8 to 23.6, thus confirming our two main predictions (see online appendix for details). Apart from the first waves, during which the testing strategy was changing, complicating comparisons, it seems there is a weak effect for the bias to be rising in prevalence. A proper test of this hypothesis would require more waves and a constant testing strategy.

In the next graph we compare the best estimates of positivity with the two methods used currently to proxy prevalence, field positivity and case counts (as a percentage of the country's population), along with the our two new methods, the debiasing estimate and a simulation of the polling method.

We simulate the polling method by taking symptom conversion parameters, as published in REACT, but from the immediately preceding wave. We use the symptoms prevalence numbers from the current wave. As long as agencies can get a polling estimate that is similarly accurate to REACT, this simulation places a lower bound on the accuracy of the polling method.

We find that the polling method is consistently closest to "true prevalence", while the debiasing estimate is further away and still inflates actual prevalence to some

[5]Both studies aim to test large, representative samples at home. Importantly, REACT sends testing kits to homes and participants can choose to send back results, while ONS sends health workers to test citizens. REACT non response, after kits are sent, is 74.6%, but unknown for ONS. Further research is needed.

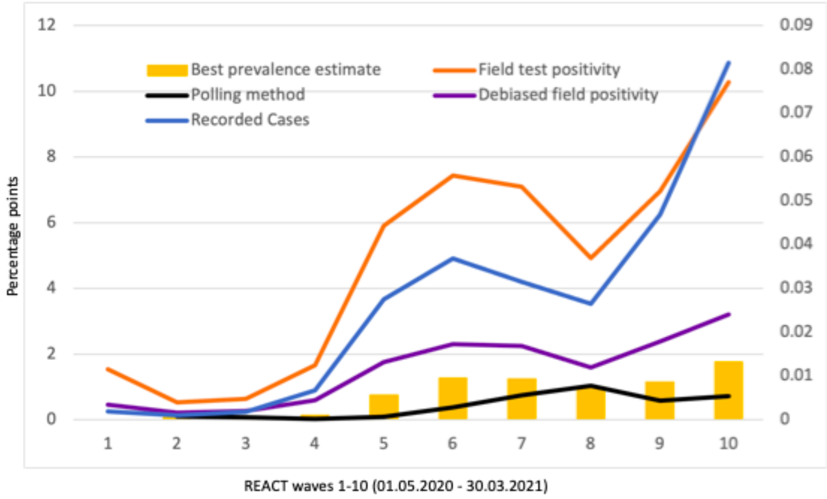


FIGURE 2.4: Comparison of the various prevalence estimates, in percentage of the UK population. The recorded cases curve refers to the right hand axis.

extent. As shown before, field positivity is an order of magnitude higher in most waves, while recorded cases are underestimating prevalence by at least an order of magnitude. Even assuming that cases sum up over several days to 10 times the daily rate, this estimate is still many times lower than estimated prevalence. Also note that all traditional methods are very variable, for example recorded cases increase almost fivefold when true prevalence doubles. Again, this is in line with our bias calculations.

A final note on the usefulness of the REACT and ONS methods: the marked difference between their prevalence estimates and common field test positivity, is driven by the fact that the monetary and non-monetary cost of testing happen are much lower in REACT and the ONS Infection Survey. Crucially, participants were able to administer the test and report symptoms without leaving the house. While this is a step in right direction, other significant non-monetary costs need to be mitigated in order to address self-selection bias. For example, for both studies, the physical unpleasantness of conducting the test may still make those not experiencing symptoms less likely to test. While it is possible to reduce other non-monetary costs of *testing*, we believe that making large-scale regular *self-reporting* of symptoms easy would be a more effective step towards achieving accurate prevalence estimates.

2.6 Application: Do Open Schools Lead to Transmission?

Closed schools cause problems to working parents, besides hindering the education of young pupils who reportedly find it hard to follow remote teaching. Studies have not yet yielded a clear, conclusive answer regarding the epidemic cost of school opening though and the debate remains heated.

Understanding the testing bias and how it varies by age group, allows us to reconcile the various pieces of evidence and solve existing puzzles. Looking at case counts, children and youngsters up to 19 years of age, seem not to be major carriers of the disease. Indeed, in a sample of 16 European countries for which data were available, children and teenagers up to 19 are always underrepresented among

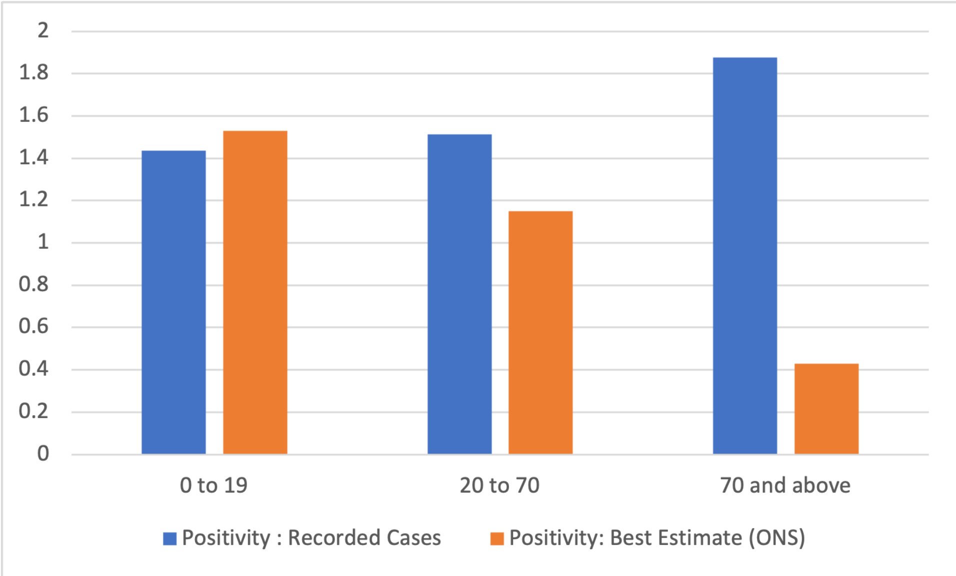


FIGURE 2.5: Recorded case positivity by age, vs best estimates.

confirmed cases.[6] Authorities around the world have used this as an argument that school opening is relatively harmless. However, our experimental results imply that young people are much less likely to test. While absolute testing propensities are similar, they are very different between those with symptoms and those without. The young have lower symptoms prevalence: conditional on testing positive, the frequency of people aged 5-17 reporting at least one of the four classic symptoms is 10.5% in REACT wave 8, vs 31.4% for those 18-54 and 22.1% for those 55 and above (also see Qiu et al., 2020; Kelvin and Halperin, 2020).[7] Combined with the large testing bias found experimentally, the different symptoms prevalence means the young test substantially less frequently.

As a consequence of the testing bias, the young are underrepresented in testing, meaning they are underreported in recorded cases. Indeed, although children aged 0-19 have the smallest presence in UK recorded cases (weighted by cohort size), high-school children are estimated to have the highest prevalence of all in the ONS survey (see figure 2.5). This example illustrates the importance of the selection-bias: how it complicates comparisons of prevalence in different age groups and can lead to wrong, in this case missing, pandemic prevention interventions.

2.7 Discussion

Using an incentivised online experiment, we found that the probability of taking a Covid-19 test for those who have symptoms (or believe they are more likely to have caught the virus) is many times higher than those who do not. In our sample, this testing propensity bias ranged from 1.5 times (for people under 30 years with no waiting time) to 42 times (for people under 30 and a 2-hour waiting time). The bias becomes larger with longer waiting times, and any cost associated with taking

[6]Finland and Norway had percentages above 15%, while the lowest were in France, Greece and Spain, below 7%. For comparison, the population share of 0-19 year olds in, e.g., Germany is 18.7%.

[7]Additonally, there seem to be reasons strictly related to the test itself that contribute to bias, due to the the under-detection of Covid-19 positivity in children, compared to that of adults (Dattner et al., 2020).

the test. Testing stations cannot readily correct this by oversampling (i.e. selecting people without symptoms to test).

A person's age also influences the testing propensity bias, which means that different locations will have different biases depending on demography. Furthermore, there have been reports of very long waiting times in some cases of community testing, which greatly exacerbates the bias and makes comparisons even within a country hard. Lastly, even keeping everything else constant, the bias depends strongly on the actual virus prevalence. All these effects combined mean the bias is very likely to be varying across space and time. Our findings imply that virus positivity results from community testing sites are heavily biased. Contrary to conventional wisdom in the health policy community that suggested the bias would be, if anything, downward, our results suggest that prevalence is inflated by up to 5 times, even under free testing.

We recognise the importance testing epidemiologically, to identify positive cases, allowing self-isolation to break disease transmission. If the goal of street testing is just to allow quick and free testing, then this possibly meets its goal. Note, however, that random testing is not efficient, economically, or epidemiologically: subsidising tests specifically for populations with a high risk of getting infected and infecting others would probably save more lives at lower cost (say, tests for young people working in service industries and living with their parents). These questions remain open for future research.

What we have shown is that "random" voluntary testing is not really random. As such, it does not provide accurate information on disease prevalence, which is important to design and implement urgent policy responses to the pandemic, in terms of type, intensity and geographic area. Since voluntary testing is always biased, aggregate results on prevalence should be corrected. We have explained a method to do such debiasing. Note that debiasing can be useful to get better estimates of prevalence in real time, but also to correct the past time series that are used to estimate and calibrate many models related to the pandemic. The object of such studies ranges from the effectiveness of measures against the pandemic (Brauner et al., 2020; Hsiang et al., 2020), proposals for new remedies such as test and tracing, to general health outcomes and economic effects. Furthermore, the probability to test is recognised as an important parameter in macroeconomic models evaluating economically optimal lockdown strategies (Alvarez, Argente, and Lippi, 2020.

Our methodology is not limited to correcting the results of community testing. We showed that the number of confirmed cases reported daily is also biased, strongly downward in this case. People might not test because of costs, or the inconvenience of going to a testing site, or even due to being afraid of losing income. According to our results, more than 85% of the people who are not feeling any symptoms, would not wait more than 30 minutes (a likely time in many street testing procedures) to have a test, even if it is provided free of charge. For people feeling symptoms the estimated percentage of non-testers is still about 40%. These percentages rise even further when tests have a non-negligible cost to the citizen.

Using polling results from a representative sample can correct the error both in recorded cases and field test positivity. Our proposed method is more accurate than these traditional proxies. Moreover the polling method is not costly, and does not require an extraordinary testing capacity, which means it can be used daily, allowing real-time prevalence estimation in myriads of communities worldwide.

The REACT and ONS studies are an interesting special case of large-scale community testing on a nationally representative sample. It is claimed that the sample is truly random. While we use such data as the best available estimate, our experimental results suggest randomness might be wanting. Even for free tests at home (see no waiting time condition in the experiment), a substantial testing bias exists. Importantly, REACT is also very expensive to run, while simultaneously less timely than our polling proposal. REACT is done monthly or less often, while our procedure can be run daily.

This Chapter also contributes to the literature on testing regimens (Mina, Parker, and Larremore, 2020). Mass testing, extending to a very large part of the population, is useful as it can provide more accurate figures, and also identifies positive cases. It has been used, among others, in Liverpool, Slovakia and South Korea (Pavelka et al., 2020; BBC, 2020; Bloomberg, 2020; Brauner et al., 2020). However, mass testing is extremely expensive, and might be infeasible, especially at frequent intervals, due to capacity and technical constraints.

In the absence of mass testing, obtaining unbiased prevalence estimates is of paramount importance for health and the economy. Underestimating disease prevalence can trigger inadequate measures and further spread of disease, while overestimating can be detrimental to economic activity. We thus urge policy makers to redesign "random" testing as a matter of priority in the effort to tackle the pandemic.

As a final note, our methodology is applicable to the prevalence measurement of any epidemic, when carriers have informative private information about their health status. Fighting disease is hard, even without the added complication of not knowing the location and magnitude of the fight. Our work offers tools to measure prevalence in real time. Further work is needed, to estimate specific selection-bias parameters for every disease, as they are necessarily related to the health burden and life expectancy reduction caused by the specific pathogen.

Chapter 3

The Best is Yet to Come: The Impact of Retirement on Prosocial Behaviour

3.1 Introduction

Retirees are often portrayed as carefree, kindly people in literature and film. Is there any substance to this portrayal - are retirees actually more prosocial? And if they are, does the increase in prosociality stem just from the wisdom that accompanies old age, or is there something special about retirement that makes people more altruistic? Given that retirees have more time, is their altruism related to spending of it helping others, or are they also more willing to sacrifice their income to support strangers and/or relatives? Does retirement cause a change in other-regarding preferences, and if yes what does this mean for the stability of preferences in general? To answer these questions, we use a unique combination of two types of data. The first one is survey data covering individuals from a large number of European countries: SHARE (longitudinal) and EU-SILC (cross-sectional). Variation in the Early Retirement Age legislation across countries, years and genders is used to identify the causal effect of retirement on various prosocial behaviour indicators. We use volunteering as a proxy for altruism and indeed find that retirees volunteer more. To identify whether the change in behaviour is really attributable to changes in preferences, we use a second type of data: we set up an incentivised experiment on a relatively large sample of retirees to identify the degree of pure altruism in the observed behavioural change, compared to other potential mechanisms such as changes in time and budget constraints, as well as the social circle.

This Chapter is the first to provide evidence on the impact of retirement on prosocial behaviour, as the latter is captured by participation in volunteering activities and contributions to charity. To our knowledge, this is the first study that measures volunteering activity in the subjects rather than relies on self-reported measures. Also, this is one of a handful experimental studies with retirees, especially involving real effort, with several methodological innovations being necessary given the age and special characteristics of the participants.

More generally, our results contribute to the literature on the stability of preferences and wider personality traits. There is a growing literature on the stability of risk preferences, across (e.g.Anderson and Mellor, 2009) and over time (Andersen et al., 2008, and Schildberg-Hörisch, 2018, 2018, for a survey). There is also some evidence regarding the response of risk preferences to shocks Hetschko and Preuss, 2020. Further afield, there is some evidence on the stability of strategic sophistication Georganas, Tonin, and Vlassopoulos, 2015. The stability of social preferences has been less investigated. Evidence exists on the stability across contexts (e.g. Wang et al., 2020), robustness to induced group identity Chen and Li, 2009. Closer to our work, Gilligan, Pasquale, and Samii, 2014 measure how social behaviour responds to heavily traumatic experiences, i.e. war. There is no evidence, to our knowledge, regarding the response to a non-violent, anticipated shock, that most of the population worldwide will necessarily experience in their lifetime, such as retirement. Note here that in the case of many countries in our sample, and especially Greece in the experimental sample, the retirement shock can be treated as exogenous. Either people were forced to retire at the official retirement age by the pension system, or the financial incentives were such that not retiring was a clearly suboptimal option. Our contribution is then to show that a mostly exogenous, relatively mild shock to people's lifestyles, that happens habitually to most people in the developed world and increasingly the developing, leads to measurable differences in their social behaviour and preferences.

3.1.1 Pro-social behaviour over age and employment status

There is no previous evidence on how transition to retirement affects prosocial behaviour. The economics literature is focused on labour market effects of volunteering using working-age samples[1]. Retirement is a major event that typically occurs later in life and for age-related reasons, and it is associated with important changes in lifestyle, consumption, activity, health and wellbeing (Battistin et al., 2009; Coe and Zamarro, 2011; Fitzpatrick and Moore, 2018).

The second theory argues for a substitution effect. It predicts a positive relationship that allows individuals to maintain their desired level of wellbeing. This substitution could be seen as a response to their increased time availability, the role of occupational loss, and the identity disruption caused by labour market disengagement. Using descriptive evidence, Sherman and Shavit, 2012 argued that changes from employment to retirement positively affect the likelihood to involve in volunteering activities. The authors discussed how the standard life-cycle hypothesis can be modified to predict the positive impact on volunteering for people who retire. While working, total consumption is the sum of all material goods plus the immaterial product of work per se, i.e. the subjective gains associated with -paid or unpaid-work. Under this assumption, total consumption will fall if the supply of unpaid working hours is zero post-retirement. As individuals smooth out their total consumption over time, they are incentivised to start engaging in some sort of prosocial behaviour, e.g. participate in voluntary activities, or increase their supply of unpaid level labour relative to their pre-retirement level[2] . Mutchler, Burr, and Caro, 2003 used data from Americans' Changing Lives (ACL) survey respondents aged 55-74 years old, to demonstrate increased volunteering activity for part-timers, those not

[1]Sauer, 2015 used data for women aged 25-55 years old from the Panel Study of Income Dynamics (PSID) and estimated that an extra year of volunteering increases full-time and part-time wage offers by 2.6% and 8.5%, respectively, and lifetime earnings by 16.7%. Cozzi, Mantovan, and Sauer, 2017 demonstrated that volunteer experience is related with higher earnings for both genders. Through a field experiment, volunteering has also been shown to increase the probability of getting hired (Baert and Vujić, 2018).

[2]Erlinghagen, 2010 reported that the effect of retirement on volunteering is exaggerated and mostly determined by the decision to continue offering voluntary work that was already taken up before retirement. However, the study pooled data from two distant waves (2001 and 2005) of the German Socio-Economic Panel, that did not allow for a direct test of this mechanism. Moreover, it did not evaluate the effect of retirement on voluntary activity, but rather reported regression coefficients - unconditional on time effects- of retirement status on binary variables indicating whether someone started or ceased volunteering between waves.

working and those who stopped working between interviews, relative to full-time employees. They suggested that this positive effect operates through formal, rather than informal, volunteering. The insensitivity of informal volunteering to work status was attributed to its obligatory nature and reductions in requests for help or support due to shrinking social networks once retired.

On the other hand, generativity was shown to increase interest in volunteering among later adulthood individuals, whilst community service motivation was significantly associated with individuals' interest in volunteering among all life stages, and social networking motivation was unique among the early and middle adulthood groups (Yamashita et al., 2019). Therefore, it is still not quite clear whether a potentially positive relationship between retirement and volunteering is due to increased time availability or enhanced prosocial preferences.

Individual behaviour is also affected by attitudes and behaviour within own social networks (Manski, 1993). Hence, peer influences from social and family networks can also affect volunteering behaviour. Friends and family members who volunteer could stimulate individual volunteering behaviour by the value of transmission, as it should be frequently encountered by their family members, friends and contacts (Van Goethem et al., 2014). Additionally, retirement exerts intra-household externalities on expenditure, home production and health behaviour (Moreau and Stancanelli, 2015; Müller and Shaikh, 2018; Stancanelli and Van Soest, 2012). Therefore, transitions to retirement could cause spillover effects on prosocial behaviour within the household.

Finally, prosocial behaviour post-retirement can be driven by previous experience and activity. This point has been raised by Erlinghagen (2010) who argued that the effect of retirement on volunteering is rather exaggerated and it is own previous experience that determines prosocial behaviour after leaving the labour market.

We are primarily interested in identifying the role of the pure altruistic component among these other factors. The literature on human altruism is very large. We have clear evidence that individuals often engage in prosocial behaviour -those promoting others' well-being (for organizations see Brief and Motowidlo, 1986, for individuals Rabin, 1993, and Fehr and Schmidt, 1999, for a general survey see Cooper and Kagel, 2016) - even when they are costly to themselves, e.g. volunteering, helping others, participating in political organisations, voting, donating to charities, etc. (Bénabou and Tirole, 2006). Age is a factor that partially explains variation in prosocial behaviour; price incentives, personality traits, social pressure, institutions, gender, and education are some others (e.g. Brañas-Garza, Capraro, and Rascon-Ramirez, 2018; Dohmen et al., 2008; Kettner and Waichman, 2016). The economics and psychology literature point to positive links between prosocial behaviour and age: motivational shifts to more emotional goals (Carstensen and Charles, 1998, empathy and prosocial behaviour Sze et al., 2012, both probability and amount of charity donations (Bellemare, Kröger, and Van Soest, 2008; Carpenter, Connolly, and Myers, 2008; List, 2004) as well as hours volunteered Katz and Rosenberg, 2005)[3]. In this study we explicitly ask: does retirement make you a nicer person – and why?

3.1.2 Economic significance of post-retirement volunteering

Our findings improve our understanding of the behaviour, and time and effort allocation decisions of a growing part of the European population that exit the labour

[3]The empirical part considered several indicators of prosocial behaviour, i.e. volunteering, providing care, active citizenship. The terms volunteering and prosocial behaviour are being used interchangeably throughout the paper.

market in later life, i.e. retirees. Transitions to retirement become more frequent with population ageing. In Europe, working-age population is decreasing and the old-age dependency ratio, i.e. people over 65 years old relative to working-age ones, will rise from 29.6% in 2016 to 51.2% in 2070. Moreover, life expectancy at retirement will increase by over five years by 2070 (European Commission, 2018)[4]. Hence, behaviour, time allocation decisions and productive capacity of retirees are central in policy-making (Centre for Ageing Better, 2018; Mutchler, Burr, and Caro, 2003). Lastly, this study contributes towards the wider agenda of understanding what drives prosocial preferences in general, highlighting heterogeneity in prosocial motivations driven by age and employment status.

Understanding the link between retirement and prosocial behaviour is important for two reasons. Firstly, for the accurate welfare analysis of retirement-related policies. Despite the fact that economists often attach an explicit zero wage to the supply of volunteer labour -or unpaid work in general-, the implications for the economy are considerable Menchik and Weisbrod, 1987. Therefore, volunteering -even later in life- can substantially contribute to the economy. In the UK, for example, volunteer work represents about 3.7% of the total labour supply in the country[5] [6].

Moreover, the 7% drop observed in dedicated volunteer time during 2012-2015 was associated with a loss exceeding £1 million in the UK. They also estimated that one hour of volunteering per week is worth £750.4 per year, and that unpaid work for this type of activity has the second highest value after childcare, and ranks before other unpaid activities, e.g. housework, adult care, transport, laundry and cooking. As the figures sketched out above refer to volunteering alone, the total contribution coming from all types of prosocial behaviour in an economy is higher, although difficult to be precisely quantified. As both the share of retirees and their life expectancy will keep growing, getting them involved in prosocial activities will enhance their role in reducing social costs and increasing welfare, let alone the positive effects that volunteering has been shown to have on own wellbeing. Apart from volunteering, the increased prosociality of retirees could mean higher transfers to their offspring and children. This means that retirement policies have obvious distributional effects, but also possibly effects on consumption patterns.

Secondly, for the practical reasons of addressing a potentially excess demand for volunteering opportunities by retirees, and harnessing the associated welfare loss. In the 2015-16 Community Life Survey (CLS), 51% of their 65-74 years old respondents participated in voluntary activities at least once a month, compared to 42% and 30% for the 50-64 and 25-34 age groups, respectively. Moreover, according to the ONS analysis, those above 65 years old reported 13.4 minutes of formal volunteering per day, on average, while those aged 25-34 reported a daily average of only 6 minutes.

[4]Consequently, labour force participation among those 20-64 years old in the EU will increase from 77.5% in 2016 to 80.7% in 2070 (European Commission, 2018). As a response, many countries encourage new forms of working in later life, e.g. partial-retirement, bridge jobs and un-retirement (Centre for Better Ageing, 2018). Wahrendorf et al. (2017) presented evidence about paid employment beyond the age of 65 becoming more common across Europe.

[5]The UK is used as an example due to data availability. The ONS analysis used data from the UK Household Satellite Accounts. More information can be found here: www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/articles/billionpoundlossinvolunteeringeffort/2017-03-16

[6]The Office for National Statistics (ONS) reported that 1.93 billion hours of volunteer work were supplied in the UK during 2015. According to the UK Labour Force Survey (LFS), the total number of actual weekly hours worked in 2015 was 1,004 million, or equivalently 52.2 billion worked hours over the year. Hence, volunteer work represented about 3.7% of the total labour supply in the country www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/timeseries/ybus/lms

The remainder of this Chapter is structured as follows: Section 2 presents the data sources and the construction of the main variables used in the analysis. Section 3 outlines the adopted identification strategy, including the experimental design. Section 4 presents and discusses the results, while Section 5 discusses the mechanisms behind the results. Section 6 concludes.

3.2 Data

3.2.1 Survey Data

The International Labour Organisation (ILO) defines volunteering as "any unpaid, non-compulsory activity to produce goods or provide services for others; that is for economic units outside the volunteer's household or family". In our main survey dataset (SHARE) individuals were asked whether they did any voluntary or charity work during the last year. SHARE is a cross-national longitudinal survey collecting information on demographics, health, and socio-economic status for individuals aged over 50 years old, which makes it an excellent fit for studying the effects of transitions into retirement. In waves 4, 5, 6 and 7 the information about respondents' voluntary activity last year was collected. Following individuals over time allows to control for unobserved heterogeneity and for dynamics in prosocial behaviour.

We supplement this longitudinal data on prosocial behaviour in the target age group with the cross-sectional European Union Statistics on Income and Living Conditions (EU-SILC) data which benefits from measuring a range of pro-social activity indicators. In the 2015 and 2016 waves, they provided information on prosocial activity, i.e. volunteering (formal and informal; 2015 sample), provision of care to others (inside and outside the household; 2016 sample) and active citizenship (2015 sample). Full details are provided in Appendix A1.

3.2.2 Experiment Data

To isolate other mechanisms driving volunteering from an increase in prosocial preference, our design also incorporates a telephone field experiment, which involved a separate subject pool. In the experiment, subjects answered a questionnaire similar to the EU-SILC and SHARE ones, which included questions about volunteering. Subjects then took part in an incentivised experiment with real in-kind and monetary donation outcomes. We describe the experiment in detail in the mechanism section below. Using social media announcements, word of mouth and local pharmacies we specified the target age to be 15 years around retirement, and randomised contacts between waves. The data collection took place over 3 waves, 4 days each: 30 April - 3 May, 14-17 May, and 24-27 September 2020 (i.e. during, straight after and a while after the COVID-19 lockdown in Greece). A total of 255 individuals aged 38-84 years old participated[7].

The first set of measurements we take are identical to those in EU-SILC and SHARE data, which enable to compare self-reported volunteering between surveyed individuals and our experiment participants. Specifically, we ask participants about self-reported volunteering (and reasons for not volunteering), time spent helping family, general social and cultural activity, health, and the standard socio-economic

[7] Response rate was 63% in the first wave (150 contacted, 94 participated), 70% in the second wave (150 contacted, 105 participated), and 60% in the third wave (93 contacted, 56 participated).

indicators such as age, gender, education, labour status etc. The second, fully incentive-compatible, set of measurements is explained in detail in the experiment set-up section below.

3.3 Estimation Strategy

The objective is to identify the causal link between retirement status and prosocial behaviour. Retirement is endogenous as individuals can opt to retire earlier or later in their lives depending on their health, wealth, time preferences, and institutions regulating the retirement eligibility criteria applying to their case. A natural experiment randomly assigning individuals to groups of retirees and non-retirees would be ideal in providing a causal answer to this empirical question. However, such experiments are not feasible and simple regression methods are likely to result in biased and inconsistent estimates regarding the effect of retirement on prosocial behaviour. Instead, a fuzzy regression discontinuity (RD) design is adopted. Endogeneity concerns are addressed by exploiting discontinuous jumps of the retirement probability at the year, country, and gender-specific ERA thresholds that apply to each surveyed individual.

To isolate the role of change in prosocial preferences from other factors, we designed an experiment that involves subjects making incentive-compatible choices about in-kind and monetary contribution to charities. Crucially, these contributions are free of any social capital, social network and other components discussed above as alternative drivers of volunteering.

3.3.1 Forcing variable: retirement eligibility

Retirement eligibility is at the heart of European welfare systems. The calculated distance between an individual's age[8] and the official early retirement age (ERA) in their country will be used as the forcing variable in the econometric design[9]. Over time, countries have implemented a series of pension reforms, including ERA increases. Hence, ERA is an important institutional threshold determining who can exit from the labour market and start claiming pensions. Moreover, it is associated with major changes in behaviour, health and lifestyle (Fitzpatrick and Moore, 2018). Using individual responses regarding own current activity status, a binary variable is constructed to indicate retirees, versus non-unemployed labour market active individuals, i.e. excluding those in military or community service, studying, disabled or performing domestic tasks.

In our sample of European countries, ERA is more frequently set after 60 years old. However, there is variation over time, across countries and by gender. For example, ERA in Austria in 2015 was 64 years and 59 years for men and women, respectively, while in 2016 both thresholds were increased by one year. In other countries, e.g. Denmark and Czech Republic, ERAs remained unchanged (see appendix C).

[8] Age is crucial because it predicts the treatment status; i.e. the retirement probability increases with age. Information about the timing of the interview and the timing of birth is also available; i.e. quarter and month in the EU-SILC and the SHARE data, respectively. Hence, we calculate the respondent's age at the time of the interview.

[9] Similar to other studies, e.g. Müller and Shaikh, 2018, information on early retirement eligibility age was collected from the Social Security Programs Throughout The World Survey (SSPTWS) which is available from the U.S. Social Security Administration (2016), as well as from OECD Pensions At A Glance reports (e.g. OECD, 2017

3.3.2 Fuzzy Regression Discontinuity Design

The implementation of an RD design relies upon information of a policy rule that determines whether an individual is potentially treated. In this context, retirement status is partially determined by whether someone's age, i.e. the forcing variable, crosses a known cutoff point c, which is the early retirement eligibility age that applies to each individual given their country, survey year and gender. The validity of the RD design relies on the fact that individuals cannot manipulate the forcing variable around c, and therefore they are considered to be randomly classified as treated and non-treated Lee and Lemieux, 2010.

In the European context, crossing the institutional cutoff point does not imply compulsory retirement. Instead, there is imperfect compliance because the discontinuity in the retirement probability is lower than 1 as someone crosses their ERA. This calls for a fuzzy RD design where the forcing variable (age) can only partially determine retirement status. Therefore, a Two-Stage Least Squares (2SLS) framework is applied in order to instrument individual retirement status using the predicted discontinuity in the probability of retirement after crossing the ERA. The following system of parametric equations is estimated:

$$y_{ic} = \beta_0 + \beta_1 D_{ic} + \beta_2 \widetilde{age}_{ic} + \beta_3 D_{ic} \widetilde{age}_{ic} + X_{ic} + \mu_c + \epsilon_{ic} \tag{3.1}$$

$$D_{ic} = b_0 + b_1 \widetilde{age}_{ic} + b_2 I_{ic} + \mu_c + v_{ic} \tag{3.2}$$

In this framework, y_{ic} is the prosocial behaviour indicator for individual i in country c, and D indicates the individual retirement status. The forcing variable is centred at the country, survey year and gender-specific ERA, i.e. $age_{ic} = (age_{ic} - c_c)$. Retirement is instrumented using binary indicators on whether the$i$-th individual has crossed the ERA threshold, i.e. $I_{ic} = 1[age_{ic} \geq 0]$. All models include a set (X_{ic}) of pre-determined characteristics such as gender, education, and ethnicity. A set of country fixed effects, μ_c, is also included to account for time-invariant differences across countries, e.g. in the institutional framework or culture. Models also control for time fixed effects to adjust for common time trends[10]. Finally, ϵ_{ic} and v_{ic} are idiosyncratic disturbance terms. In the cases where Equation 3.1 is estimated using longitudinal rather than cross-sectional data, models control for individual fixed effects and, in some cases, lagged own and partner's prosocial indicators.

Under Equation 3.1, the impact of retirement on prosocial behaviour is given by β_1. Equation 3.2 is the first-stage regression indicating how retirement probability changes at the cutoff, i.e. b_2 . Linear interaction terms between the forcing variable and the instrument are also included as additional instruments, i.e. $age_{ic} I_{ic}$, in order to allow for different slopes at both sides of the cutoff point. In this case, additional first-stage regressions are estimated for the interaction terms. Models using higher-order polynomials of the forcing variable are also estimated, although their use is avoided in RD designs due to poor performance, especially when samples are not sufficiently large around the cutoff (Gelman and Imbens, 2019).

Given this framework, the estimated β_1 coefficient in Equation 3.1 is interpreted as the Local Average Treatment Effect (LATE) of retirement on prosocial behaviour. In other words, it is the average treatment for the compliers, i.e. for those individuals who exit labour market and retire once they cross the ERA threshold.

[10]The EU-SILC data are cross-sectional. However, when using those data a set of quarter-of-survey fixed effects is included. Year fixed effects are included in models estimated using the SHARE data.

3.3.3 Mechanism

Prosocial behaviour can be driven by a variety of factors, as discussed in the Introduction. This behaviour change may be caused by people accommodating their existing preference in their new lifestyle. A retiree, for example, may be volunteering more to compensate for the social interaction previously consumed at work. Alternatively, the same observed behaviour can be driven by a change of preferences upon retirement. We first take advantage of the data on partner?s activity in the SHARE sample. We thus examine the effect of social network by controlling for intensity of a partner's volunteering in the longitudinal model. We also test for retirement spillover effects on prosocial behaviour within the household by controlling for partner's retirement. Furthermore, we use the longitudinal nature of the SHARE data to check for the role of own past volunteering experience in current behaviour (see Appendix E for the details).

In the experiment, we fully eliminate the effect of social and strategic components which allows us to isolate the role of change in purely pro-social preference. Conceptually, we consider three types of consumption goods. This is similar to Sherman and Shavit, 2012 who in their study of volunteering motivations divide consumption goods into material and immaterial. In our framework, we further distinguish between two types of immaterial goods: altruistic and other. The other category comprises all previously discussed motivations not driven by changes in prosocial preferences, such as social interaction and signalling of social capital. In the experiment, we eliminate the possibility of consuming the *immaterial-other* good from the in-kind and monetary donations. This leaves the *material* good and *immaterial-altruistic* good as the only consumption options.

Incentivised elicitation of the mechanism

Along with the measurements identical to those in the EU-SILC and SHARE data, we develop a second set of fully incentive-compatible measurements. These involve participants' choices about (i) in-kind and (ii) monetary donations they can make to real charities. The set of recipients for each task's earnings covers all possible consumption options: (a) keep it for themselves; (b) give it to relative/friend; (c) donate to Church of Greece; (d) donate to environmental charity; (e) donate to refugees charity; and (f) donate to cancer charity[11].

The first measurement is a real effort task designed to gauge participants' preference for volunteering in an incentive-compatible way. Participants were given a simple, but time-consuming, task which allowed them to earn up to €5. First, they had to indicate a recipient from the list above for their task-related earnings. Those who have a preference for donating their time to charity in real life are expected to be more willing to volunteer their time (and labour) towards a charity recipient in our experiment. Crucially, this measurement isolates the purely altruistic preference for time donation from alternative explanations involving non-monetary utility gains from the time spent volunteering. Under the assumption that the design of the real effort task (described below) provides the pure disutility of time use, we rule out the potential complementary utility gained, e.g. a social aspect of volunteering.

[11] In the cases of (d), (e) and (f) participants could nominate the charity of their choice.

Experiment Design

We created a number of novel experiment design solutions to ensure the inclusivity of the sample, and the credibility of the incentivised measures. Both are crucial for the external validity of the results. Specifically, we expected some retired participants to be less tech-savvy and less willing to travel, compared to the average subject pool of a study reliant on a real-effort task. Additionally, COVID-19 quarantine measures came into place, which eliminated the option of administering a face-to-face study. It is not surprising that experiments involving elder and, in particular, retired people are very rare. To our knowledge, there is only one controlled laboratory experiment involving retirees in their late sixties. Sutter and Kocher, 2007 use a lab setting to study the relationship between age and trust. Their sample of 64 retirees was recruited from athletic courses for retired persons at the local Department of Sports and from an adult education institution, which organizes seminars on various topics and for various age groups[12]. Our experiment is the first to administer a controlled experiment administered to a sample of over 250 retirees and persons of close-to-retirement age.

We believe that the combination of methods we developed, i.e. telephone interview, adaptation of real-effort task and adaptation of a credible random number draw to be administered over the phone, can be successfully used in the future to reach people often under-sampled in research, mostly due to age and other relevant demographics.

Real Effort Task

We selected the real effort task that satisfied the following criteria that ensured the representativeness of the sample: (i) skill-independent; (ii) free of intrinsic value (to avoid unobserved heterogeneity in enjoyment, sense of purpose, etc.); (iii) suitable for elder people; and (iv) easy to administer over the phone, rather than online or in the lab. Whilst most of the existing real effort tasks satisfy requirements (i) and (ii), requirements (iii) and (iv) were specific to both our research question and the data collection timing. We wanted to minimise the exclusion from our sample based on technology (for example, online participation requires a certain level of internet proficiency) or willingness to travel (for lab participation). To address this challenge, we designed a novel real-effort task that was administered over the phone.

The real-effort task involved counting the number of vowels in common words of the local language, i.e. in Greek. Subjects could earn up to €5 by completing up to seven sets of tasks, four vowel counts in each. To ensure equal time and effort cost, we pre-recorded audio clips of the research assistant pronouncing the words, with gaps in-between for the subject to provide the answer. The participants first learned the nature of the task and then they could choose one recipient of their real-effort task's earnings. Then, they listened to the audio clips and reported the vowel counts. Participants were free to stop the task at any stage and move on to the next section of the study. See Supplementary Materials for the instructions and screenshots of the tasks.

[12]Similarly, Holm and Nystedt, 2005 administered a mail-based semi-experimental trust game with participants of 20 and 70 years old, using a public database in Sweden. Charness and Villeval, 2009 explore cooperation and competition in a sample involving 39 elder (over 50 years old) employed people.

Lottery-over-phone

Designing how to credibly administer the lottery over the phone was non-trivial. The interviewer would ask the participant to find any banknote in their wallet and read out the digits of the serial number, apart from the last two. The last two digits were the lottery number. The interviewer would then generate a random number and tell it to the participant. The participant knew that they would win the lottery if the last two digits of the serial number on their banknote (unknown to the interviewer at that point) were the same as the random number the interviewer had just given to them (Appendix Figure A6).

Addressing demand effects

Being observed (by the experimenter or by other subjects) creates an additional cost of not complying with social norms, and can thus change the subject's behaviour (Georganas, Tonin, and Vlassopoulos, 2015). This is particularly problematic if studying charitable giving. DellaVigna, List, and Malmendier (2012) estimated the social pressure cost of refusing to donate to lie between $1.40-$3.80, depending on the charity type. Consistent with this, in the pilot of this study we observed most of the subjects donating the endowment of ¤5 to charity. We address this concern in two ways, (i) using a high-stakes lottery instead of a small certain endowment, (ii) adding a more socially ambiguous option of donating winnings to a family member or a friend.

The main function of the higher stakes lottery is to mitigate the "peanuts effect" , i.e. people's tendency to underweight small gains (Prelec and Loewenstein, 1991; Green and Myerson, 2004; Weber and Chapman, 2005). Research shows that people prefer a gamble over a certainty equivalent when the latter is a small gain. Conversely, most people have a higher valuation of a 10% chance to win $1 than of getting $.10 for sure. Higher stakes reduce the (unusually) high giving in Dictator Games, as long as the increase is substantial. For example, there is no effect from a difference between $5 and $10, but there is from $1.22 to $122 (see meta-study by Larney, Rotella, and Barclay, 2019).

We also consider the potential effect of introducing uncertainty on pro-social contributions. There is some evidence that more risk-averse people make more inequality-averse choices in the Dictator game (Van Koten, Ortmann, and Babicky, 2013). We do not worry about this too much since we are looking at the difference between retirees and non-retirees exposed to the same incentives.

The list of possible recipients of participant's earnings is designed to cover all possible consumption options. The option "give to relative or friend" is special in terms of both effect and interpretation. We hope that participants who want to keep earnings for themselves but are affected by social pressure (e.g. observability of their choices by the interviewer or the experimenter) would use this as a less-salient selfish option. Notice, however, that introducing this option comes with a tradeoff of more ambiguous interpretation of the monetary donations. Giving money to a relative in need has prosocial motivation underpinning it, whilst giving money to a family member to then derive private benefit from it does not. Hence, we focus on money kept to self as the main measure of prosociality in the analysis.

3.4 Results

3.4.1 Descriptive statistics

The survey data samples have been restricted to include individuals within a 10-year window around their country, survey year and gender -specific ERAs. This time window is used throughout the empirical analysis. Based on this sample restriction there are 121,182 individuals in the SHARE sample, and 58.2% of them is retired. Overall, 19.6% of the SHARE sample reported some voluntary work. Retirees appear to be more involved in voluntary work than non-retirees and the difference is statistically significant. For a more disaggregated information on prosocial behaviour, statistics on variables from the EU-SILC sample are provided. After applying the same sample restriction, 87,768 individuals are left in the sample with 47.1% of them being retired[13]. In the EU-SILC sample, 27.3% and 23.8% of the individuals report offering informal and formal voluntary work, respectively. Although higher, volunteering incidence among EU-SILC participants is comparable to the SHARE one.

Regarding informal voluntary work, EU-SILC retirees seem to be more involved in a statistically significant way. The incidence of formal volunteering and care provision to other people within the household is more balanced between retirees and non-retirees. Non-retirees are more likely to provide care to non-household members and be more active citizens.

The experimental data sample is fairly balanced in terms of labour market status. There is an indication of the retirees being less likely to offer formal volunteering than non-retirees ($p < .10$); the difference is negligible when considering informal volunteering (for the full descriptive statistics see Appendix D).

3.4.2 Results overview

The endogenous decision to retire is modelled as a function of whether an individual's age has crossed the legislated ERA threshold that applies in their case, as discussed in the identification strategy section. In this way, the discontinuous jump of the retirement probability at institutionally set age thresholds will allow the identification of the impact of retirement on prosocial behaviour indicators. Our first set of 2SLS and FE-2SLS results includes evidence of both survey samples that are quite comparable and indicate that there is a positive relationship between retirement and prosocial behaviour. The relationship is robust to sensitivity tests regarding own past prosocial activity, partner's volunteering activity and partner's retirement. Retirement positively affects the probability of volunteering, by around 10%, and the relationship is marginally stronger when considering formal volunteering.

The positive link between retirement and prosociality is also confirmed with the experimental data. Our second set of results demonstrates that retirees have higher prosociality in monetary donations along with the in-kind ones, captured by volunteering. This points towards the mechanism of increased prosocial preference post-retirement, rather than a mechanism related to increased time availability or preference for the same amount of labour.

[13] To further check that individuals identified as retirees have actually withdrawn from the labour force, the variable recording whether someone worked at least one hour during the week before the interview. In the EU-SILC sample 93.7% of retirees reported no hours of work last week. In a similar question 85.9% of the SHARE retirees sample reported no hours of paid work during last 4 weeks. Excluding those retirees reporting paid hours of work from the estimation samples does not affect the results.

3.4.3 First stage results and RD validity checks

Because retirement status is endogenous it is instrumented with the ERA indicator. This variable must be relevant and valid in order to be a suitable instrument. Throughout European countries, ERAs are exogenously set by the governments, hence validity cannot be formally tested. Therefore, ERAs are assumed to be linked to prosocial behaviour only through transitions to retirement. For the instrument to be relevant, a strong first-stage relationship between the endogenous variable and the instrument is required, i.e. the probability of being retired must be strongly predicted by the ERA indicator, I_{ic}.

Figure 3.1 scatters the shares of retired individuals across the ERA-normalised age (specific to year, country and gender) using: (a) EU-SILC (waves 2015 and 2016); (b) SHARE (waves 4-7); and (c) our experiment data. SHARE means lie slightly above the EU-SILC ones, because SHARE surveys individuals older than 50 years old and closer to their retirement age. The experiment data also follow these patterns, although in a bit noisy way due to a considerably smaller sample size[14]. The graph confirms that there is no perfect treatment compliance as some individuals retire before they reach their ERA while others stay in the labour market even after having crossed it. However, the share of retirees increases disproportionately at the eligibility age cutoff, providing reassurance about the instrument strength.

Panel A in Table 3.1 displays first-stage results from Equation 3.2 using the EU-SILC sample to support this claim. The probability of retirement increases by 28% when a local constant age function is specified (column 1). Using a linear age function at both sides of the ERA cutoff (column 2) suggests that the discontinuity in retirement probability is 31% higher for those having crossed their ERA. Discontinuities are also significant when quadratic and cubic age functions are specified at both sides of the cutoff. In all cases, the instrument relevance condition is satisfied. The retirement eligibility indicator strongly predicts retirement status, it is always statistically significant at the 1%, and the first-stage F-statistics of excluded instruments are sufficiently high. We get similar results using the SHARE sample, in Panels B and C.

The identifying assumption in an RD design is that individuals cannot manipulate the forcing variable around the cutoff age. Therefore, all observable characteristics should be balanced around the cutoff and individuals below it should be a valid control group for those above it, i.e. the treatment is considered to be as good as random (Lee and Lemieux, 2010). Examination of the forcing variable density around the cutoff can validate that local assignment could be considered as random. Appendix Figure E.2 displays a normalised age histogram within a time window of -/+ 10 years around ERA. The forcing variable is smooth around the cutoff age providing no evidence of forcing variable manipulation[15].

[14] The similarity between the experiment and the survey data lines is more evident if the EU-SILC and SHARE samples are restricted to only participants from Greece. Results available upon request.

[15] We also test if predetermined individual characteristics, i.e. gender, nationality and education, are locally balanced around the cutoff. Individuals around the cutoff should not be systematically different if the treatment is locally randomised. We obtained some 2SLS retirement estimates using predetermined covariates as outcomes, and focusing on a short time-window around the cutoff age. All parameters were not statistically significant indicating that treated and control individuals are balanced in terms of observables. Results are available upon request.

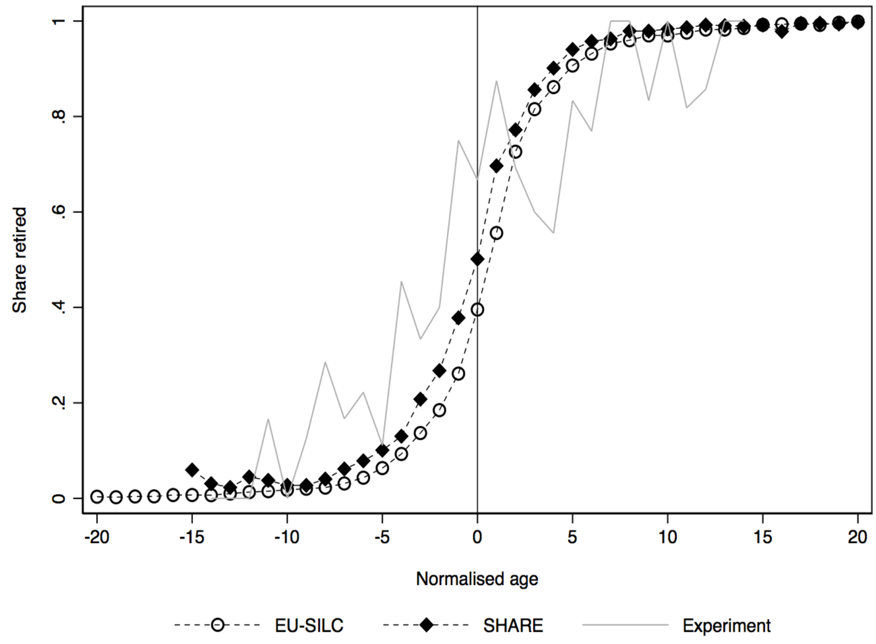


FIGURE 3.1: Share of retirees by age. *Source*: EU-SILC; SHARE; experiment. *Notes*: EU-SILC waves: 2015, 2016; SHARE waves: 4, 5, 6, 7; Experiment waves: all. Means are weighted by the survey (EU-SILC and SHARE) weights where relevant. Age is normalised by the country, survey year and gender-specific ERA of each individual.

3.4.4 Impact of retirement on prosocial behaviour

After establishing the existence of a strong first-stage relationship between retirement status and early retirement eligibility, we examine the impact of retirement on prosocial behaviour indicators. First, we make use of the detailed volunteering measurements of EU-SILC data to test the impact of retirement on the incidence of formal voluntary work, informal voluntary work and the overall voluntary work indicator (constructed using the formal and informal volunteering variables).

The 2SLS results suggest a positive and significant relationship (Table 3.1, Panel A). Retirement increases the probability of engaging into voluntary work by approximately 8% based on the local linear specification in column 2[16]. Distinguishing between informal and formal voluntary activity does not uncover any notable differentiation regarding the impact of retirement. Retirement increases the probability of informal voluntary work by 6%-8%, depending on how the forcing variable is specified. The incidence of formal voluntary work post-retirement increases by about 7%.

[16] The result remains positive when second order polynomials of the forcing variable are used, however it is less precisely estimated. Using higher-degree polynomials of the forcing variable returns much noisier estimates. Hence, models with local linear age functions will be used as the preferred specifications.

Although not reported here, active citizenship is not affected by retirement status. The same holds when considering provision of care to other people inside and outside the household. Providing care to household members only, is associated with a positive although not significant coefficient. Hence, these results are not provided here but are available upon request.

After establishing the effect of retirement across the range volunteering measurements, we turn to SHARE data, re-estimating Equation 3.1 in order to see whether the impact of retirement on voluntary activity is comparable between databases. (Panels B-C in 3.1). The SHARE sample is considerably larger than the EU-SILC one, and covers the period 2013-2017. The results of retirement on voluntary work are remarkably similar to those obtained using the EU-SILC sample[17]. Retirement increases the probability of voluntary work by about 9% (columns 1-3), depending on how the age function is specified. Moreover, because SHARE follows respondents over time, Panel C reports parameter estimates conditional on individual fixed effects. These should capture any unobserved time-invariant heterogeneity correlated to both volunteering activity and retirement status. The FE-2SLS results confirm the positive relationship between retirement and voluntary work, suggesting that retirement increases the probability of voluntary work by 10%-13%, depending on the local age function specification[18]. Models also control for individual characteristics, time of survey, country fixed effects, and for individual fixed effects (Panel C). Regressions are weighted using the relevant survey weights and they are estimated over a 10-year time window around the cutoff age[19].

TABLE 3.1: Retirement and prosocial behaviour: Evidence from survey data.

	[1]	[2]	[3]	[4]
	Panel A: 2SLS estimates; EU-SILC sample			
Retired (outcome: voluntary work)	.097** (.039)	.079** (.035)	.081 (.092)	.322 (.409)
Retired (outcome: informal voluntary work)	.088** (.034)	.061** (.030)	.067 (.081)	.121 (.348)
Retired (outcome: formal voluntary work)	.073** (.035)	.071** (.031)	.079 (.081)	.251 (.346)
First-stage: Age>ERA	.281*** (.010)	.311*** (.009)	.167*** (.011)	.132*** (.013)
First stage: F-statistic	847.83	1547.04	1431.96	1154.28
Observations	85,695	85,695	85,695	85,695

[17]The lists of countries covered by the two surveys overlap to a great extent. The EU-SILC estimation sample covers 22 countries and the SHARE one covers 19 countries (observations for Norway, Ireland and The Netherlands are not available). When the EU-SILC sample is forced to cover the countries covered by the SHARE one, the estimation sample reduces from 85,695 observations to 79,331 observations but the results remain practically the same. The retirement coefficient is .087 and the standard error is .037 (compared to the result in Table 3, panel A, column 2). A full list of the countries included in both samples is provided in the Appendix.

[18]All the baseline estimates are robust to the exclusion of retirees reporting hours of work in the last week (EU-SILC sample) or the last month (SHARE sample). More specifically, using the same bandwidth and a local linear age function, the 2SLS retirement coefficient is .076 (standard error = .036) in the EU-SILC sample. In the SHARE sample, the respective FE-2SLS parameter is .122 (standard error = .041).

[19]Individual survey weights have been adjusted based on the distance of each individual's age from the ERA threshold, so that individuals closer to it (from either side) are attached to a greater weight. However, results are robust to alternative weighting schemes.

Local age function	Constant	Linear	Quadratic	Cubic
Individual characteristics	Yes	Yes	Yes	Yes
Survey quarter fixed effects	Yes	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes	Yes
Panel B: 2SLS estimates; SHARE sample				
Retired (outcome: voluntary activity)	.087*** (.023)	.086*** (.024)	.092* (.048)	.114 (.113)
First stage: Age>ERA	.349*** (.008)	.338*** (.009)	.253*** (.010)	.229*** (.011)
First stage: F-statistic	1,705.33	1,025.08	1,823.05	1,409.93
Observations	121,182	121,182	121,182	121,182
Local age function	Constant	Linear	Quadratic	Cubic
Individual characteristics	Yes	Yes	Yes	Yes
Individual fixed effects	No	No	No	No
Country fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Panel C: FE-2SLS estimates; SHARE sample				
Retired (outcome: voluntary activity)	.100*** (.035)	.105*** (.035)	.127** (.060)	.135 (.117)
First stage: Age>ERA	.225*** (.009)	.227*** (.009)	.182*** (.010)	.170*** (.010)
First stage: F-statistic	609.59	304.81	404.82	314.19
Local age function	Constant	Linear	Quadratic	Cubic
Observations	98,840	98,840	98,840	98,840
Individual characteristics	Yes	Yes	Yes	Yes
Individual fixed effects	Yes	Yes	Yes	Yes
Country fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes

Source: EU-SILC; SHARE. Notes: Results are weighted using survey weights. Robust standard errors in parentheses. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

In sum, the results obtained from all sources of data using instrumental variables methods provide strong evidence to the claim that retirement increases the probability of volunteering, either formally or informally. A series of robustness and sensitivity tests confirms the validity of this finding (see section E.3 in the appendix).

3.5 Mechanisms

3.5.1 Survey data

To examine whether the impact of retirement varies with observable characteristics, we split the samples accordingly. The baseline effect of retirement holds at 8%-10% greater volunteering probability when several characteristics and fixed effects are controlled for. However, this probability increase varies with a number of factors. The relationship is stronger for females, tertiary educated, people with good self-reported health status and those not limited in their activities by health-related issues. There is no evidence that retirement impacts on other types of prosocial behaviour such as assistance or care to people outside the household, or active citizenship (see section E.1 in the appendix).

The impact of retirement is higher for people whose partners are also volunteers. For those whose partners volunteered in the same year, retirement increases the chance of own volunteering by 30% (see E.2 in the appendix). Also, the impact of retirement is higher for those who volunteered in the past. Controlling for dynamics, columns 3-4, confirms that previous volunteering experience is a very strong predictor of current activity (appendix section E.2). However, including a lagged dependent variable leaves the retirement status coefficient unaffected, compared to the static specification. Hence, this indicates an autonomous impact of retirement on the probability of offering volunteering work.

3.5.2 Experiment

After providing baseline evidence and robustness using the survey data, we turn to the experiment to investigate the source of the retirement effect. Is it pure altruism or a lower cost of time? As descriptives go, the sample is fairly gender balanced (111 males, 142 females) and the mean (median) age was 62.2 (63) years old. The majority (75%) completed the study over a phone interview, the rest opted into receiving a link over email and completing online. To preview the main result, in Figure 3.2 we present the unconditional mean amount that subjects kept for themselves, by group. Retirement indeed seems to have an effect, that is different (actually, opposite) to the raw effect of age. In the rest of this section, we seek to establish the effect econometrically.

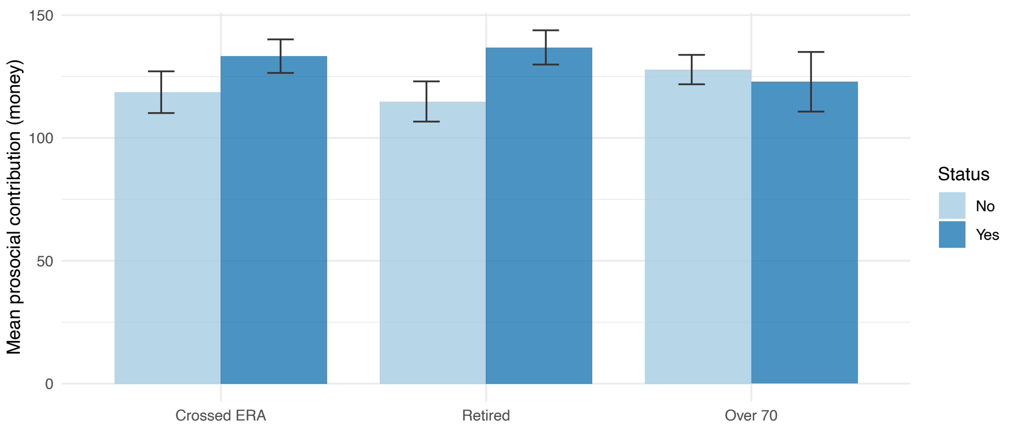


FIGURE 3.2: Mean money kept to self (+/- standard error), by age and ERA-crossed status.

Self-reported volunteering

First, we use the experiment data to replicate the survey-based evidence. About 27% (25%) reported having participated in formal (informal) volunteering in the past year. This is comparable with the EU-SILC sample, where

27.3% (23.8%) of respondents offer informal (formal) voluntary work. However, a limitation here is that the ERA threshold does not vary. Experiment participants come from a single country (Greece) and the threshold remained unchanged (at 62 years old) during data collection. In the absence of exogenous variation in eligibility thresholds that would allow to predict individual retirement status, we estimate the following model using OLS:

$$y_i = a_1 Post62_i + a_2 \widetilde{age}_i + u_i \tag{3.3}$$

where y indicates whether participant i volunteers or not, $Post62$ equals to 1 if the participant is older than 62 years old and 0 otherwise, and $\widetilde{age}$ is a local linear age function.In this framework, the coefficient a_1 gives the change in the probability of volunteering at the age of 62, and it is interpreted as the Intent-To-Treat (ITT) effect of early retirement eligibility on volunteering[20].

Table 3.2 confirms the survey-based evidence of crossing the ERA threshold being associated with an increased probability to volunteer. We find that being past the early retirement threshold increases the probability of having engaged in volunteering by over 20%. Moreover, consistently with the survey data results, the effect is driven by an increase in formal volunteering, estimated at 22-26% depending on specification – while retirees behave similarly to working individuals with respect to informal volunteering.

TABLE 3.2: Retirement and prosocial behaviour: Evidence from experiment data.

	Total sample	Excluding early retired	Excluding early retired & -/+ 15 years around ERA
	[1]	[2]	[3]
Crossed ERA (outcome: voluntary work)	.222*** (.081)	.231** (.089)	.238** (.092)
Crossed ERA (outcome: formal voluntary work)	.227*** (.076)	.259*** (.084)	.265*** (.088)
Crossed ERA (outcome: informal voluntary work)	.084 (.073)	.097 (.078)	.098 (.081)
Observations	255	224	214
Local age function	Linear	Linear	Linear
Individual characteristics	Yes	Yes	Yes
Wave fixed effects	Yes	Yes	Yes

Source: Experiment. Notes: OLS estimates. Robust standard errors in parentheses. Individual characteristics include gender, higher education and age controls. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Moreover, to test whether this systematically increases with age, we specify a number of placebo tests, by setting alternative ERA thresholds (from 55

[20]Several studies use age-based thresholds to uncover the intent-to-treat impact of policies on various outcomes. Fitzpatrick and Moore, 2018 adopted a similar framework to study the mortality effect of crossing the Social Security eligibility age. Results using the 2SLS estimator are comparable and available upon request.

to 71 years old) and adjusting the treatment and local age functions in Equation 3.3 accordingly. The results are in Figure 3.3. The probability of volunteering is not statistically different from zero away from the cutoff. It sharply becomes positive and significant for ages 61-63 years old, taking its highest value for the age of 62 (0.361; std.err = 0.096) which is the ERA threshold applying to the experiment participants. As these are the ages where the retirement probability also jumps disproportionately (Figure 3.1) we read those results as the ITT effects of retirement on volunteering.

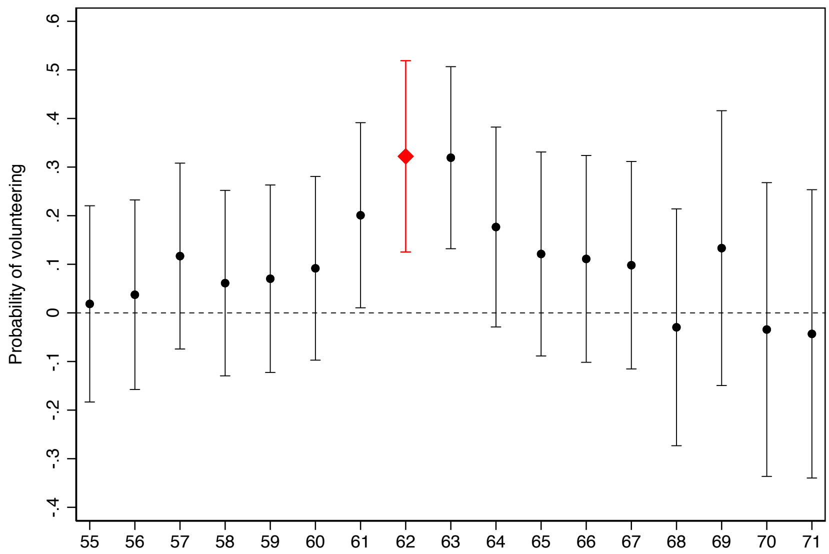


FIGURE 3.3: Intent-to-treat effect of retirement and placebo tests. *Source*: Experiment. *Notes*: The actual ERA threshold is set at 62 years old (red line). Thresholds around that are placebo ones and the model specification in Equation 3.3 has been adjusted accordingly. Outcome is volunteering (either formal or informal) and the sample excludes those early retired. Real ERA is set at 62 years old. Vertical lines represent 95% CIs based on robust standard errors. Black dots and vertical lines correspond to fake ERAs. All estimates are conditional to individual characteristics and wave fixed effects.

3.5.3 Incentive-compatible prosocial behaviour

Effort

Most of the participants chose to produce effort (to earn, gift or donate money) whilst a minority (n=47) refused to take part. Over three quarters (76.7%) of those produced effort chose to donate earnings to charity, whilst the rest of them (23.3%) opted to keep the earnings for themselves or give them to a friend or relative (Appendix Figure A4.1). Among the charities, the cancer research one attracted most in-kind donations overall and, also, highest effort intensity. The relationship between the effort intensity and earnings recipient was significant in the OLS analysis (model 1, Appendix Table A4.1). Overall, effort intensity was significantly lower among retirees compared to non-retirees, which is expected given that any task is likely to become more difficult with age. People who participated online, compared to those participated by phone, provided significantly more effort. We verified that retirees were not more likely to complete the study by phone (model 2, Appendix Table A4.1).

Money

Money donations were multimodal. The majority of participants chose to keep, gift or donate sums of €0, €100, or €200 (a handful of people did not participate in the lottery). On average, participants chose to keep €52.124 to themselves, but, as expected, there was a high variation in these amounts (s.d=80.075). We use winnings-kept-to-self as the main outcome variable, which is the reverse of winnings-given away[21] In line with the pattern observed in the in-kind donations above, the cancer research charity attracted most monetary donations compared to other charities, as well as the highest shares of the total lottery pot (Appendix Figure A4.2).

Retirement and prosociality

Next, we test if retirement increases prosociality by comparing the amount of winnings that retirees and non-retirees intend to keep for themselves (as opposed to donating to charities or family/friends). A decrease in money-to-self would indicate that retirement leads to increased prosocial gifting across both in-kind and monetary domains. Results in Table 4 support the hypothesis of increased prosociality driving the increased volunteering at retirement. Those eligible for retirement keep on average between €40-€56 less to themselves, compared to those not eligible (columns 1-3). These choices of monetary gifts are also significantly associated with choices of the in-kind donation (column 4). The experimental subjects kept on average €41 in expected lottery winnings for every €1 kept of real-effort task earnings. Of course, the real-effort earning of up to €5 allows for value-signalling, hence we are careful to not interpret this result beyond the outcome consistency.

TABLE 3.3: Monetary and in-kind contributions upon retirement: Evidence from experiment data.

	Total sample	Excluding early retired	Excluding early retired & using -/+ 15 years around ERA	
	[1]	[2]	[3]	[4]
Crossed ERA	-39.718** (17.224)	-46.034*** (18.639)	-55.997*** (19.788)	-57.590*** (18.151)
Real-effort earnings (in €) kept to self	-	-	-	41.614*** (3.348)
R-squared	0.059	0.06	0.065	0.24
Observations	249	219	209	209
Local age function	Linear	Linear	Linear	Linear
Individual characteristics	Yes	Yes	Yes	Yes

[21] We also consider an alternative outcome variable which is the difference in money-to-charity and money-to-self which yields similar results.

Wave fixed effects	Yes	Yes	Yes	Yes

Source: Experiment. Notes: OLS estimates. Outcome is lottery winnings (in €) kept for self. Individual characteristics include gender, higher education and age controls. Robust standard errors in parentheses. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively. a €0-€5.0, calculated based on 28 units of real-effort task with €5.0 maximum earnings.

Overall, the data provided consistent evidence that transition to retirement (and retirement eligibility) increases the likelihood of engaging in prosocial activities. Moreover, this increase is due to enhanced prosociality rather than merely due to having more free time. Given that (a) these activities have been shown to be beneficial for both individual well-being and societal welfare, and (b) population ageing will intensify, policy interventions should aim at increasing participation in post-retirement unpaid work.

3.6 Concluding Remarks

Retirement is one of the major single lifestyle shifts in most people's lives, yet not much is known about its effect of preferences. A possible reason for this knowledge gap is the endogeneity of the retirement choice, combined with the fact that many relevant factors change synchronously to retirement. Lab experiments would be optimal to control important factors, but simulating the retirement experience is hardly feasible in the lab (given the major changes in income and passage of time involved). Our paper addresses this by sampling directly from the population of people who experienced retirement and tapping directly into the preference change. Using large scale survey data from European countries we find evidence of volunteering increasing substantially upon retirement. We build on this evidence by using an incentivised experiment to uncover the causal links in this behavioural change.

In the literature, social preferences are usually assumed to be stable, as is the case with any fundamental economically relevant trait. Any preference shifts that have been shown to date are driven by major negative lifestyle shocks, e.g. military conflicts, that are not common for most populations. This paper is the first to identify the impact on preferences of a mild but globally relevant shock, retirement. Combining large scale survey data and controlled, incentivised experiments we separate the effect of retirement from other factors, like age, on two manifestations of prosociality: volunteering and cash donations. We find retirement does not just lead to people donating more of a resource that they have plenty of, time. Retirees also donate more in cash, although retirement presumably lowers their endowment in that dimension. All things considered, retirement seemingly *makes you a better person*.

Given the continuing population ageing in the West, but increasingly other countries too, our findings are policy relevant. Reforms aiming to change (almost without exception, raise) the average retirement age have to take into consideration the effect on overall welfare. Even though economists often attach a zero wage to various prosocial activities, e.g. volunteering, the implications of raised prosociality for the economy are substantial. Survey evidence suggests that these activities represent a considerable part of the overall labour supply and generate value. Although disengaged from paid employment, retirees can have significant contributions to the public good and reduce social costs through such activities, apart from the benefits on their own well-being. Our evidence on time and effort allocation decisions of retirees, as well as their productive capacity, suggest that the gain for pension systems from higher retirement ages, will come at a significant loss for members of society benefiting from retiree volunteering, as well as from direct money transfers. Raising the retirement threshold does not only increase the working proportion of the population, it also decreases the retired and volunteering proportion of population - which should be considered.

In studies like this, there is always a subtle balance to strike between external validity and uncovering causality. In the experiment we find strong evidence for a change in preferences, in survey data we find that this likely extends across countries and cultures. More research is needed to investigate how exactly the retirement effect changes across cultures, and also what the exact goals of retiree charity are. In our sample retirees donated substantially to cancer related charities but less for refugees and the environment (see Appendix Figure A10), indicating that their preferences might be influenced in part by awareness of issues.

A major question remains as to *why* retirement would make you nicer to others? Is this shift driven by feeling happier about shifting into a more relaxed part of life or is it reciprocal to the benefits retirees are now receiving, from the working population? This question remains open for future research, but we conjecture that retirees benefitting from pension plans with higher replacement rates (i.e. people who are presumably getting more back from society comparing to what they gave) would be exhibiting a stronger pro-social effect.

Identifying how exactly retirement leads to preference change, can also help identify other shocks that affect preferences. Transition to unemployment, for example, is similar to retirement in that people experience a lifestyle shift, with more free time and lower income. However, this shock is often involuntary, unanticipated and mostly perceived by the individual to be unfair. Do unemployed people volunteer more? If unemployment makes people more miserable, does it lead to lower prosociality and, if so, would this effect be mitigated by more generous unemployment benefits? These questions remain open for future research.

Chapter 4

Who is miserable now? Identifying clusters of people with the lowest subjective wellbeing in the UK

4.1 Introduction

Over the past decade or so, there has been increasing interest in going beyond standard economic measures of welfare, such as income, to consider wellbeing in a broader sense (Stiglitz, Sen, Fitoussi, et al., 2009). Among these efforts is research which uses subjective wellbeing (SWB) measures i.e. reports about how individuals think and feel about their lives. This work typically examines the relationship between a single measure of SWB and a range of characteristics. Key findings from the literature include a substantive negative association between unemployment and life satisfaction (Knabe et al., 2010), as well as a strong association between health and a range of measures of SWB (Steptoe, Deaton, and Stone, 2015).

Whilst this regression-based approach has yielded important insights, it presents an incomplete picture of limited use to policy. First, it ignores the multidimensionality of wellbeing (Dolan and Kudrna, 2016), typically by focussing only on life satisfaction and neglecting day to day experienced wellbeing. Second, it provides insights into "average Joe's" wellbeing (Binder and Coad, 2011) and is silent on whether the main determinants of wellbeing differ for those of greatest policy concern, namely with the lowest SWB. Third, it focuses on single determinants of wellbeing ceteris paribus, thus ignoring heterogeneities across different groups and interactions between determinants. Against this background, the current work fills an important gap by: first, using a range of measures of SWB; second, focusing on those who report the lowest SWB across all measures; and third, employing Latent Class Analysis (LCA) to partition the sample into groups (classes) most similar to each other in terms of life circumstances (age, health, socio-economic status, SWB etc.).

According to our multidimensional definition of misery, we find that just over 1% of the APS sample are miserable, a substantively smaller group than the almost 5% who report low life satisfaction. Our LCA results suggest that the overall sample can be summarised into seven main groups. We find that two of the seven classes have an above-average probability of being the most miserable: 1) unemployed/inactive people over 30 with severe health problems and/or a disability, who live in rental accommodation and are not in a partnership and have up to compulsory level education; and 2) 16-59 year olds, who are in employment and have GCSE or above compulsory level education but are facing some issues with health and disability, are not in a relationship, and who live in rental accommodation or have a mortgage. Policymakers interested in improving the lives of the worst off in society should pay special attention to these two groups as they account for 96% of the miserable in our sample. The rest of this Chapter is organised as follows. Section 2 further details the background to the study. Section 3 describes the data and presents our methods. Section 4 presents the results, Finally, Section 5. discusses the findings in the context of the existing SWB literature.

4.2 Background

4.2.1 What measures?

In the current work, we take a person-centred approach to examining the worst off in society, defining the worst off according to a novel definition of misery that spans different dimensions of SWB. In so doing, we build on existing literature in terms of what measures, who matters, and in what ways?

Existing research typically focuses on single measures of SWB and SWB is an umbrella term for how people think about their lives and feel as they go about them (Diener, Lucas, Oishi, et al., 2002). An important distinction in the literature is between evaluative measures, which elicit global reports of happiness or, more commonly, life satisfaction, and experience-based measures, which elicit more granular reports of happiness in the moment (Dolan and Metcalfe, 2012). Some measures of SWB also tap into a "eudemonic" account of wellbeing, which assesses the purpose or meaning people have in their lives or experiences (see Dolan and Kudrna, 2016, for how to categorise the different measures).

Most large-scale surveys around the world solely elicit evaluations of life satisfaction (e.g. World Values Survey, General Social Survey, German Socio-Economic Panel, and Understanding Society) although some additionally capture reports of SWB in the moment (most notably the American Time Use Survey). As such, these surveys provide only a partial picture of SWB. When the Office for National Statistics (ONS) in the UK began measuring wellbeing in 2011, it took the decision to elicit reports of evaluative, experiential and eudemonic wellbeing, thus facilitating a more comprehensive assessment of individual's SWB (Dolan and Metcalfe, 2012). The four questions used are as follows:

1. Overall, how satisfied are you with your life nowadays?
2. Overall, to what extent do you feel the things you do in your life are worthwhile?
3. Overall, how happy did you feel yesterday?
4. Overall, how anxious did you feel yesterday?

The responses to all four SWB questions are measured on a 0-10 scale where 0 is "not at all" and 10 is "completely". The ONS data allow for a multidimensional approach to modelling individuals' SWB which spans evaluative and experiential measures, as well as both hedonic and eudemonic wellbeing types of wellbeing (Dolan and Kudrna, 2016). However, being restricted to four survey measures, the set does not tap into all of the dimensions of SWB that many would consider relevant, including sad or joyous moments, which would arguably be best assessed using naturalistic monitoring tools, or evaluations such as the meaning of one's life overall (Stone and Mackie, 2013). As a result, the ONS measures can be understood as provided a richer but still incomplete picture of SWB in comparison to that offered by many large scale surveys.

4.2.2 Who matters?

Previous research has tended to look for average effects, treating a population as if it was a single representative citizen (Oswald et al., 2003). Some limited work using quantile regression techniques, however, demonstrates that looking at averages alone provides an incomplete and, at times, misleading picture of the relationship between SWB and determinants of interest at different points in the wellbeing distribution. For example, Binder and Coad, 2011 find that education is positively associated with happiness at the bottom end of the wellbeing distribution but negatively so at the top, and Lamu and Olsen, 2016 find that both income and health are relatively more important at the lower end of the wellbeing distribution.

Those with the lowest wellbeing will be of more concern to policymakers than the average citizen. Most "common sense conceptions of justice" are seen to lie somewhere between the utilitarian social welfare function (SWF), which solely emphasises improving overall welfare as much as possible irrespective of its distribution, and the Rawlsian SWF, which focuses only on improving the welfare of the worst-off individual and disregards how efficiently resources might be used elsewhere (Sugden, 1993; see also Dolan and Tsuchiya, 2011 for an empirical investigation of the SWF). The wellbeing of the worst off is therefore a policy concern on distributional (equity-related) grounds.

Understanding the SWB of the worst-off also matters for efficiency reasons as it provides insights into how best to target resources at the worst-off. As existing research indicates that the determinants of SWB for at the lower end of the distribution differ to those at the top, research comparing the determinants of SWB across the distribution can highlight the different strategies aimed at improving wellbeing at different points. Such research may also highlight the potential differential feasibility or costs of improving the wellbeing of the worst off compared to those who already have high levels of SWB. If, for example, the determinants of wellbeing at the bottom of the distribution are such that policy interventions are likely to have limited potential impact, i.e. that the worst off are inelastic suppliers of wellbeing, then it is important for policymakers to be aware of this.

The ONS defines low wellbeing on each question according to the underlying distribution of the data. For the three positively framed questions, a score of four or less is deemed as low. For the anxiety question, a score of six or more is defined as having high levels of anxiety and therefore low wellbeing according to this measure. Whilst the research community, policymakers, and individuals might disagree about which the best measure of SWB is, they will surely all agree that someone doing badly on all four questions has low wellbeing. Someone who reports low SWB on all four ways of tapping into it is clearly doing at least as badly as someone who reports low SWB if they were only asked one of those questions, and arguably worse. A person who reports being both dissatisfied with their lives overall and as well as experiencing unhappiness day to day is more miserable than their counterpart who is dissatisfied but relatively happy day to day. Moreover, measuring SWB is a fuzzy concept: an individual has no objective indication of where to rate themselves on the scale (for a thorough discussion of fuzzy sets, and empirical examples of how much trouble people have even with notions of "tallness" and "beauty" see Norwich and Turksen, 1984). Consider, for instance, an individual who is repeatedly asked to evaluate the same stimulus, e.g. SWB on a given day. Empirical evidence shows that 75% of the time she may rate it as a "5" and 25% of time as a "4". Hence, one-quarter of occasions she would fall into the miserable subset, although three-quarters of the time she is above the cut-off point. Defining misery over four questions creates a stricter criterion for identifying the individual as a member of the miserable population: an individual with the same response pattern would only be misclassified as miserable $(1/0.25)^4 = 1/256$ percent of the time (assuming that the inconsistency in reporting is perfectly random and uncorrelated between measures). A stricter definition of misery means that policy-makers can be more confident in identifying the worst off in society by mitigating noise in who they categorise as the most miserable (be that due to the fuzziness of the concept or simply the mistakes people may make when filling in a form). In this setup, the strategy of defining misery through $n + 1$ measures weakly dominates defining it through n measures. In other words, the stricter definition yields at least as good an outcome as the more lenient definition; and when individuals may make unintentional mistakes, it yields a superior outcome (for further discussion of strategies in response to such "trembling hand" errors see Selten, 1975). On this basis, we posit that identifying the worst-off individuals through all available SWB measures is the best strategy for examining misery.

Importantly, as the ONS bases their definitions of low wellbeing for all four measures on the underlying wellbeing distribution it is worth highlighting that this group represents the worst off in the population – a relative rather than an absolute measure of low SWB. As such a subgroup of the population will always fall into this category regardless of improvements to wellbeing, much like there are always people living in relative poverty. Our focus on people who report low wellbeing across all four measures acknowledges the multifaceted nature of SWB and hones in on those that are doing badly across evaluative, experiential and eudemonic dimensions.

4.2.3 In what ways?

The extant literature on SWB focuses on each determinants of wellbeing ceteris paribus and so it ignores "clusters" of determinants that interact in important ways with one another. What predicts the SWB of a middle-aged man, for example, may be different to what predicts the SWB of a retirement-age woman, and the relationship between SWB and being in poor health might vary greatly depending on a person's level of income. Furthermore, there may be an unobserved factor (such as, personality) that affects both the observable characteristics – such as health and socioeconomic status – and SWB. Although some SWB work does present subgroup analysis and interactions between specific factors, the challenge of treating the individual in a holistic way remains largely unaddressed in the SWB literature (Clark et al., 2005).

The current work uses LCA to identify clusters of people within society that report low wellbeing. LCA is increasingly used to deal with the challenges of heterogeneity and endogeneity by allowing the latent (unobserved) characteristic to partition the data into clusters united by combinations of observable characteristics (see e.g. Anand, Krishnakumar, and Tran, 2011; Brown et al., 2014; Clark et al., 2005; Fernandez-Blanco, Orea, and Prieto-Rodriguez, 2009; Giovanis, 2014). LCA splits respondents into homogenous groups (latent classes) such that individuals in the same latent class will have similar response patterns to the independent variables whilst individuals across latent classes will have different response patterns to each other. The relationship between being in a certain class and an outcome of interest (such as being among the worst off in terms of SWB) can then be examined. This approach is consistent with our contention that it is important, for academic and policy purposes, to think about SWB of whole people rather than exploring specific determinants ceteris paribus.

4.3 Data and Methods

4.3.1 Data

This paper uses the pooled[1] three-year dataset of the ONS's UK representative APS[2]covering the period April 2014 to March 2016[3].The dataset contains observations for 284,456 adults, 280,003 (over 98%) of whom responded to all four SWB questions.

4.3.2 Methods

Latent Class Analysis

In step one, we perform Latent class analysis (LCA). This allows us to examine unobserved heterogeneity; that is, to identify groups of participants who represent the greatest similarity on the same set of observed variables within a given group and the greatest dissimilarity between other participants' LCA is a type of a finite mixture model which makes it particularly well-suited for categorically scored data and variables with different scale types (Vermunt and Magidson, 2003). For a more detailed comparison of alternative clustering approaches, see Appendix H. LCA has been applied in research concerning a wide range of outcomes such as self-reported consumer taste (Fernandez-Blanco, Orea, and Prieto-Rodriguez, 2009)[4], financial satisfaction across life stages (Brown et al., 2014), and the relationship between self-reported well-being and recycling in Britain (Giovanis, 2014).

Here we assume that a latent variable of "person's life circumstances" (the X in Figure 4.1) defines the combination (A, B, C, D in Figure 4.1) of a person's socioeconomic status (their employment status, education, etc.), their reference group (age, marital status, etc.) and the way they perceive their life (SWB and self-reported health). For full details of the model specification and the intuition behind the method, see Appendix **??**. We assume the latent class to be "person's life circumstances" and use as the manifest variables the major socio-economic and personal characteristics available in the dataset and traditionally used in SWB literature. These are age, sex, health and disability status, SWB, employment status, socio-economic status, housing tenure, marital status, education, and income.

[1]Note that the dataset does not contain the time variable, as it is pooled and intended to be treated as point-in-time, according to the ONS guidelines. This does not present a challenge to our design: since the dataset is pooled from independently representative waves, we do not consider it to be an issue even if the overall SWB changes with time. Each person is observed once and the structure of nationally representative selection does not change.

[2]APS combines results from the Labour Force Survey (LFS) and the English, Welsh and Scottish Labour Force Survey boosts with the aim of providing estimates between censuses of main social and labour market variables at a local area level.

[3]No individuals are included more than once in the dataset.

[4]Although more commonly used in panel data – to observe whether classes recovered remain consistent over time variable.

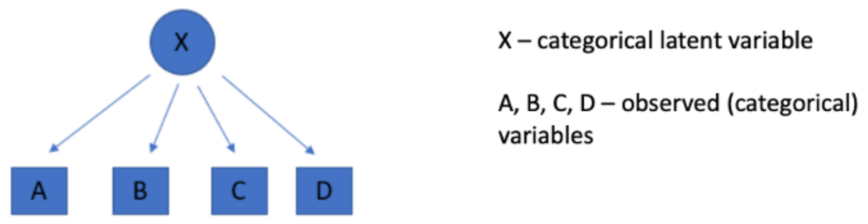


FIGURE 4.1: A generic path diagram for an unconditional latent class model.

The LCA model assumes conditional independence. For two independent categorical variables – A (with categories) and B (with categories), the joint probability of observing both responses should be equal to the product of the probabilities of observing each response: $Pr(Y_{11} \cap Y_{jk}) = Pr(Y_{11}) \cdot Pr(Y_{jk})$.

We recognise that conditional independence of the observable characteristics is a rather bold assumption for our dataset. From the technical point of view, however, the literature agrees that the tenability of the local independence assumption may also be partially relaxed (see, e.g., Huang and Bandeen-Roche, 2004). More importantly, from the conceptual point of view, it fits with our goal of being able to step back from the well-established negative correlations between SWB and bad health or unemployment, for example, and explore if there are groups of these factors that are associated with misery in combination.

In step two, we then examine the proportion of miserable people in each of the classes emerging from the LCA. It is common in LCA to use the model-driven partition to then compare the emergent classification on the variables used in the model (e.g. to verify significant differences between classes) as well as the variables not used in the clustering step (e.g. to explore further differences between classes)[5]. Whilst we favour looking at the intersection of all four measures in our definition of misery, we also examine an alternative threshold of low life satisfaction. In the results, we report the observable characteristics of individuals belonging to each class, together with the proportion of those in misery, for these two different thresholds. This allows us to achieve our main goal of identifying observable characteristics of individuals that are more likely to be miserable without making claims on what exactly causes it.

In comparison to regression analysis, LCA is person-centred, looking at groups of factors rather than individual variables, and allows for the ranking of groups against thresholds of interest, e.g. proportions of misery. It is also particularly suitable for reducing multidimensionality in the data (Masyn, 2013). For comparison, see Appendix **??**, which present binary logistic regressions predicting misery.

[5] See, for example, typology of high school dropouts (Bowers and Sprott, 2012), detection of Internet-addicted and at-risk Internet-use groups in general population sample (Rumpf et al., 2014), self-reported anxiety and depression in general population sample (Lang et al., 2006).

4.3.3 Model Selections

We start by using all the variables traditionally used in the SWB literature: age, gender, SWB, health and disability status, employment status, socio-economic status, housing tenure, marital status, education, and income. Whilst household income is not reported for most respondents, we retain this variable with the inclusion of the "NA" category. We interpret it in conjunction with the housing tenure which is commonly used as a proxy for income[6]. In choosing the optimal number of classes, we look to Bayesian information criterion (BIC) statistic and its variations (ABIC, CAIC, Chi-squared and Likelihood Ratio), which is commonly used to balance the gain in log-likelihood through an increase in the number of classes and the loss of degrees of freedom from the greater number of parameters (Lin and Dayton, 1997). Thus, a lower value of the information criterion suggests a better balance between model fit and parsimony.

A good latent class indicator is one for which there is a strong relationship between the item and the latent class variable. Strong item-class relationships must have a particular item response (e.g. item endorsement in the case of binary items, epitomises members in at least one of the K latent classes in the model) and the item must be able to be used to distinguish between members across at least one pair of classes among the K latent classes in the model. The first quality is referred to as latent class homogeneity[7] and the second quality is referred to as latent class separation[8] (Collins and Lanza, 2009). Gender, sexual orientation, ethnicity and religion did not present a separation into classes above a 70% probability threshold, but we do retain gender as a control variable.

To establish the appropriate class number, we take 50 random samples of 60'000 (~33% of total sample) and run the clustering code (*poLCA* package, R version 3.6.3) for a number of classes (n) from 1 to 10 on each of them. For each run, we set the number of repetitions (*nrep*) to 30 and maximum iterations (*maxiter*) to 4000. A high number of repetitions and iterations allows the model to re-start from new random initial values which is crucial for finding the global rather than local maximum. We then compare the values of the information criteria across the set of fitted models and selected the model with $n = 7$ classes, that had the lowest value, BIC, which in our case was consistent across the criteria (see Figure K.1 and Table K.1 in the Appendix K).

[6]Housing tenure is often used as a proxy for income, in the UK in particular there is a strong association between the two (see e.g. Macintyre et al., 1998)

[7]For example, consider a class with an estimated class-specific item probability of 0.90. This means that in that class, an estimated 90 % of individuals will endorse that particular item whereas only 10 % will not. You could then consider this item endorsement as "typical" or "characteristic of" that class and could say that class has high homogeneity with respect to that item. Now consider a class with an estimated class-specific item probability of 0.55. This means that in that class, only an estimated 55 % of individuals will endorse that particular item whereas 45 % will not.

[8]It is possible to have high class homogeneity and still have low class separation. For example, consider two classes, one of which has an estimated class-specific item probability of 0.90 and another class with an estimated class-specific item probability of 0.95. In this case, since item endorsement is 'typical' for both of these classes and the two classes can be characterized by a high rate of endorsement for that item, they are not distinct from each other with respect to that item. Now consider two classes, one of which has an estimated class-specific item probability of 0.90 and another with an estimated class-specific item probability of 0.05. In this case, each class has good homogeneity with respect to the item and they also have a high degree of separation because the first class may be characterized by a high rate of item endorsement whereas the other class may be characterized by a high rate of item non-endorsement.

To ensure that the classes we find represent naturally occurring subgroups in the population rather than being a sample specific statistical artefact, we conduct clustering on multiple random samples ($n = 40k, 50k, 70k$) from the dataset, to ensure consistent separation into classes of same sizes and characteristics. Given that the same classes appear consistently when conducting the same analysis with multiple subsets of the same sample, the classes are considered reliable (Bauer and Curran, 2004; Lenzenweger, 2004). For full details see Appendix K.

4.4 Results

4.4.1 Descriptive Statistics

Looking at the percentage of the sample reporting low wellbeing across the different dimensions of SWB included in the APS, we see that 5% of the sample report low life satisfaction, 3.8% report low worthwhileness, 8.9% report low happiness yesterday and 20% of the sample report high anxiety yesterday. That only 1.1% of the sample reports low wellbeing across all four measures highlight the strict nature of our definition and suggests that those identified as miserable according to all four measures in combination are unlikely to represent false positives. Importantly, it does however also highlight that there are many people experiencing low SWB that will not be captured using this definition due to their not fulfilling all four criteria. Full descriptive statistics for the entire sample and for the miserable subset can be found in Appendix I.

4.4.2 LCA Results

Table 4.1 provides conditional item response probabilities by outcome variable for each class. The columns represent the latent classes. The model assumed 7 latent classes in this case, therefore there are 8 columns. The rows indicate the categories of each indicator variable. The table shows the probability of having a given life circumstance conditional on belonging to the different classes. For example, 99% of respondents in class 5 are aged 60+, compared to only 3% of respondents in class 4, and 34% of the sample average. This difference from the sample average can be interpreted as one of the defining characteristics of the class. For ease of interpretation, we highlight the proportions on which the classes differ most from the dataset average in grey. For example, class 5 contains predominantly individuals over 60 who are retired and homeowners. LCA also allows us to observe the estimated size of each class in relation to the sample (bottom row in Table 4.1). For example, class 3 makes up 11.2% of the sample, and these individuals are predominately self-employed, in good health, and three-quarters of them have above compulsory education.

4.4.3 Applying the misery threshold

According to the definition of misery using all four ONS wellbeing questions and the ONS thresholds (0-4 on Happy, 0-4 on LS, 6-10 on Anxious, 0-4 on Worthwhile), in the three-year dataset 3076 individuals (out of 280,003) or approximately 1.1% fall in this subgroup. Looking at another possible threshold, when we define misery as low LS in isolation, just under 5% of the sample is miserable (Table ??). This comparison highlights the strictness of our preferred definition of misery.

TABLE 4.1: The probability of belonging to a given latent class on each of observable characteristic

			Latent Class						
		Sample Average	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Satisfied	High (7-10)	**0.81**	0.32	0.26	0.87	0.8	0.89	0.95	0.95
	Medium (5-6)	**0.14**	0.38	0.56	0.11	0.16	0.09	0.05	0.04
	Low (0-4)	**0.05**	0.31	0.18	0.02	0.03	0.01	0	0
Worthwhile	High (7-10)	**0.84**	0.43	0.44	0.9	0.85	0.9	0.95	0.96
	Medium (5-6)	**0.12**	0.32	0.44	0.08	0.13	0.09	0.05	0.04
	Low (0-4)	**0.04**	0.26	0.12	0.01	0.02	0.01	0	0
Happy	High (7-10)	**0.75**	0.35	0.25	0.81	0.77	0.85	0.86	0.84
	Medium (5-6)	**0.16**	0.29	0.42	0.14	0.16	0.12	0.11	0.13
	Low (0-4)	**0.09**	0.36	0.34	0.06	0.07	0.04	0.03	0.03
Anxious	Low (0-3)	**0.64**	0.33	0.29	0.68	0.64	0.72	0.73	0.7
	Medium (4-5)	**0.16**	0.21	0.27	0.16	0.17	0.14	0.13	0.16
	High (6-10)	**0.2**	0.46	0.44	0.17	0.19	0.15	0.13	0.15
Age	16-29	**0.2**	0.07	0.26	0.12	0.56	0	0.34	0.24
	30-59	**0.45**	0.52	0.65	0.64	0.41	0.01	0.53	0.69
	60+	**0.34**	0.41	0.09	0.24	0.03	0.99	0.13	0.07
Disability	No	**0.73**	0.06	0.67	0.86	0.84	0.61	0.88	0.92
	Yes	**0.27**	0.94	0.33	0.14	0.16	0.39	0.12	0.08
Health	Good/Very Good	**0.74**	0.06	0.6	0.86	0.85	0.65	0.87	0.92
	Fair	**0.19**	0.36	0.31	0.13	0.14	0.29	0.12	0.07
	Bad/Very Bad	**0.08**	0.58	0.08	0.01	0.01	0.06	0.01	0
Econ. Activity	Employee	**0.48**	0	0.99	0	0.01	0	1	1
	ILO Unemployed	**0.03**	0.06	0.01	0.02	0.25	0	0	0
	Inactive	**0.08**	0.14	0	0.1	0.51	0.03	0	0
	Inactive (LT sick/disab)	**0.05**	0.55	0	0	0.01	0.01	0	0
	Retired	**0.25**	0.22	0	0.06	0	0.96	0	0
	Self-employed	**0.09**	0.02	0	0.82	0.01	0	0	0
	Student	**0.02**	0	0	0	0.21	0	0	0
	Managerial/ Professional (H/L)	**0.29**	0.06	0.3	0.4	0.05	0.13	0.06	0.76
	Intermediate/ Lower Superv	**0.18**	0.08	0.31	0.01	0.14	0.09	0.4	0.19
	Semi/ Routine	**0.18**	0.16	0.37	0	0.23	0.09	0.49	0.05

Socio-Econ Status	Small Employer/ Own Acc	**0.08**	0.05	0	0.59	0.01	0.03	0	0
	FT Student	**0.03**	0	0.03	0	0.24	0	0.05	0
	Never worl/ LT Unempl	**0.03**	0.1	0	0	0.19	0.01	0	0
	NotClassif (Retired)	**0.21**	0.55	0	0	0.15	0.64	0	0
Housing Tenure	Owner	**0.36**	0.23	0.17	0.36	0.14	0.8	0.23	0.19
	Mortgage	**0.33**	0.12	0.4	0.44	0.23	0.04	0.36	0.66
	Rent	**0.31**	0.66	0.43	0.2	0.63	0.16	0.41	0.15
Income	Above median	**0.21**	0	0.33	0	0	0	0.1	0.74
	Below median	**0.21**	0	0.53	0	0	0	0.78	0.15
	NA	**0.58**	1	0.14	1	1	1	0.12	0.11
Marital status	Single	**0.27**	0.29	0.4	0.21	0.57	0.05	0.38	0.26
	Married/ Partnership	**0.51**	0.3	0.34	0.62	0.34	0.59	0.44	0.62
	Divorced/ Dissolved/ Widowed	**0.22**	0.41	0.26	0.17	0.09	0.35	0.17	0.12
Education	Above compulsory	**0.56**	0.3	0.65	0.76	0.55	0.22	0.52	0.91
	GCSE	**0.19**	0.23	0.27	0.18	0.31	0.08	0.36	0.09
	Basic/None	**0.09**	0.34	0.08	0.06	0.13	0.09	0.11	0
	NA	**0.16**	0.14	0	0	0	0.62	0	0
Class population	Estimated		8.6	7.3	11.2	9	23.5	17.3	23
shares (%)	Predicted		8.5	6.9	11.2	8.9	23.7	17.6	23.2

On average (before any classes are considered), a person in our sample had about 1/100 chance of being miserable. Once the classes are considered, the risk of misery can be roughly organised into 3 tiers which we summarise in Table 4.2, along with the accompanying combinations of characteristics. Both classes 1 and 2, have an above-average proportion of miserable people. Together they make up 15.4% of the sample, but account for 96% of the total miserable group. Classes 3 and 4, consisting of 20% individuals, have a below-average but non-negligible proportion of miserable people. Finally, classes 5-7, comprising over 63% of the sample, contain almost no miserable individuals. Importantly, we do not claim that the combination of characteristics that each class represents are causally related to SWB.

TABLE 4.2: The representation of the miserable among the overall sample and the seven classes

			Latent Class						
		Sample Average	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Miserable (%)	**All 4 SWB measures**	**1.1**	10	2.96	0.23	0.15	0.02	0	0

Life Satisfaction	**4.9**	31.09	19.14	2.34	3.36	1.22	0.16	0.08

4.4.4 Comparing misery on all four to low life satisfaction

About 5% of the sample falls under the “low” threshold on life satisfaction measure, while only 1% also fall beyond the threshold on all four measures (see Table 4.2). When we compare the order rankings of proportions of the miserable using our definition of misery to those produced using low life satisfaction, they are largely consistent. While the two first classes are the most vulnerable to misery across the different definitions, misery on LS dimension ranks class 4 above class 3, with the proportion difference being significant between these two classes. Whilst we are careful to avoid over interpreting this exploratory analysis, this result is suggestive of a difference in priority ordering of most vulnerable groups depending on the chosen definition of misery.

We also observe the ratio of miserable by the LS definition to miserable on all four SWB measures increasing from class 1 to class 7. For example, class 1 contains three respondents reporting low LS to each one reporting misery on all four SWB scales. This ratio increases to over 6/1 in class 2, 10/1 in class 3, and so on. This suggests that the two definitions of misery are closely aligned for the classes containing the largest proportion of individuals with low SWB. As the proportion of individuals with low SWB falls, the definitions diverge, with the proportion of miserable on all four SWB scales falling at a much higher rate than the proportion of people reporting low LS.

TABLE 4.3: The representation of different classes among the miserable

Class	Description	Class Population (out of 100%)	Proportion miserable in the class	Low LS
1	Age 30-60+, fair/bad/very bad health, disability, not economically active with long-term sickness/disability, rented accommodation, single or divorced/dissolved/ widowed, basic or up to compulsory education.	8.50%	10%	31.09%
2	Age 16-59, fair health but greater than sample’s average proportion of disability, employed in managerial, mid- and lower level occupations, mortgage/rental accommodation, single or divorced/dissolved/ widowed, below-median income, GCSE level or higher than compulsory education.	6.90%	2.96%	19.14%
3	Age 30-59, good health, not disabled, self-employed or managerial role, mortgage holders, married/have partner, above compulsory education.	11.20%	0.23%	2.34%
4	Age 16-29, good health, not disabled, unemployed, inactive or student or semi-routine employment, job single, renting, basic/none or GSCE-level education.	8.90%	0.15%	3.36%

5	Age 60+, fair health, disability, married or had partner, homeowners.	23.70%	0.02%	0.34%
6	Age 16-59, good health, employed in intermediate to semi/routine jobs, renting accommodation, below-median income, single/married, basic/none or GCSE education.	17.60%	0%	0.16%
7	Age 16-59, good health, employed in higher and mid-level management jobs, mortgaged accommodation, above-median income, married, above compulsory education.	23.20%	0%	0.08%

4.5 Discussion

In this paper, we take a person-centred approach to investigating who's miserable now. We define misery using the four measures of SWB used by the ONS. We only consider someone to be miserable if they report low well-being on all four measures. In this way, we partly circumvent the debate about which of the four questions best reflects SWB and address concerns surrounding fuzzy preferences and simply mistaken subjective reports. According to this definition, 1.1% of the total sample are miserable. We examine who is among the worst of in society by using LCA to identify groups of people united by specific observable characteristics and highlighting those characteristics that differentiate groups more vulnerable to misery from those at lower than average risk of being miserable.

The LCA highlights two groups that are at higher-than-average risk of being miserable. By far the most vulnerable are those belonging to class 1. Of the miserable people included in our analysis, class 1 account for 77%. Members of this group tend to be aged 30+, economically inactive, face disability and health problems, live in rented accommodation, have compulsory or lower levels of education and tend not to be in a partnership. Those in class 2 are also vulnerable to misery, making up 19% of the miserable people in our sample. People in this class share some but not all of the characteristics which define class 1. Members of class 2 also report some health issues and have a higher-than-average risk of disability. They also tend not to be in a partnership. However, this group tends to be employed, is younger, more educated and is as likely to have a mortgage as to be renting.

Together the members of these two classes make up just over 15% of the sample but they account for 96% of the miserable. These people, therefore, largely represent the answer to the question of who is miserable now. Their shared characteristics are perhaps unsurprising given the existing SWB literature. Health and marital status are both long-established factors associated with SWB (Dush and Amato, 2005; Steptoe, Deaton, and Stone, 2015) and employment and job security have also been identified as key predictors (see e.g. Knabe et al., 2010; Dawson, Veliziotis, and Hopkins, 2017). Other work which has examined the relationship between homeownership and SWB and separately higher education and SWB has found mixed results (Oswald et al., 2003; Zumbro, 2014; Kristoffersen, 2018). The current work builds on existing studies by highlighting the substantive risk of misery facing those who concurrently lack a number of these different protective factors.

Much of the existing literature has examined the determinants of LS. An analysis of the most miserable 5% of the population on LS yields similar results, with classes 1 and 2 remaining the classes which are the most vulnerable to misery. Among those with lower, but non-negligible, chance of being miserable, however, differences do emerge. While the low SWB definition of misery ranks class 3 – well-educated, married, middle-aged self-employed people with a mortgage – over class 4 – young single people still in education, unemployed or in semi-routine employment who are renting – in terms of vulnerability to misery, for the low LS the order is reversed. The reasons for the priority ranking reversal across the two definitions are unclear. It might, for example, reflect the worries that self-employed people report facing over their job and financial security (Binder and Coad, 2013), or their greater susceptibility to negative emotions and stress (Patzelt and Shepherd, 2011;Blanchflower, 2004). Equally the difference may be explained by unemployment being negatively associated with evaluative but not experiential SWB (Knabe et al., 2010). It is important to emphasise, however, that the major difference in the response to our overarching question of who is miserable when we look across the two definitions of misery is one of scale rather than composition: Many more people are miserable when we define misery as low life satisfaction, compared to reporting low SWB across all four measures, but class 1 and 2 still account for the vast majority of the miserable in both cases.

This paper reflects an initial step on a journey towards a goal of improving the wellbeing of the worst off by providing insights into who is miserable. This approach is not without limitations. In terms of identifying who are the most miserable, we must rely on the APS survey questions on people's life circumstances and we must rely on those surveyed in the sample. The APS includes a broad range of questions but it does not cover all of the dimensions of wellbeing of potential interest, nor all of the determinants of SWB that have been identified in the literature. For example, the APS it is lacking in terms of indicators on people's evaluation of the meaning of their lives and how people spend their time, which existing work identifies as an important dimension and predictor of SWB respectively (Stone and Mackie, 2013; Laffan, 2018). Furthermore, those interested in SWB and misery must do more to get at populations who do not participate in population surveys, such as the homeless and those in institutions such as care homes and prisons, many of whom we might to be among the worst off in society. For example, homeless people, which, depending on the definition, constitute about 0.5% (320,000) of the UK population (Shelter, 2018) and we do not capture them in our analysis[9].

In terms of establishing the factors associated with who is the worst off, LCA helps us to identify groups of individuals at the highest risks of misery but like most data science tools it requires large volumes of complete observations. This means that once a person fails to answer one of the survey questions (e.g. housing tenure) their entire entry is dropped from the clustering analysis, which can be a problem for the cases where the non-response to certain questions is group-specific (Heffetz and Rabin, 2013). This can be particularly challenging if the non-response behaviour is correlated with the

[9]Interestingly, the limited evidence that exists suggests that the homeless people are not necessarily miserable on all four SWB dimensions: Biswas-Diener and Diener, 2006 finds that whilst the levels of negative affect are higher among the homeless people in both India and the US, only in the US their life satisfaction is below neutral.

variable of interest, i.e. if the miserable tend to avoid answering certain questions about themselves.

Finally, we cannot make causal claims based on our analysis. Like other correlational SWB research, the associations we present are vulnerable to reverse causality and omitted variable bias. As a result, insights from the current work do not suggest how to address people's misery but rather identify those group of people that policymakers should pay particular attention to. In particular, our results emphasise the importance of considering how and why individual factors may interplay to make people more or less vulnerable to misery. For example, the misery of those in poor health whilst in active workforce may be driven by the manner in which they combine causing daily concern about job security. In contrast, misery of those individuals whose poor health prevents them from participating in the workforce may be, in part, caused by the resulting loneliness they experience. Optimal policies to address misery should be informed by evidence on the way combinations of such factors influence people's SWB.

Importantly, even if the combination of characteristics that the analysis identifies as being predictive of misery do represent causal impacts on wellbeing, some characteristics more susceptible to policy intervention than others, for example, job security compared to marital and disability status. Others like age and sex are not at all. Several of the shared characteristics in both groups with a higher-than-average percentage of miserable – including a relatively high risk of being in poor health and having a disability – suggest that members of this group may be inelastic suppliers of wellbeing and the potential for policy intervention to improve their wellbeing may be limited.

Notwithstanding these limitations, the current work makes significant contributions to our understanding of who's miserable now. One of the most important yardsticks for judging a society is how well it treats its worst-off. By looking across the four ONS wellbeing questions, we classify just over 1% of APS sample as being in the most miserable group. By identifying which clusters of people are most vulnerable, we hope to have provided researchers and policymakers with insights which can assist them in more accurately identifying who to target when trying to improve the lives of the worst-off.

Chapter 5

Conclusion

To conclude, this thesis applied novel experimental and econometric techniques to investigate three questions that are difficult to address by conventional methods. In the first Chapter we used an incentivised experiment to investigate the role of self-selection into free testing for Covid-19 on prevalence estimates relied upon by governments, healthcare authorities and the public worldwide. Methodologically, such investigation required creating an incentive-compatible setting where (i) the non-monetary costs of testing, such as waiting time are non-negligible and (ii) there is a systematic difference in people's willingness to incur these costs that is correlated with the outcome variable. In case of Covid-19, it is intuitively plausible that people with symptoms (or beliefs of being at risk) are *both* more willing to incur the non-monetary cost to get tested *and* are more likely to test positive, compared to people who feel healthy. We designed an experiment which elicited incentive-compatible willingness to wait for a free test conditional on the person's current symptoms – and the hypothetical willingness to wait conditional on alternative symptoms and/or beliefs about likeliness of having Covid-19. We showed that self-selection bias leads to majorly inflated prevalence estimates, and varies by age group and other socio-demographic characteristics. We then validated the results using independently collected governmental data. This allowed us to demonstrate that the bias is time-varying, which renders any conclusions based on changes across time at best uninformative. Finally, we developed a polling-based alternative to random testing, which circumvents the bias altogether. Since this approach has currently been tested on European data; for a wider application, conversion rates from symptoms to positivity and their correlation with socio-demographic characteristics must be measured.

The subsequent chapter addresses the factors affecting prosocial behaviour. Other-regarding preferences and prosocial behaviour vary across people and countries. Systematic differences are often related to fundamental personal characteristics such as gender and cultural background, but for a given person, preferences are traditionally assumed to be stable. To what extent do important lifestyle changes affect preferences, such as prosociality? In this chapter we demonstrated that retirement increases the probability of volunteering and found evidence suggesting that enhanced prosocial preferences, rather than increased time availability, explain this empirical observation. Given the population ageing, our results are relevant and useful to inform the design of policies that seek to exploit the productive capacities of individuals who withdraw from the labour market and enter retirement. Considering differences in the generosity of the pension system between countries and across time, further research could provide valuable insights into the potential role of reciprocity in the retirees' prosocial behaviour.

The final Chapter used latent class analysis, a descriptive methodology common in social sciences other than economics, to identify eight latent classes of the population and detect the combinations of characteristics most associated with misery. Together with estimating the relative sizes of these subgroups in the population this, we believe, provided a direction for further investigation of causality behind low SWB in these groups – and the way to alleviate it. We stress the importance of the direction, as both ours and past research show evidence that different factors would be prioritised when looking to increase the average SWB in the population. Of course, the welfare gains may not be equally costly across the wellbeing distribution – which lends further motivation for better understanding the determinants of wellbeing of the worst off. Notably, by nature of a household survey design, some of the likely worst-off such as homeless people, are not represented in our data. One important direction or further research would be to seek out the factors affecting the SWB of such "off-grid" subgroups – as well as investigate whether such factors can be detected at earlier stages, such as in survey data.

Appendix A

Factors associated with testing bias

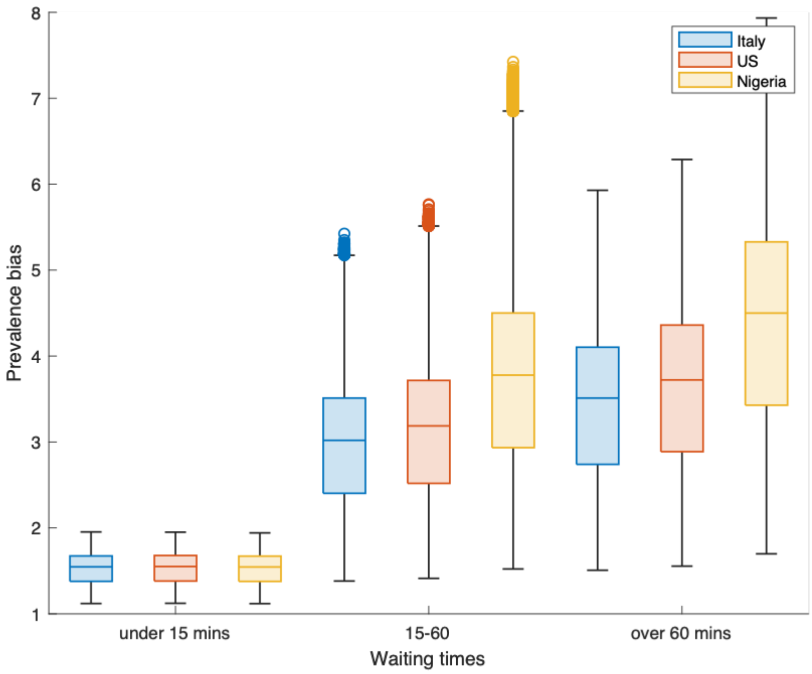


FIGURE A.1: Estimate of the prevalence bias in field testing. Test positivity divided by the best prevalence estimate using REACT and ONS data.

TABLE A.1: Odds of dropping out from (hypothetical) wait for a free Covid-19 test, by age group and symptoms.

	Dependent variable:
	Odds of not waiting for test (Ref: Age 30-50 | No symptoms)
Under 30 | No symptoms	−0.114
	(0.099)
Under 30 | Symptoms	-1.342^{***}
	(0.109)
30-50 | Symptoms	-1.279^{***}
	(0.105)
50+ | No symptoms	−0.107
	(0.105)
50+ | Symptoms	-1.403^{***}
	(0.119)
Observations	1,150
R^2	0.264
Max. Possible R^2	1.000
Log Likelihood	−6,177.707
Wald Test	351.360^{***} (df = 5)
LR Test	352.721^{***} (df = 5)
Score (Logrank) Test	388.121^{***} (df = 5)
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

TABLE A.2: Proportional hazard ratio for dropping out from (hypothetical) wait for a free Covid-19 test, by age group and symptoms.

	Dependent variable:
	Odds of not waiting for Covid-19 test (Ref: Age 30-50 ∣ No symptoms)
Under 30 ∣ No symptoms	−0.114
	(0.099)
Under 30 ∣ Symptoms	-1.342^{***}
	(0.109)
30-50 ∣ Symptoms	-1.279^{***}
	(0.105)
50+ ∣ No symptoms	−0.107
	(0.105)
50+ ∣ Symptoms	-1.403^{***}
	(0.119)
Observations	1,150
R^2	0.264
Max. Possible R^2	1.000
Log Likelihood	−6,177.707
Wald Test	351.360^{***} (df = 5)
LR Test	352.721^{***} (df = 5)
Score (Logrank) Test	388.121^{***} (df = 5)
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Appendix B

Descriptive statistics for the experiment

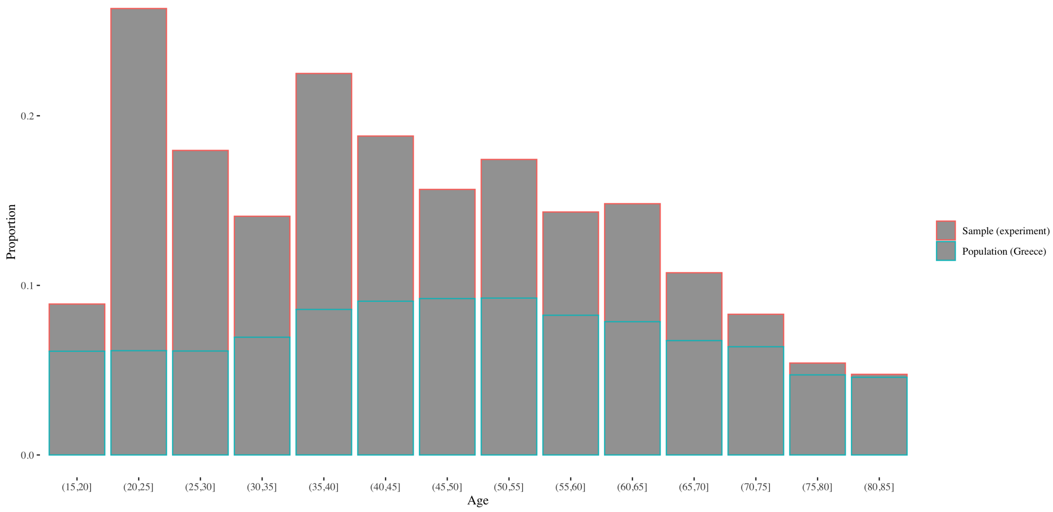


FIGURE B.1: Age distribution in the experiment (n=578) and in population of Greece (source: populationpyramid.net).

	Mean	Std dev	min	max
Covid symptoms (0=no; 1=yes)	1.38	0.12	0	1
Age	40.404	15.302	18	84

TABLE B.1: Summary statistics of sample demographics and symptoms for the experimental data.

Hypothetical willingness to test (N=575)		
	By symptoms	**By waiting time**
	No Symptoms	
Mean (SD)	2.96 (1.48)	2.39 (2.04)
Median (Min, Max)	3.00 (1.00, 5.00)	2.00 (0, 8.00)
	Flu Symptoms	
Mean (SD)	2.00 (1.20)	3.81 (2.26)
Median (Min, Max)	2.00 (1.00, 5,00)	4.00 (0, 8.00)
	Covid Symptoms	
Mean (SD)	1.46 (0.951)	5.19 (2.35)
Median (Min, Max)	1.00 (1.00, 5,00)	5.00 (0, 8.00)

By-symptoms key: 1: certainly yes; 2: probably yes; 3: maybe; 4: probably no; 5: certainly no
By-wait-time key: : 0: would not wait at all; 1: would only take it if available immediately; 3: 5 - 15 minutes; 4: 15 - 30 minutes; 5: 30 - 45 minutes; 6: up to an hour; 7: 1 - 2 hours; 8 over 2 hours

TABLE B.2: Summary statistics for hypothetical willingness to wait to take the test, by symptoms and waiting time.

Prize	Not entered	Dropped upon learning waiting time	Dropped after some wait	Swapped prize for cash	Kept prize	N
Book voucher	103	31	38	46	48	266
Test voucher	138	25	31	62	12	268
Bias	1,263	1,267	1.28	4.03	0.25	534

TABLE B.3: Willingness to wait for a 1/30 chance of winning a prize. Number of subjects by level of task completion and incentive (rows 1-2), bias by incentive (row 3).

Appendix C

Survey data

TABLE C.1: Early retirement ages and sample sizes for countries in the EU-SILC and SHARE data.

Country	EU-SILC sample	SHARE sample	2011	2013	2015	2017
Austria	3,236	8,740	M:62; F:58	M:64; F:59	M:64; F:59	M:65; F:60
Belgium	2,936	11,304	60	60	60	62
Switzerland	3,700	7,080	M:63; F:62	M:63; F:62	M:63; F:62	M:63; F:62
Cyprus	2,323	576	63	63	63	63
Czech Republic	5,122	11,984	60	60	60	60
Germany	7,483	9,151	63	63	63	65
Denmark	2,259	8,045	65	65	65	65
Estonia	3,408	11,484	60	60	60	60
Greece	6,781	3,380	M:60; F:55	M:60; F:57	62	62
Spain	5,713	7,689	61	61	61	61
Finland	3,464	1,301	63	63	63	63
France	6,523	9,543	60	60	61	61
Ireland	1,498	-	66	66	66	66
Italy	5,452	6,290	M:62; F:61	M:63; F:62	M:63; F:62	M:63; F:62
Luxemburg	1,834	2,093	57	57	57	60
Netherlands	3,114	-	65	65	65	65
Norway	1,752	-	67	67	67	67
Poland	5,813	4,579	M:65; F:60	M:65; F:60	M:65; F:60	M:66; F:61
Portugal	4,686	1,933	55	55	65	65
Sweden	1,749	7,616	61	61	61	61
Slovenia	2,821	6,846	M:58; F:57	M:58; F:57	M:59; F:58	M:60; F:59
Slovakia	4,028	1,548	60	60	60	60
Total	85,695	121,182	-	-	-	-

Source: Social Security Programs Throughout The World; OECD Pensions At A Glance.

Notes: M is for Males and F for Females. Ireland, Norway and the Netherlands are not included in the SHARE estimation samples. Sample sizes correspond to the baseline estimation samples. For the EU-SILC sample, figures refer only to 2015.

Appendix D

Outcome variables and descriptive statistics

D.1 Volunteering indicators

Within the official International Labour Organisation (ILO) definition of volunteering, EU-SILC data measures two sub-types of volunteer work: informal and formal. The first prosocial behaviour indicator is about participation in formal voluntary work. Respondents aged 16 years old and over reported if, during the last 12 months, they carried out any unpaid non-compulsory work for or through an organisation, a formal group, a club as well as for a charitable or religious organisation. Activities related to people, the environment, animals and the wider community, and attending meetings related to those activities were considered. Unpaid internships in profit-making companies were not considered. Respondents justified their non-participation in formal volunteering due to lack of interest, lack of time or other reasons. Their answers were grouped to construct a binary outcome indicating participation (or not) in formal volunteering.

The second prosocial indicator records participation in informal voluntary activities. Respondents aged 16 years old and over in the 2015 EU-SILC wave were asked whether, during the last 12 months, they undertook any informal unpaid activities that were not arranged, organised or motivated by any organisation. These activities include helping other people including family members living outside their household (e.g. cooking for others, taking care of people in hospitals or at home, taking people for a walk, shopping etc.), taking care of homeless or wild animals, and participating in other informal voluntary activities (cleaning a beach, a forest etc.). Informal volunteering excludes any activity related to own household, work or undertaken within charitable organisations. Respondents also reported the reasons for not being engaged in informal volunteering, i.e. due to lack of interest, time or other reasons.

D.2 Other prosocial behaviour indicators

The 2015 EU-SILC wave also reports if individuals participated in political or local interest group activities, public consultation, peaceful protest, petition signing, participation in demonstration, writing letters to politicians or the media. Active participation using the internet and attending meetings related to these activities were also considered. Voting and participation in elections were not considered. Not participating in such activities was justified due to lack of time, interest or other reasons. An active citizenship indicator was constructed using those responses. Similarly, the 2016 EU-SILC wave collected information on providing care or assistance to others (excluding childcare). Three outcome variables were constructed using those responses: (a) whether the respondent provided care or assistance to people from inside or outside their household relative to those who not engaging in such activities; (b) whether someone provided care or assistance only to household members relative to those who do not provide any care or assistance; and (c) whether someone provided care or assistance only to people from outside the household relative to those who do not engage in any caring activity.

D.3 Descriptive statistics

Table D.1 presents descriptive statistics, by retirement status, on the outcome variables available in the surveys and experiment data. Statistics are weighted by the respective survey weights where relevant. As throughout our analysis, samples have been restricted to include individuals within a 10-year window around their country, survey year gender-specific ERAs.

TABLE D.1: Descriptive statistics on outcome variables.

	Retirees	Non-retirees	Difference	Observations
	[1]	[2]	[3]	[4]
Voluntary work (SHARE)	0.207	0.184	.023***	-
Observations	70,557 (58.2%)	50,625 (41.8%)	-	121,182
Informal voluntary work (EU-SILC)	0.294	0.256	.038***	-
Observations	41,291 (47.1%)	46,477 (52.9%)	-	87,768
Formal voluntary work (EU-SILC)	0.242	0.234	.008*	-
Observations	41,293 (47.1%)	46,475 (52.9%)	-	87,768
Provide care inside household (EU-SILC)	0.068	0.066	0.001	-
Observations	42,207 (49.3%)	43,383 (50.7%)	-	85,590
Provide care outside household (EU-SILC)	0.091	0.1	-.010***	-
Observations	43,454 (48.4%)	46,264 (51.6%)	-	89,718
Active citizenship (EU-SILC)	0.147	0.167	-.020***	-
Observations	41,290 (47.1%)	46,462 (52.9%)	-	87,752
Formal voluntary work (experiment)	0.315	0.211	-.104*	-
Observations	143 (56.7%)	109 (43.2%)	-	252

Source: EU-SILC; SHARE; experiment. Notes: EU-SILC waves: 2015, 2016; SHARE waves: 4, 5, 6, 7; Experiment waves: all. Means are weighted by the respective survey weights (EU-SILC and SHARE). Survey samples cover individuals 10 years around their country, survey year and gender-specific ERAs.

Appendix E

What drives volunteering after retirement? Survey data

E.1 Observable characteristics

We also examine, by splitting the survey samples accordingly, whether the impact of retirement varies with observable characteristics (Table E.1). The baseline effect on volunteering is driven by females, especially in SHARE, and this is confirmed in the EU-SILC data. However, the EU-SILC data suggest that retirement affects informal volunteering for females and formal volunteering for males. Volunteering after retirement is more likely for those who have completed tertiary education. Regarding the retirement impact on informal volunteering, the parameter estimate is higher, although significant at the 10%, for those having completed only secondary education.

Results by subjective health status indicate the existence of a health gradient behind the baseline effects. Individuals were grouped based on having reported bad or very bad health status, or fair, good or excellent health. The retirement effect is strong and positive only for those with good self-reported health status, and it is not significant for poor health individuals. This is confirmed when the samples are split using a variable indicating whether respondents are limited in their activities due to health issues. In both samples, the baseline effect is driven by those reporting that their activities are not limited due to health-related issues. In the case of formal volunteering the only significant estimate comes from those reporting not being severely limited in their activities due to health problems. For those severely limited due to health issues, the estimated parameters are either very low or negative, and always not significant.

TABLE E.1: Retirement and prosocial behaviour: Analysis by group.

	Voluntary work (SHARE sample)	Voluntary work (EU-SILC sample)	Informal voluntary work (EU-SILC sample)	Formal voluntary work (EU-SILC sample)
	FE-2SLS	2SLS	2SLS	2SLS
Sub-group:	[1]	[2]	[3]	[4]
Males	.055 (.053)	.071 (.049)	.038 (.040)	.081* (.044)
Females	.150*** (.047)	.087* (.051)	.081* (.046)	.062 (.044)
Primary or less education	.168 (.110)	-.104 (.092)	-.063 (.082)	-.007 (.069)
Secondary education	.026 (.042)	.079* (.042)	.065* (.036)	.043 (.036)

Tertiary education	.259*** (.076)	.191*** (.092)	.112 (.078)	.235*** (.088)
Fair/Good/Very Good health status	.093** (.037)	.091** (.037)	.063** (.032)	.087*** (.033)
Bad/Very Bad health status	.052 (.320)	-.038 (.131)	.063 (.103)	-.109 (.112)
Not limited in activities due to health	.099** (.048)	.079* (.042)	.060 (.036)	.058 (.038)
Not severely limited in activities due to health	.072 (.084)	.099 (.075)	.078 (.067)	.156** (.065)
Severely limited in activities due to health	.119 (.179)	-.039 (.125)	.006 (.107)	-.080 (.108)

Source: EU-SILC; SHARE. Notes: Results are weighted using survey weights. Models use a local linear age function, and control for the usual set of individual characteristics and fixed effects. Robust standard errors in parentheses. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

E.2 Social network and spillover effects

Individual social networks cannot be identified in the data. Nevertheless, using the SHARE sample we examine whether a person's volunteering activity is affected by their partner's activity. Table E.2 (Panel A) displays the FE-2SLS results using a local linear age function and a 10-year window around ERA. Column 1 confirms the positive retirement impact and demonstrates a sizeable positive effect for individuals whose partner is volunteering. Then, the sample is split based on the partner's volunteering activity during the same year (columns 2-3). Retirement has a positive effect regardless the partner's prosocial behaviour, however, its impact is considerably higher when partners have also volunteered during the same year. Furthermore, the longitudinal design of the SHARE data allows to calculate how intense is the partner's volunteering activity across the 4 waves of the survey. Column 5 suggests that retirement increases volunteering when partners of retirees tend to volunteer more often. The results hold when we use samples of individuals whose partners' retired more than 2 and more than 3 times in the period.

Next, we empirically test the hypothesis that transitions to retirement could cause spillover effects on prosocial behaviour within the household. We construct a binary indicator on whether an individual's partner is retired and instrument it the usual way, i.e. a dummy on having crossed the respective ERA. Panel B in Table E.2 displays the results. There is a strong first-stage evidence for partners as well, however, own volunteering is not affected by partner's retirement (column 1). Controlling for both own and partner's retirement (column 2) confirms this result; the probability of volunteering is only affected by own retirement. Column 3 provides further evidence showing that own retirement does not have an impact of partner's volunteering. In columns 4-5 the sample is split based on partner's retirement status. Own retirement does not affect the volunteering probability for individuals whose partners are retirees (column 4). On the contrary, own retirement has a strong positive relationship with volunteering for those whose partners are still in the labour market. This could be an indication of a substitution effect for couples of retirees towards more home-oriented activities, in line with the evidence presented by Moreau and Stancanelli, 2015.

TABLE E.2: Retirement and prosocial behaviour: Couple complementarities.

	[1]	[2]	[3]	[4]	[5]
	Panel A: The role of partner's volunteering activity				
	Total sample	Partner did not volunteer in same year	Partner volunteered in same year	Partner did not volunteer in period	Partner volunteered ≥1 time in period
Retired	.113** (.048)	.096* (.050)	.302* (.159)	.083** (.037)	.191** (.089)
Partner volunteering	.123*** (.013)	-	-	-	-
Observations	51,249	39,130	7,329	79,676	19,164
	Panel B: The role of partner's retirement				
	Own volunteering (total sample)	Own volunteering (total sample)	Partner's volunteering (total sample)	Own volunteering (partner retired)	Own volunteering (partner not retired)
Retired	-	.115** (.057)	.048 (.049)	.095 (.111)	.224*** (.080)
Partner retired	.035 (.052)	-.012 (.061)	-	-	-
First stage: Own age>ERA	-	.231*** (.013)	.241*** (.013)	.184*** (.020)	.193*** (.019)
First stage: F-statistic	-	97.88	183.03	82.54	96.5
First stage: Partner's age>ERA	.239*** (.013)	.232*** (.013)	-	-	-
First stage: F-statistic	171.99	93.97	-	-	-
Observations	52,212	52,212	52,093	19,271	29,570

Source: SHARE. Notes: FE-2SLS estimates. Results are weighted using survey weights. Models use a local linear age function and control for the usual set of individual characteristics and fixed effects. Robust standard errors in parentheses. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Another hypothesis could be that prosocial behaviour post-retirement is affected by own previous experience and activity. This point has been raised by Erlinghagen, 2010 who argued that the effect of retirement on volunteering is rather exaggerated and it is own previous experience that determines prosocial behaviour after leaving the labour market. The longitudinal design of the SHARE data allows to test this argument. Therefore, a dynamic version of Equation 3.1 is estimated by including a one-year lagged dependent variable alongside the rest regressors and fixed effects. Due to the fact that the sample consists of thousands of individuals followed for a relatively short period of time, i.e. 4 waves, the 2SLS and the 2SLS-FE estimators will be upwards and downwards biased, respectively (Nickell, 1981[1]). In cases with

[1]In the 2SLS case, the lagged dependent variable would be correlated with the individual fixed effect in the error term. Demeaning the data would eliminate the time invariant effect, however, the lagged dependent variable will remain correlated with the disturbance term. Even if the number of individuals is large, this sort of correlation induces a bias of order 1/T, which is quite sizeable in small panels as in here where T=4 ((Nickell, 1981).

"small T, large N" effects, serial correlation within individuals, endogenous regressors and, possibly, predetermined lagged explanatory variables, the system Generalised Method of Moments (GMM) estimator has been shown to be quite consistent; especially when $T \geq 3$ (Arellano and Bond, 1991, Blundell and Bond, 1998, Bond, 2002, Roodman, 2009).

Table E.3 (Panel A) displays the results. For reference, results for a static specification are also provided. SHARE is not a balanced panel of individuals, hence results are reported using both the original panel, a more balanced version of it where individuals are observed at least in 3 waves, and a fully balanced panel of individuals observed in all waves, i.e.$T = 4$[2]. Regarding the static specifications, the results confirm a positive effect of retirement on volunteering activity. Although the GMM evidence suggests that the relationship is weaker relative to the 2SLS-FE estimates, the retirement status coefficient estimates are statistically significant at the 1%. Controlling for dynamics, columns 3-4, confirms that previous volunteering experience is a very strong predictor of current activity. However, including a lagged dependent variable leaves the retirement status coefficient unaffected, therefore indicating that there is an autonomous impact of retirement on the probability of offering volunteering work, regardless of past activity[3].

TABLE E.3: Retirement and prosocial behaviour: Past activity and volunteering frequency.

	Panel A: The role of past volunteering activity			
	[1]	[2]	[3]	[4]
Retired	.024*** (.007)	.040*** (.012)	.028*** (.008)	.039*** (.012)
Volunteered last year	-	-	.212*** (.022)	.227*** (.026)
Panel time dimension	T ≥ 1	T = 4	T ≥ 1	T = 4
First stage: F-statistic	83.38	28.67	4.49	38.12
Instrument count	28	21	29	25
Hansen test	10.2	16.67	15.42	27.64
Observations	121,182	44,244	61,311	31,407
	Panel B: The role of volunteering frequency			
	Volunteer less than every month	Volunteer almost every month	Volunteer almost every week	Volunteer almost every day
% retired among those who:	48.2	55.12	67.24	70.51
% volunteer among retired:	4.07	5.66	8.29	3.39
Retired	-.005 (.007)	.001 (.008)	.047*** (.010)	.011* (.006)

[2]Results are robust when using samples with $T \geq 3$. Also, we obtained FE-2SLS estimates using subsamples of SHARE individuals observed in at least 3 and all 4 waves. All first-stage relationships are strong and the impact of retirement is higher as compared to the baseline results using the total -unbalanced- panel. In the case where $T = 4$ the impact of retirement is statistically significant even when higher order local functions of age are used at both sides of the cutoff. All tests are available upon request.

[3]Moreover, past volunteering activity is a stronger predictor of today?s behaviour in the case of non-retirees. After splitting the sample by retirement status (and using those within 10 years before or after their ERA), the coefficient of the lagged dependent variable is .224 (standard error = .039) for non-retirees and .198 (standard error = .026) for those retired.

Volunteered last year	.070*** (.020)	.074*** (.020)	.126*** (.024)	.037** (.016)
Panel time dimension	T = 4	T = 4	T = 4	T = 4
First stage: F-statistic	5.96	9.65	26.2	6.71
Instrument count	25	25	25	25
Hansen test	15.53	18.61	13.01	10.37
Observations	25,823	26,350	27,080	25,116

Source: SHARE. Notes: System Generalised Method of Moment (GMM) estimates. Results are weighted using survey weights. Models use a local linear age function and control for the usual set of individual characteristics and fixed effects. In dynamic specifications, lagged variables are instrumented using instruments dated $t-2$ and earlier. Samples include individuals within 10 years at both sides of the ERA cutoff. Windmeijer-corrected cluster-robust standard errors in parentheses. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

Another concern could be around the impact of retirement on the intensity -or frequency- of offering unpaid work. Although the actual number of volunteer hours is not available, SHARE respondents reported how often they provided voluntary or charity works within the last 12 months. Those who volunteered, were given the following options: (a) almost every day; (b) almost every week; (c) almost every month; and (d) less often. Based on these responses, four binary indicators are constructed. ?hose not volunteered were the reference group in each case. Then, the dynamic version of Equation 3.1 was estimated, using those four indicators as outcomes and controlling for past volunteering behaviour (regardless of its frequency and treating it as predetermined). Panel B of Table E.3 displays the system GMM estimates. The fraction of retirees among volunteers increases with the frequency of volunteering activity. For example, 48% of the sample of those who volunteer less than once per month (column 1) are retirees. However, the fraction of retirees in the sample of those who report volunteering activity almost every day rises to 70%. This could be partially attributed to increased time availability post-retirement, although a distinction in the SHARE sample among volunteering types, i.e. formal, informal, household oriented or not, would be quite useful to look deeper in their activity patterns. The prevalence of volunteers in the sample of retirees follows a hump-shaped pattern as volunteering frequency increases. More specifically, 4% of retirees volunteer less often than every month, 8.3% of them volunteer every week, and 3.4% volunteer almost every day.

Using the sample of individuals observed in all 4 SHARE waves, reveals that the impact of retirement also follows a hump-shape profile as the frequency of volunteering activity increases. The effect is zero in columns 1-2 where only those volunteering every month or less often are used. However, retirement has a positive and significant impact on volunteering almost every week, relative to non-volunteers (column 3). There is also a lower, and less precisely estimated, positive impact of retirement on the probability of volunteering almost every day. Moreover, past volunteering activity is always a strong predictor of current volunteering frequency ? especially for those volunteering almost every week.

E.3 Robustness checks

The results so far have been estimated using a 10-year bandwidth around the ERA cutoff. To check their sensitivity to the bandwidth choice, baseline models using a local-linear age function are re-estimated using a range of alternative bandwidths. Figure E.1 displays the results. 2SLS coefficients are plotted with their 95% confidence intervals. Horizontal dashed lines represent the baseline effects. For all volunteering indicators, the results are robust alternative bandwidths although point estimates become noisier as time windows narrow.

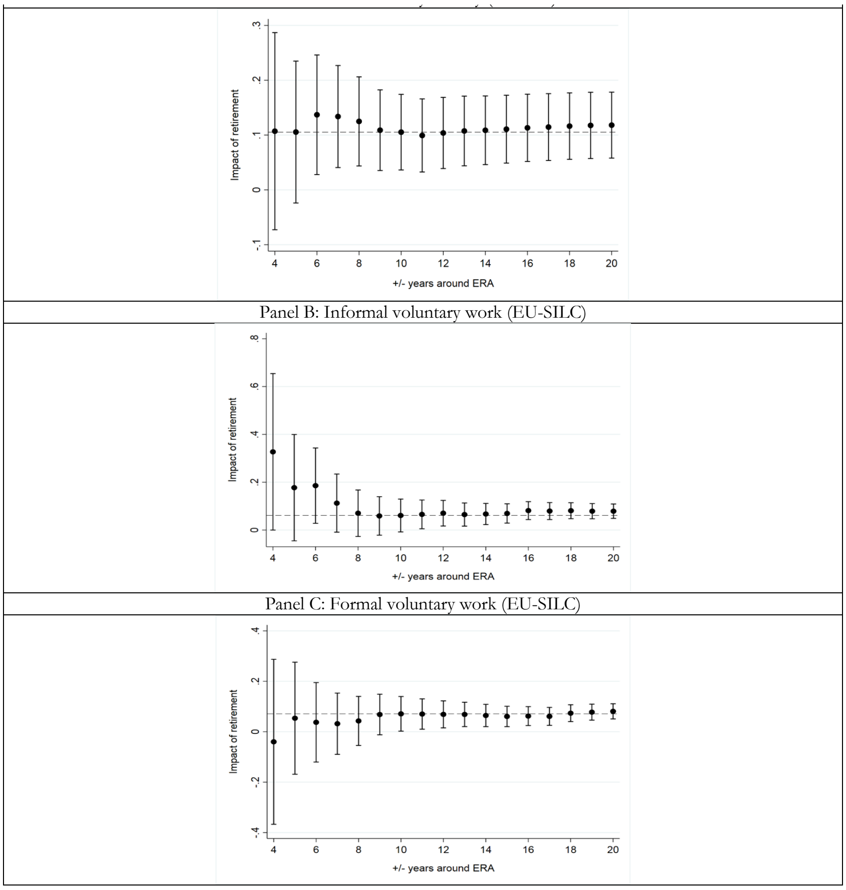


FIGURE E.1: Retirement and volunteering using alternative bandwidths around ERA. *Source*: SHARE; EU-SILC. *Notes*: 2SLS (EU-SILC) and FE-2SLS (SHARE) estimates using a local linear age function. Horizontal dashed lines correspond to the baseline effects (Table 2). Vertical lines represent the 95% confidence intervals based on robust standard errors.

Another robustness test is to replace the actual ERA for each individual with fake ones ranging a few years back. This will indicate whether prosocial behaviour is affected before crossing the official ERA, because individuals might opt to retire earlier, or they start adjusting their behaviour as they prepare to exit from the labour market. Table E.4 displays the results. Regarding informal volunteering, there are some statistically significant estimates up to four years before the actual ERA that disappear for earlier years. The results are in accordance with Mutchler, Burr, and Caro, 2003 who reported that informal volunteering is not affected by working status, mainly due to its obligatory nature. Moreover, they argued that as people grow older and retire, they should be receiving less requests for informal help because their social networks shrink post-retirement. This could also be implied by the diminishing parameter estimates in column 1; recall that the actual baseline effect is .061, i.e. lower than the placebo one estimated for $t-1$.

Contrary to the above effects, there are no significant estimates when placebo ERAs are used when formal volunteering is the outcome (column 2). This indicates that people tend to change only their informal volunteering behaviour as they approach their ERA. This is not the case for formal volunteering as the latter is more likely to be more structured and scheduled, and hence less compatible with working and commuting patterns. 2SLS and 2SLS-FE estimates for volunteering using the SHARE sample are also positive and significant up to four years before the actual ERA but the effect disappears after that. This could be conflating retirement implications on informal volunteering, however, no further disaggregation into volunteering types is possible as in the EU- SILC data. Therefore, these results can provide some support to the claim that people tend to change their prosocial behaviour as they as they approach their ERA, at least regarding the incidence of their volunteering and charity activity.

TABLE E.4: Retirement and prosocial behaviour: Falsification tests.

	Informal volunteering (EU-SILC sample)		Formal volunteering (EU-SILC sample)		Volunteering (SHARE sample)			
	2SLS	Obs.	2SLS	Obs.	2SLS	Obs.	FE-2SLS	Obs.
Fake ERA set at:	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
t-1	.068** (.033)	84,436	.065* (.033)	86,435	.084*** (.024)	116,691	.094*** (.036)	94,644
t-2	.088** (.037)	86,821	.058 (.037)	86,820	.082*** (.026)	111,370	.100*** (.038)	89,707
t-3	.118** (.047)	86,845	.063 (.047)	86,848	.088*** (.030)	105,541	.092** (.043)	84,335
t-4	.154** (.067)	86,831	.045 (.068)	86,829	.110*** (.041)	99,051	.106** (.052)	78,355
t-5	.202 (.123)	86,642	.094 (.125)	86,648	.117 (.075)	92,299	.124 (.081)	72,300
t-6	.415 (.440)	86,229	.111 (.447)	86,241	.356 (.302)	85,052	.330 (.200)	65,866

Source: EU-SILC; SHARE. Notes: Models are weighted using survey weights. Models use a local linear age function and control for the usual set of characteristics and fixed effects. Robust standard errors in parentheses. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

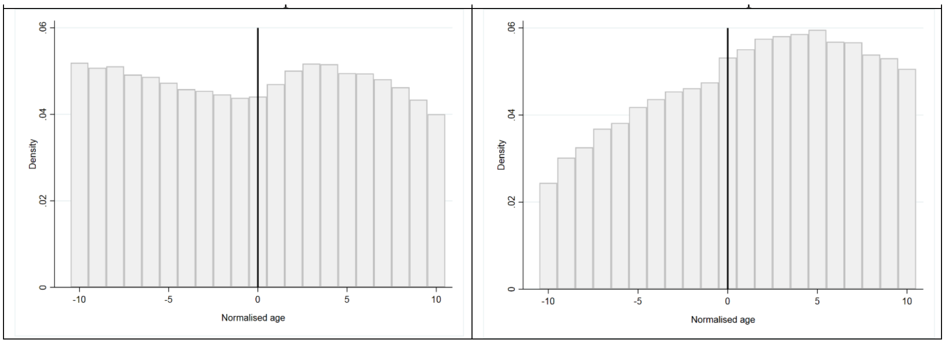


FIGURE E.2: Density of the forcing variable around the cut-off. *Source*: EU-SILC; SHARE.

Appendix F

Experiment data

TABLE F.1: Real effort and recipients.

	In-kind contribution (€0-5)[a]	Completed study over email (vs phone)
	[1]	[2]
Retired	-2.822** (1.401)	.051 (.085)
Recipient: charity (Ref category: Self)	6.831*** (1.034)	
Recipient: Relative/Friend	10.127*** (2.386)	
Completed study over email (vs phone)	13.147***(1.648)	
Age	.068 (.095)	-.015*** (.005)
R-squared	0.531	0.378
Observations	253	253
Local age function	Linear	Linear
Individual characteristics	Yes	Yes
Interview wave fixed effects	Yes	Yes

Source: experiment. Notes: Robust standard errors in parentheses. Asterisks ***, ** and * denote statistical significance at the 1%, 5% and 10% level, respectively.

a Calculated based on 28 units of real-effort task with maximum earnings of €5.

Appendix G

Reasons for not volunteering

The EU-SILC asked respondents why they were not engaged in any kind of volunteer work, i.e. due to lack of interest, due to lack of time, or due to any other reason. Figure G.1 graphs those trends by age. Among all non-volunteering individuals, the fraction of those being time- constrained decreases with age. Regarding non-volunteering retirees, the share of those being constrained by time, is small but relatively stable around ERA and starts decreasing quite late. Similar patters hold for those not engaged in volunteering due to other reasons. To rule out any health-related reasons, shares were also calculated using only those not being limited in their activities by some health condition but the trends are identical.

However, after crossing the ERA, there is a considerable increase in the fraction of those who do not volunteer, either formally or informally, due to lack of interest. The fraction of retirees not engaged in informal volunteering due to lack of interest decreases as the approach their ERA but to a much lesser extent after crossing it. Moreover, the fraction of retirees not offering formal voluntary work remains stable after crossing the ERA.

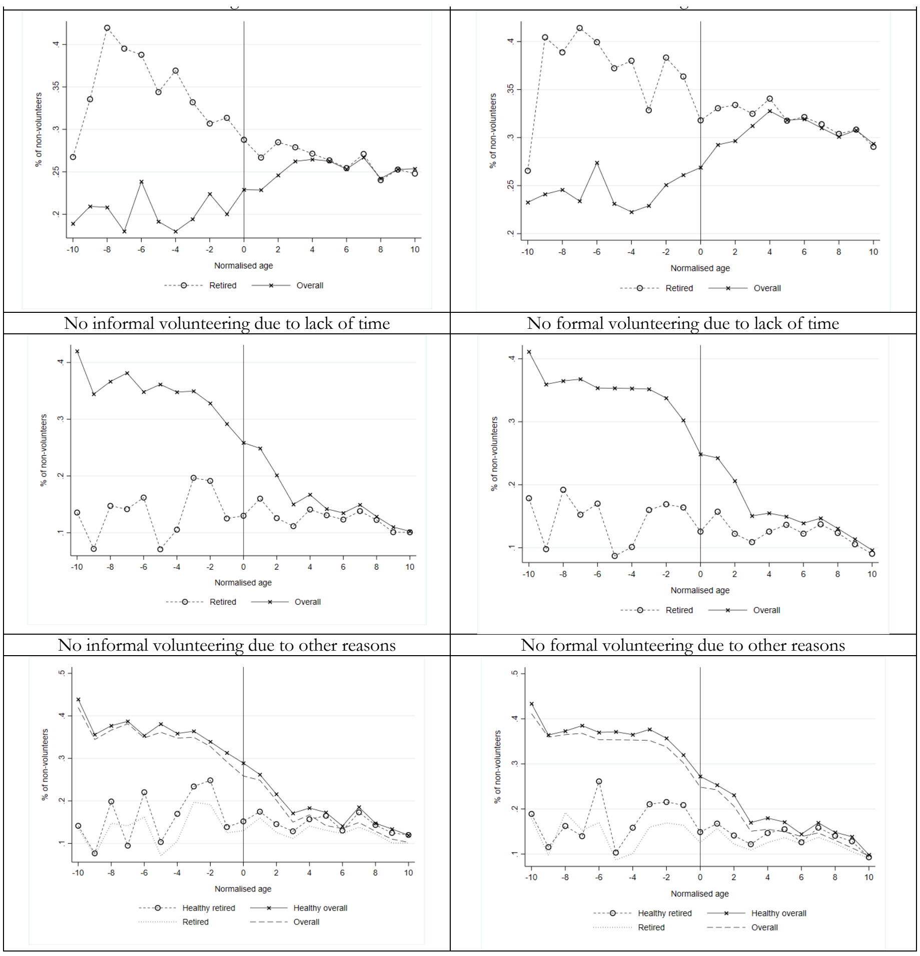


FIGURE G.1: Reasons for not volunteering. *Source*: EU-SILC
Notes: Means by normalised age are weighted using the survey weights. Overall refers to both retired and non-retired non-volunteers. Healthy refers to non-volunteers who report not being limited in their activities by any health-related conditions.

Appendix H

LCA compared to other classification approaches

Latent Class Analysis is a type of a Finite Mixture Model. Main differences between the LCA and other types of cluster analysis are: (i) model-based (rather than distance-based) grouping of data, (ii) probabilistic (rather than deterministic) assignment of class/group membership. Model-based grouping is well-suited for categorical variables, since there are no 'distances' between nominal categories, in contrast to continuous variables. Even though in some earlier literature SWB scales were assumed to be continuous, the general consensus now is that we should treat them as categories. Naturally, this approach is more dependent on the initial selection of the cut-off points for the categories (Hagenaars and McCutcheon, 2002).

Another advantage of a latent class model is that it is a probabilistic model for clustering. Probabilistic assignment allows for 'fuzzy sets' where we can measure to what extent we are sure that the individual belongs to a particular group. One may then assign the case to the latent class with the highest a posteriori probability (modal assignment), or leave classification 'fuzzy' -i.e, view the case as belonging probabilistically to each latent class to the degree indicated. Because the latent class model is probabilistic, it gives additional alternatives for assessing model fit via likelihood statistics, and better captures/retains uncertainty in the classification (Linzer, Lewis, et al., 2011).

Appendix I

Descriptive statistics for the APS data

TABLE I.1: Descriptive statistics for the 2014-2016 Annual Population Survey dataset, with full breakdown of the variables as per the ONS coding

	Overall (n=274732)
Sex	
Male	121425 (44.2%)
Female	153307 (55.8%)
Age	
16-29	31160 (11.3%)
30-39	41009 (14.9%)
40-49	47730 (17.4%)
50-59	51399 (18.7%)
60-69	53680 (19.5%)
70-99	49754 (18.1%)
Health	
Good/VeryGood	195953 (71.3%)
Fair	50502 (18.4%)
Bad/VeryBad	20290 (7.4%)
Missing	7987 (2.9%)
Disability Status	
Not Disabled	187530 (68.3%)
Disabled	67428 (24.5%)
Missing	19774 (7.2%)
Marital Status	
Single	69921 (25.5%)
Married/CPartner	140758 (51.2%)
Married/Partner(separated)	9071 (3.3%)
Divorced/Dissolved	31213 (11.4%)
Widowed	23729 (8.6%)
Missing	40 (0.0%)
Education	
Basic/None	22871 (8.3%)
Alevel	50692 (18.5%)
Degree/Professional	87459 (31.8%)
GCSE	46437 (16.9%)
Other qualification	19814 (7.2%)
Missing	47459 (17.3%)

Economic Activity	
Employee	127160 (46.3%)
ILO Unemployed	8145 (3.0%)
Inactive	20628 (7.5%)
Inactive(LT sick/disab)	13925 (5.1%)
Retired	75957 (27.6%)
Self-employed	23251 (8.5%)
Student	4386 (1.6%)
Missing	1280 (0.5%)
Socio-Economic Status	
Managerial/Professional(H/L)	67099 (24.4%)
Intermediate/Lower Superv	43474 (15.8%)
Semi/Routine	48184 (17.5%)
Small Employer/Own Acc	19962 (7.3%)
FT Student	8042 (2.9%)
Never worl/LT Unempl	7908 (2.9%)
NotClassif	61639 (22.4%)
Missing	18424 (6.7%)
House Ownership	
Rent	84344 (30.7%)
Mortgage	85432 (31.1%)
Owned	100999 (36.8%)
Missing	3957 (1.4%)
White British	
1	239405 (87.1%)
0	35064 (12.8%)
Missing	263 (0.1%)
Religious	
No	84721 (30.8%)
Yes	189662 (69.0%)
Missing	349 (0.1%)
Nonheterosexual	
0	4652 (1.7%)
1	242569 (88.3%)
Missing	27511 (10.0%)
Satisfied	
0-4	13628 (5.0%)
05-Aug	178489 (65.0%)
09-Oct	82615 (30.1%)
Worthwhile	
0-4	10511 (3.8%)
05-Aug	165845 (60.4%)
09-Oct	98376 (35.8%)
Happy	
0-4	24504 (8.9%)
05-Aug	153487 (55.9%)
09-Oct	96741 (35.2%)
Anxious (Reversed)	
0-4	55034 (20.0%)
05-Aug	106134 (38.6%)
09-Oct	113564 (41.3%)
Miserable	

0	271656 (98.9%)
1	3076 (1.1%)
Country	
England	208058 (75.7%)
Scotland	36295 (13.2%)
Wales	30379 (11.1%)

TABLE I.2: Descriptive statistics for the Miserable subsample of the 2014-2016 Annual Population Survey dataset.

	Overall (n=3076)
Sex	
Male	1360 (44.2%)
Female	1716 (55.8%)
Age	
16-29	177 (5.8%)
30-39	362 (11.8%)
40-49	706 (23.0%)
50-59	954 (31.0%)
60-69	513 (16.7%)
70-99	364 (11.8%)
Health level	
Good/VeryGood	502 (16.3%)
Fair	692 (22.5%)
Bad/VeryBad	1803 (58.6%)
Missing	79 (2.6%)
Disability Status	
Not Equality Act Disabled	593 (19.3%)
Equality Act Disabled	2296 (74.6%)
Missing	187 (6.1%)
Marital Status	
Single	1066 (34.7%)
Married/CPartner	723 (23.5%)
Married/Partner(separated)	220 (7.2%)
Divorced/Dissolved	727 (23.6%)
Widowed	340 (11.1%)
Education	
Basic/None	699 (22.7%)
Alevel	473 (15.4%)
Degree/Professional	569 (18.5%)
GCSE	598 (19.4%)
Other qualification	358 (11.6%)
Missing	379 (12.3%)
Economic Activity	
Employee	575 (18.7%)
ILO Unemployed	208 (6.8%)
Inactive	271 (8.8%)
Inactive(LT sick/disab)	1364 (44.3%)
Retired	497 (16.2%)
Self-employed	120 (3.9%)
Student	10 (0.3%)
Missing	31 (1.0%)

Socio-Economic Status	
Managerial/Professional(H/L)	306 (9.9%)
Intermediate/Lower Superv	333 (10.8%)
Semi/Routine	608 (19.8%)
Small Employer/Own Acc	173 (5.6%)
FT Student	27 (0.9%)
Never worl/LT Unempl	260 (8.5%)
NotClassif	1246 (40.5%)
Missing	123 (4.0%)
House Ownership	
Rent	1863 (60.6%)
Mortgage	507 (16.5%)
Owned	659 (21.4%)
Missing	47 (1.5%)
White British	
1	2798 (91.0%)
0	275 (8.9%)
Missing	3 (0.1%)
Religious	
0	1112 (36.2%)
1	1958 (63.7%)
Missing	6 (0.2%)
Non Heterosexual	
0	86 (2.8%)
1	2719 (88.4%)
Missing	271 (8.8%)
Country	
England	2289 (74.4%)
Scotland	407 (13.2%)
Wales	380 (12.4%)

Appendix J

Latent Class Analysis Model – Specification and Intuition

FigureJ.1 details the components of the LCA model in our specification. The 'dependent' variables that are being partitioned into classes comprise the APS measures reflecting the subjective and objective reports of life circumstances. Latent class partition is estimated by fitting the model. At the stage of selecting the variables for the clustering exercise, sex did not show distinct partition between classes which motivated its exclusion from the clustering part of the model. We retained it as a covariate variable in the model, which effectively estimates probabilities of belonging to each class given person's sex.

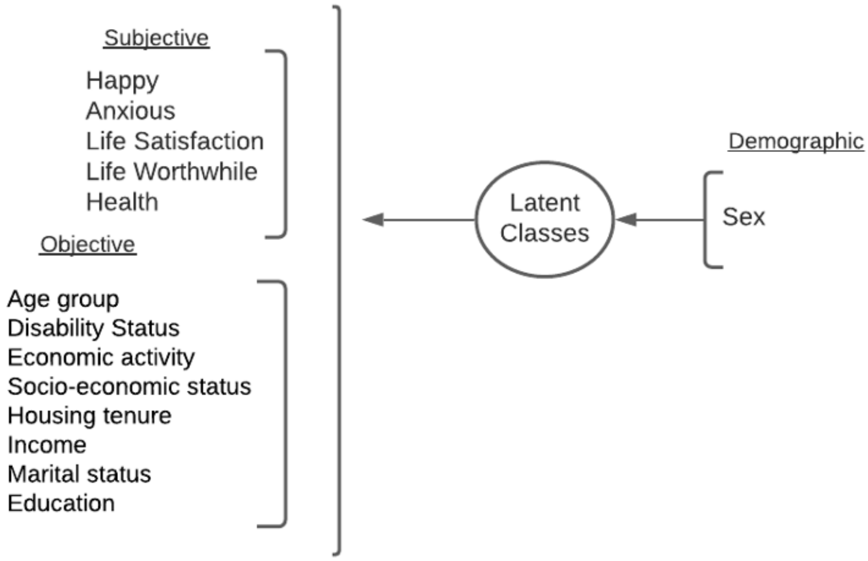


FIGURE J.1: Latent class analysis (LCA) model for the life circumstances typology.

Mixture modelling is a widely applied data analysis technique used to identify unobserved heterogeneity in a population. The key function of the finite mixture models is to express the overall distribution of one or more variables as a mixture of (or composite of) a finite number of component distributions, usually simpler and more tractable in form than the overall distribution (Masyn, 2013). The central idea is to fit a model in which any confounding between the manifest variables can be explained by a single unobserved "latent" categorical variable.

As an example, consider the dataset (1) in Figure **??**. It lists three variables that have three response levels (Life satisfaction (*High, Medium, Low)*, Happy Yesterday (same as LS), House Ownership (*Owner, Mortgage, Rent*)), one variable with two response levels (Disability (*Yes, No*)), and one variable with K response levels (Variable J (Responses 1 to K_j)). Just for these five manifest variables, and if we assume that $K_j = 4$, there are $33 \times 2 \times 4 = 72$ possible response patterns that we might observe in the individuals 1 to N. Latent class analysis enables the researcher to group (or cluster) these responses patterns (and, thus, the individuals with those response patterns) into a smaller number of R latent classes (K<< 72) such that the response patterns for individuals within each class are more similar than response patterns across classes. For example, response patterns of Person 1 (*High, Medium, No, Mortgage, ..., Response*$_3$) and Person 3 (*High, Medium, No, Mortgage , . . . , Response*$_1$) might be grouped in the same latent class, different from that which would comprise responses of Person 2 (*Medium, High, Yes, Mortgage , . . . , Response*$_{K_j}$) and Person i (*Low, High, Yes, Mortgage, . . . , Response*$_{K_j}$). Because grouping the observed response patterns is tantamount to grouping individuals, this framing of LCA makes it more person-oriented (Masyn, 2013).

We will now explain the intuition behind the LCA-based data partition model. Table (2) in Figure **??**. presents the same example dataset reformatted to highlight the components of the optimisation process. Note how person 1's response "Life Satisfaction: High" in panel (1) transforms into three variables in table (2): Y_{111} = 1, Y_{112} = 0, Y_{113} = 0. The model takes all the Y_{ijk} observations and estimates the model parameters in table (3) Figure **??**, using the following log-likelihood function:

$$lnL = \ln L = \sum_{i=1}^{N} \ln \sum_{r=1}^{R} p_r = \prod_{j=1}^{J} \prod_{k=1}^{K_J} (\pi_{jrk})^{Y_{ijk}} \tag{J.1}$$

With respect to p_r and π_{jrk} using the expectation-maximization (EM) algorithm (Dempster, Laird, and Rubin, 1977) . Here, p_r (the bottom row of table (3) Figure J.2) denotes the class-specific proportions of the sample. For example, $p_1 = 0.15$ indicates that 15% of the sample were classified as Class 1. In turn, π_{jrk} denotes the estimates of outcome probabilities conditional on belonging to class r. For example, values ($\pi_{111} = 0.9$,π_{112}= 0.04, $\pi_{113} = 0.01$) would suggest that conditional on belonging to Class 1, an individual would rate their LS as high with a 90% probability, whilst as medium with 4%, and low with 1% probability. Likewise, values ($\pi_{121} = 0.3$, $\pi_{131} = 0.1$) would indicate that individuals in classes 2 and 3 would rate their LS as high with probabilities 30% and 10% respectively.

N by J	Variable 1: **Life Satisfaction** *(High, Medium, Low)*	Variable 2: **Happy Yesterday** *(High, Medium, Low)*	Variable 3: **Disability** *(Yes, No)*	Variable 4: **House Ownership** *(Rent, Mortgage, Owner)*	...	**Variable J** (Responses *1 to K_j)*
Person 1	High	Medium	No	Mortgage	…	Response 4
Person 2	Medium	High	Yes	Mortgage	...	Response 3
Person 3	High	Medium	No	Mortgage	...	Response 1
⋮	⋮	⋮	⋮	⋮	⋮	⋮
Person N	Low	High	Yes	Mortgage	...	Response K

(1)

N by K_j	Life Satisfaction (j = 1)			...	Variable J (Responses 1 to K)				
	High (Y_{11})	Medium (Y_{12})	Low (Y_{13})	...	Y_{J1}	Y_{J2}	Y_{J3}	…	Y_{JK}
Person 1	1 (Y_{111})	0 (Y_{112})	0 (Y_{113})	...	0	0	0 (Y_{1J3})	…	1 (Y_{1JK})
Person 2	0 (Y_{211})	1 (Y_{212})	0 (Y_{213})	...	0	0	1(Y_{2J3})	…	0 (Y_{2JK})
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Person N	0 (Y_{N11})	0 (Y_{N12})	1 (Y_{N13})	...	0	0	1 (Y_{NJ3})	…	Y_{NJK}

(2)

K_j by R	Class 1	Class 2	...	Class R
Y_{11} (LS: High)	π_{111}	π_{121}	...	π_{1R1}
Y_{12} (LS: Medium)	π_{112}	π_{122}	...	π_{1R2}
Y_{13} (LS: Low)	π_{113}	π_{123}	...	π_{1R3}
⋮	⋮	⋮	⋮	⋮
Y_{JK} (Var J: K)	π_{J11}	π_{J21}	...	π_{JRK}
Proportion of sample in the class	p_1	p_2	...	p_K

(3)

FIGURE J.2: (1) Illustration of the initial data structure; (2) illustration of the same dataset, as the input for the model; (3) illustration the output of the LCA model.

Appendix K

Fitting the LCA model

The LCA model is fitted by Maximum Likelihood (ML) using the EM algorithm with the following steps. (1) start with random initial probabilities (i.e. random split of people into classes on all observable characteristics), (2) maximize the log-likelihood (LL) function (reclassify people based on an improvement criterion), (3) update the probabilities (based on the posterior distribution), (4) repeat (2) and (3) until no further improvement is possible more (LL is at the maximum value).

The analysis used the raw matrix of the variables available in the APS dataset (see Appendix A.3 for the list of variables). Fitted models with different numbers of classes were compared on the goodness-of-fit statistics reported in Tables K.1 and K.2 . The ABIC, CAIC, Chi-sq and Likelihood Ratio are all used to measure the goodness-of-fit and differ with respect to how additional parameters are penalize. Overall, a lower value of the information criterion suggests a better balance between model fit and parsimony. The second function of this process was to select the variables that allow for maximally distinct classes to emerge – according to Nagin, 1999 the rule of thumb for the acceptable group classification is that the average posterior probability of correct group membership assignment is $\geq$0.80. In general, entropy with values approaching 1 indicate clear delineation of classes (Celeux and Soromenho, 1996). Hence, we run the clustering algorithm to achieve the optimal goodness-of-fit statistics for the given model specification and then collapsed variable categories that tended to same class into broader categories for this variable.

Specifically, to establish the appropriate class number, we took 50 random samples of 60,000 (~33% of total sample) and run the clustering code (*poLCA*) for number of classes (n) from 1 to 10 on each of them. For each run, we have set the number of repetitions (*nrep*) to 30 and maximum iterations (*maxiter*) to 4000. A high number of repetitions and iterations allows the model to restart from new random initial values which is crucial for finding the global rather than local maximum.

TABLE K.1: Numeric values of the Information Criteria

Model	BIC	Chi_2	Lik_ratio	ABIC	CAIC	Entropy
Model 1	1387713	1932453083	391778.2	1387612	1387745	0
Model 2	1260129	1941158447	263824	1259919	1260195	0.933
Model 3	1209784	8169251	213164.3	1209466	1209884	0.904
Model 4	1184085	6958206	187195.4	1183659	1184219	0.861
Model 5	1166762	6430045	169716.2	1166228	1166930	0.888
Model 6	1154862	5453477	157529.7	1154220	1155064	0.873
Model 7	1144518	6842326	147127.4	1143768	1144754	0.87
Model 8	1262374	1767018171	263730.7	1261515	1262644	0.933
Model 9	1238450	11161315	221672.5	1237484	1238754	0.913

Model 10	1246519	9172346	225792.7	1245445	1246857	0.933

Once obtaining the optimal number of classes for the selected model specification, in accordance with the best practices (see e.g., Hagenaars and McCutcheon, 2002) we run the model with seven classes multiple times to be reasonably certain that we have found the parameter estimates that produce the global maximum likelihood solution. A well-known drawback of the Expectation-Maximisation (EM) algorithm is that depending upon the initial parameter values chosen in the first iteration, the algorithm may only find a local, rather than the global, maximum of the log-likelihood function (McLachlan and Krishnan, 2007.

To avoid these local maxima, it is a standard to run *poLCA* with the same model specification and same number of classes multiple times using different starting values, to locate the estimated model parameters that correspond to the model with the global maximum likelihood. Upon rerunning the model 50 times we observe convergence to the same maximum log-likelihood value. Hence, we can be reasonably sure we found the global maximum for the given specification. Additionally, we looked to the smallest estimated class size to verify that is not close to zero which would indicate non-convergence of the model. In our specification, the smallest class size in the full-sample model was estimated as 7.3% which indicates a successful convergence.

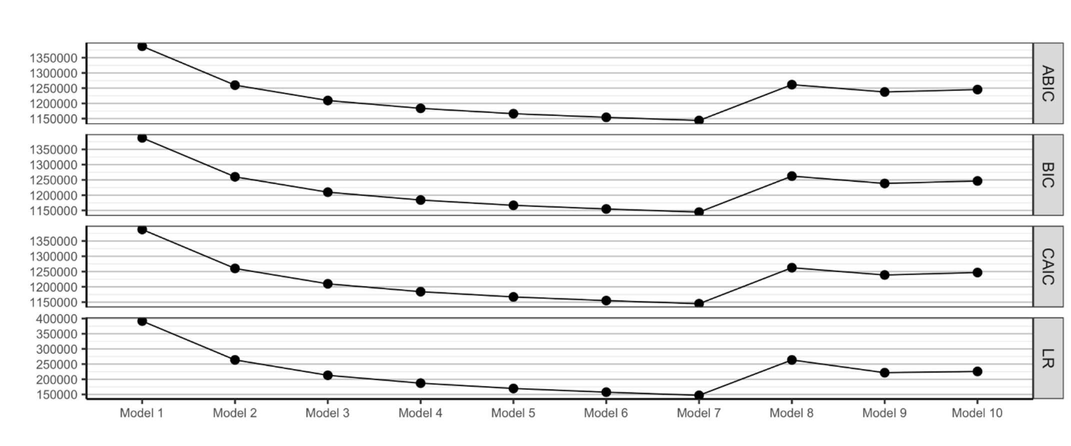


FIGURE K.1: Elbow plot of Information Criteria, n=60’000 subsample. The lowest value indicated the optimal number of classes.

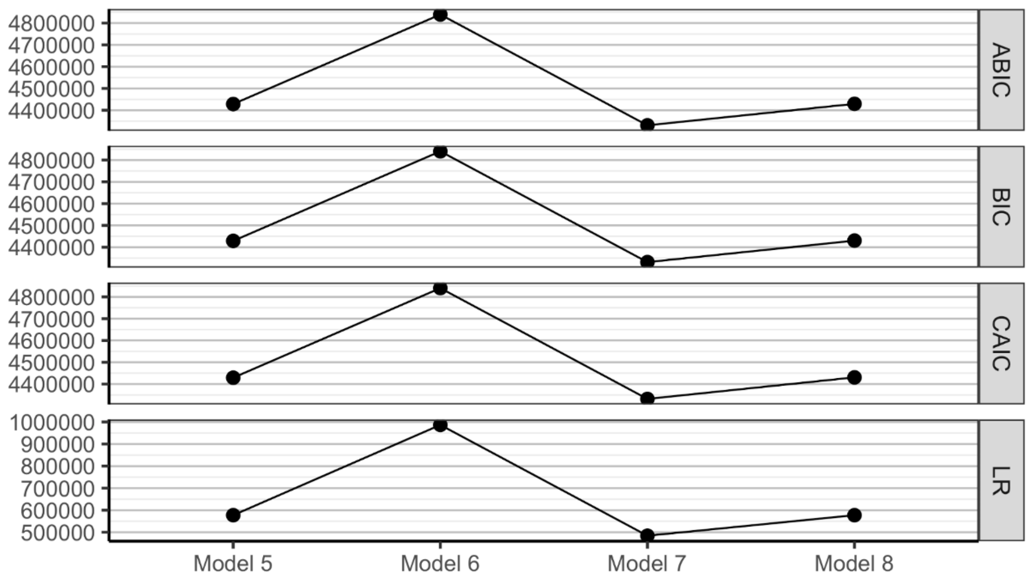


FIGURE K.2: Elbow plot of Information Criteria, full sample ($n = 215,758$).

Appendix L

Consistency check using binary logistic regressions

In a robustness check, we use logistic regression with a binary outcome of being/not being miserable to examine the odds associated with a given predictor controlling for all the others. We include the standard determinants from the SWB literature: age, sex, sexual orientation, marital status, health, employment status, socio-economic status, property ownership, religion, and ethnicity.

The logistic regression with misery as outcome variable estimates how an individual's characteristics relate to their odds of being miserable. Health, disability status and employment status emerge as the factors most strongly associated with misery once other covariates are controlled for (Table **??**, model 1). A person in fair health had 3.45 times greater odds of being miserable than the same person in good or very good health. The odds of misery for a person in bad or very bad health were over 12 times higher than those of a person in good or very good health. Having a disability had a smaller, albeit sizeable effect on misery – the odds of being miserable for a disabled person were 1.82 those of a non-disabled one.

Unemployment is strongly associated with misery: compared to the employed individuals, the unemployed had 2.91 times greater odds of being miserable. The APS questionnaire allows for a distinction between being just economically inactive and inactive with a long-term sickness or disability. Predictably, being both economically inactive and having a long-term sickness and disability was associated with a greater increase in odds of misery than inactivity alone. Economically inactive individuals had 1.56 times greater odds of misery than the employed ones, whereas for those in category 'inactive (long-term sick/disabled) the odds of misery are 2.47 greater than those of the employed ones. For average marginal effects see Table L.2.

While health, disability and employment status are the most important predictors of misery, we observed significant effects in other common life factors. Ranked in terms of vulnerability to misery these are: socio-economic status (having a semi-routine or routine occupation was associated with greater odds of misery, compared to holding a managerial job), education (individuals with an A-level were less likely to be miserable than those with basic or no education; interestingly, education beyond A-level did not appear to matter), housing tenure (house owners were less likely to be miserable), ethnicity and religiosity (non- white British and non-atheist individuals were less likely to be miserable than their white British and religious counterparts), and marital status (couples were less miserable than single people).

We also observed the classic U-shape relationship between SWB and age9:

the odds of misery of individuals aged 30-39 were 1.54 those of individuals aged 16-29; they increased further for the individuals aged 40-49, whose odds of misery were 1.80 those of the reference group; and around retirement age odds of misery decreased again, not differing significantly from those of the reference group aged 16-29. Notably, in the LCA analysis we identify heterogeneity behind this overall coefficient in the 16-29 age group. Class 4 comprising predominantly respondents of this age, healthy, yet out of employment or in lower SES had larger proportion of respondents classified as miserable compared to the subsequent three classes (5, 6, and 7) comprising more than half of the sample.

TABLE L.1: Covariates of Misery: (1) unweighted, (2) weighted

	Dependent variable: Miserable *logistic* -1	*survey-weighted logistic* -2
Sex (Ref: Male)Female	-0.09* (0.05)	-0.13** (0.06)
Age (Ref: 16-29) 30-39	0.43*** (0.11)	0.44*** (0.13)
Age: 40-49	0.59*** (0.10)	0.63*** (0.13)
Age: 50-59	0.51*** (0.11)	0.59*** (0.14)
Age: 60-69	0.002 (0.12)	0.03 (0.16)
Age: 70-99	-0.82 (0.59)	-0.73 (0.65)
Health (Ref: Good/Very Good) Fair	1.24*** (0.08)	1.34*** (0.11)
Health: Bad/Very Bad	2.50*** (0.09)	2.60*** (0.12)
Disability Status (Ref: Not Disabled) Disabled	0.60*** (0.08)	0.65*** (0.11)
Marital Status (Ref: Single) Married/Civil Partnership	-0.69*** (0.06)	-0.72*** (0.08)
Married/Partner(separated)	0.14 (0.09)	0.03 (0.11)
Divorced/Dissolved	0.06 (0.06)	0.02 (0.08)
Widowed	0.17 (0.11)	0.09 (0.13)
Education (Ref: Basic/None) GCSE	-0.16** (0.07)	-0.18* (0.09)
Degree/Professional	-0.001 (0.08)	-0.01 (0.10)
GCSE	-0.07 (0.07)	-0.05 (0.08)
Other qualification	0.04 (0.08)	0.03 (0.09)
Economic Activity (Ref: Employee) Unpaid Family Worker	-8.04 (134.83)	-8.66*** (0.79)
Unemployed	1.07*** (0.11)	1.04*** (0.13)
Inactive	0.45*** (0.09)	0.33*** (0.12)
Inactive(LT sick/disab)	0.90*** (0.08)	0.80*** (0.10)
Retired	-0.20 (0.13)	-0.29* (0.15)
Self-employed	0.10 (0.14)	0.0002 (0.17)
Student	-0.19 (0.37)	0.06 (0.56)
Socio-Econ Status (Ref: Managerial) Intermediate/Lower Supervisory	0.11 (0.10)	0.08 (0.12)
Semi/Routine Occupation	0.22** (0.09)	0.25** (0.12)
Small employer/Own Account	0.19 (0.13)	0.18 (0.16)
Full Time Student	-0.03 (0.24)	-0.25 (0.38)
Never worked/LT Unemployed	0.12 (0.12)	0.15 (0.16)
Not Classifiable	0.12 (0.09)	0.14 (0.12)
Housing Tenure (Ref: Rent) Mortgage	-0.10 (0.07)	-0.15* (0.08)
House owner	-0.15** (0.07)	-0.07 (0.10)
Not White British	-0.26*** (0.08)	-0.24** (0.10)
Religious	-0.16*** (0.05)	-0.13** (0.06)
Non-Heterosexual	-0.15 (0.13)	-0.29* (0.17)

Constant	-5.84*** (0.19)		-5.79*** (0.25)
Observations		177,760	177,760
Log Likelihood		-8,915.80	
Akaike Inf. Crit.		17,903.60	
Note: *p<0.1; **p<0.05; ***p<0.01			

TABLE L.2: Average marginal effects (average effect of the covariate on the probability (0-100%) of misery) of the unweighted logistic model

	AME (change in probability (0-100% likely) of misery)	SE
Sex: Female	-0.10*	0.05
Age: 30-39	0.10***	0.1
Age: 40-49	0.65***	0.1
Age: 50-59	0.54***	0.1
Age: 60-69	0	0.1
Age: 70-99	-0.49**	0.25
Disabled	0.63***	0.08
Health: Fair	0.86***	0.07
Health: Bad/Very Bad	3.72***	0.22
Married/Civil partner	-0.70***	0.07
Married/Partner (separated)	0.19	0.13
Divorced/ Partnership Dissolved	0.08	0.09
Widowed	0.25	0.16
Education: GCSE	-0.08	0.08
Education: A-level	-0.18**	0.08
Education: degree/Professionsl	0	0.09
Education: Other	0.05	0.09
Econ. Act.: Unemployed	1.36***	0.18
Econ. Act.: Inactive	0.42***	0.09
Econ. Act.: Inactive (Long-term sick/disabled)	1.06***	0.1
Econ. Act.: Family work[1]	-0.79***	0.06
Econ. Act.: Retired	-0.14	0.09
Econ. Act.: Self-employed	0.08	0.12
Econ. Act.: Student	-0.14	0.24
SES: Intermediate/Lower Superv	0.12	0.1
SES: Never work/LT Unempl	0.13	0.13
SES: Not classifiable	0.12	0.1
SES: Semi/Routine	0.25**	0.1
SES: Small Employer	0.22	0.15
Housing tenure owner	-0.16**	0.08
Housing tenure: Mortgage	-0.11	0.08
Non-Heterosexual	0.17	0.15
Not White British	-0.27***	0.08
Religious	-0.18	0.05

Note: *p< 0.1, **p< 0.05, ***p<0.01. Delta-method estimated standard errors.

Appendix M

Experimental Instructions

M.1 Experimental instructions: Self-selection bias

Experimental Instructions - English

Introduction

Welcome to our survey!
Thank you very much for your participation!

This research is carried out by the City, University of London and is not expected to last more than 15 minutes.

The research is anonymous , i.e. the data collected will not be linked to any information that may be related to you personally. This anonymous data can be accessed by other academics and summarized in an article - a fact that is a regular practice in this field.
If you have any concerns, questions or concerns about this research, you can contact the head of the research team[...].

The present research has been approved by the Ethics Research Ethics Committee of the City University of London.

If you are over 18, have read and understood the above and would like to participate, please click "Agree" to start the survey.

Please insert the participant ID you received with the invitation for this survey, exactly as you received it.

Background Questions

(detailed answer options available in the Supplementary Materials - Coding file)

- Have you been testing Covid-19 in the last 5 days?
- Do you have symptoms right now?
- How likely do you think it is to have COVID-19 right now?
- How often have you left your home in the last 7 days?
- Have you had COVID-19 in the past?
- If a COVID-19 vaccine, available free of charge from local health authorities, is available in January, would you like to have it?
- How likely do you think it is to get COVID-19 within the next year?

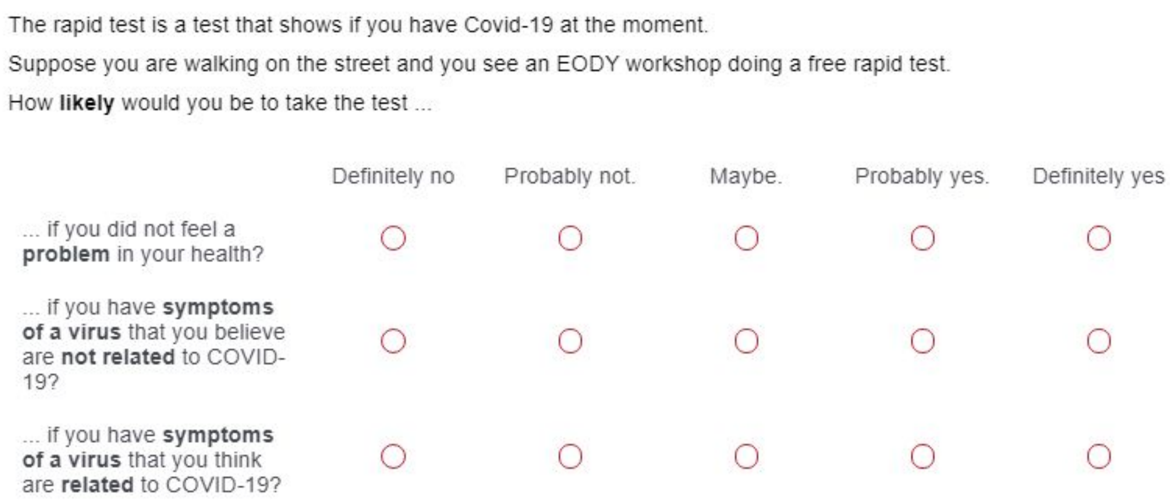
- Have you had face-to-face contact with a confirmed Covid-19 case in the last 7 days ?

Willingness to take the test, conditional on **Symptoms**

The rapid test is a test that shows if you have Covid-19 at the moment.
Suppose you are walking on the street and you see an EODY workshop doing a free rapid test.
How likely would you be to take the test …

- ... if you did not feel a **problem** in your health?
- ... if you have **symptoms of a virus** that you believe are **not related** to COVID-19?
- ... if you have **symptoms of a virus** that you think are **related** to COVID-19?

Screen View



The rapid test is a test that shows if you have Covid-19 at the moment.

Suppose you are walking on the street and you see an EODY workshop doing a free rapid test.

How **likely** would you be to take the test ...

	Definitely no	Probably not.	Maybe.	Probably yes.	Definitely yes
... if you did not feel a **problem** in your health?	○	○	○	○	○
... if you have **symptoms of a virus** that you believe are **not related** to COVID-19?	○	○	○	○	○
... if you have **symptoms of a virus** that you think are **related** to COVID-19?	○	○	○	○	○

Willingness to Wait for the test - conditional on **Symptoms**

- In the above scenario, in which you walk on the street and you see an EODY workshop doing a free Rapid test, how long would you wait in line until your turn comes ...
  - ... if you did not feel a **problem in** your health?
  - ... if you have **symptoms of a virus** that you believe are **not related** to COVID-19?
  - ... if you have **symptoms of a virus** that you think are **related** to COVID-19?

Screen View

In the above scenario, in which you walk on the street and you see an EODY workshop doing a free Rapid test, **how long** would you **wait in line** until your turn comes: ...

	I would not do a test, even if there was no queue / waiting	I would only take the test if there was no queue / waiting.	Less than 5 minutes.	5-15 minutes	15-30 minutes	30-45 minutes	Up to an hour.	1-2 hours.	More than 2 hours.
... if you did not feel a **problem in** your health?	○	○	○	○	○	○	○	○	○
... if you have **symptoms of a virus** that you believe are **not related** to COVID-19?	○	○	○	○	○	○	○	○	○

Willingness to Wait for the test - conditional on **Beliefs**

In the above scenario, in which you walk on the street and you see an EODY workshop doing a free Rapid test, how long would you wait in line until your turn comes...

- ... If you had **no reason** to believe you have COVID-19 at the time?
- ... If you had **some reasons** to believe that you may have COVID-19?
- ... If you had **serious reasons** to believe you have COVID-19?

Screen View

In the above scenario, in which you walk on the street and you see an EODY workshop doing a free Rapid test, **how long** would you **wait in line** until your turn comes: ...

	I would not do a test, even if there was no queue / waiting	I would only take the test if there was no queue / waiting.	Less than 5 minutes.	5-15 minutes	15-30 minutes	30-45 minutes	Up to an hour.	1-2 hours.	More than 2 hours.
If you had **no reason** to believe you have COVID-19 at the time	○	○	○	○	○	○	○	○	○
If you had **any reason** to believe that you may have COVID-19	○	○	○	○	○	○	○	○	○
If you had **serious reasons** to believe you have COVID-19	○	○	○	○	○	○	○	○	○

Risky Environments

(detailed answer options available in the Supplementary Materials - Coding file)

- Are you currently using Public Transport ? If so, how often?
- As part of your job , how often do you interact in person (not over the phone or online) with people you do not live in the same house with?
- As part of your work , how many people do you interact with in person?
- Of the people you live with **in the same house** , is there anyone who interacts in person with other people in the context of their work?
- Do you belong to a vulnerable group (regarding COVID-19)?
- Does one of the people you live in the same house belong to a vulnerable group (as far as COVID-19 is concerned)?
- Please check the box below if you can take a free COVID-19 test at work.
- Do you currently see family members with whom you do not live in the same house?

- At Christmas, do you plan to see family members with whom you do not live in the same house?

Demographics

(detailed answer options available in the Supplementary Materials - Coding file)

- What is your gender?
- What is your marital status?
- Do you live with your partner?
- What is the highest level of education **you have received or started**?
- What is your year of birth?
- How would you describe your general state of health, regardless of covid-19?
- Are there any restrictions on your activities due to health problems?
- Have you ever worked?
- What is your current professional status?
- If you retired in what field was your main job?
- What is your height (in centimeters)?
- This question is purely for statistical reasons - remember: research is anonymous.
- What is your weight (in kilos)?
- This question is purely for statistical reasons - remember: research is anonymous.
- What is your post code? (optional)

Breaking lockdown rules - 3 treatments

1. Consider an app that for next week would pay you 5 euros for each day you do not leave the house between 11 and 12 at night, unless it was a real emergency. Would you put it on your cell phone?
2. Consider an application that for the next week would pay you 5 euros for each day you did not meet friends at your home or home, unless it was a real emergency. Would you put it on your cell phone?
3. Consider an application that for the next week would pay you 5 euros for each day you would NOT meet friends or relatives (other than roommates), in your home or home, unless of course it was a real emergency. Would you put it on your cell phone?

Screen View

Consider an app that for next week would pay you 5 euros for each day you do not leave the house between 11 and 12 at night, unless it was a real emergency.

Would you put it on your cell phone?

- Yes
- No

Experts question - 2 treatments

This question is OPTIONAL. You can skip it if you want.There is a correct answer and if you find it, an additional € 0.30 will be added to your fee.

1. In a 10-question questionnaire about Japanese geography, Dimitris answered 4 questions correctly.

There are two advisors, K and Z.
One is an expert on the subject and answered 8 questions of the above questionnaire correctly.
The other has no corresponding specialization and answered only 5 questions correctly.
Dimitris **does not know** which of the two advisers is the expert on the subject and which is not.
But he **knows** that with consultant K he had 4 common answers, while with consultant Z 6 common.
Based on **this information alone** , which of the two advisers, K or Z, is most likely to be an expert on the subject?

2. In a 10-question questionnaire about Japanese geography, Dimitris answered 4 questions correctly.

There are two advisors, K and Z.
One is an expert on the subject and answered 8 questions of the above questionnaire correctly.
The other has no corresponding specialization and answered only 5 questions correctly.
Dimitris **does not know** which of the two advisers is the expert on the subject and which is not.
But he **knows** that he had 3 common answers with consultant K, while he had 6 common answers with consultant Z.
Based on **this information alone** , which of the two advisers, K or Z, is most likely to be an expert on the subject?

Incentivised waiting task - 2 treatments

1. Covid-19 test treatment

The obligatory part of the research is completed - thank you for your time!
Your answers have been recorded and you will receive a fee of € 1.5 for your time, plus € 0.3 for the last question, if you answered it correctly.
Now, you have the OPTION to take part in our lottery, with a 1 in 30 chance of winning a € 80 voucher for a COVID-19 molecular test at home .
If you win you will receive the voucher within 36 hours!
The next screen will follow a process that will take a few minutes. You will be informed exactly about the duration and what you need to do.

Screen View

The obligatory part of the research is completed - thank you for your time!
Your answers have been recorded and you will receive a fee of € 1.5 for your time, plus € 0.3 for the last question, if you answered it correctly.
Now, you have the OPTION to take part in our lottery, with a 1 in 30 chance of winning a **€ 80** voucher for a **COVID-19 molecular test at home** .
If you win you will receive the voucher within 36 hours!
The next screen will follow a process that will take a few minutes. You will be informed exactly about the duration and what you need to do.

2. Bookshop voucher treatment

The obligatory part of the research is completed - thank you for your time!
Now, you have the OPTION to take part in our lottery, with a 1 in 30 chance of winning a € 80 voucher from Public stores .
If you win you will receive the voucher within 36 hours!
The next screen will follow a process that will take a few minutes. You will be informed exactly about the duration and what you need to do.

Screen View

The obligatory part of the research is completed - thank you for your time!
Now, you have the OPTION to take part in our lottery, with a 1 in 30 chance of winning a **€ 80** voucher from **Public stores** .
If you win you will receive the voucher within 36 hours!
The next screen will follow a process that will take a few minutes. You will be informed exactly about the duration and what you need to do.

Incentivised waiting task - Implementation

1. Instructions at Qualtrics end screen

Please copy your entry code (your email) to use it in the lottery!
For security reasons it is a new password and not your email
Your password is: ${q://QID70/ChoiceTextEntryValue}

Click here to proceed:
http://georgana.net/sotiris/task/atten/
You can now close this window.

2. Instructions at task starting screen

What is your password?

No password, you DO NOT PAY for this stage.

In total the process will take XXX seconds.

XXX will be revealed when you start. If you complete the process to the end, you will be able to participate in the lottery.

Close the window whenever you want, but there is no second chance to participate
(do not press refresh, participation is canceled).

A button will appear at unknown times.

You have 4 seconds to press the button when it appears.

The button will appear here

In this form Button

Okay;

Start!

3. Instructions during waiting task

Wait time = 300 seconds

Your password is: EXAMPLE

In total the process will take 300 seconds.

You have 4 seconds to press the button when it appears.

They live 277

Okay, wait for it to show up again.

Wait time = 600 seconds

Your password is: MY_EXPERIMENT_ID

In total the process will take 600 seconds.

You have 4 seconds to press the button when it appears.

There are 595 left

Press me

3.1 Instructions if task failed

Your password is: MY_EXPERIMENT_ID

In total the process will take 600 seconds.

You have 4 seconds to press the button when it appears.

Unfortunately you failed to press the button.

Do not restart the test, attempts are recorded and there is no second chance.

You can close the window.

And remember, covid-19 is a real threat. Follow the instructions of the authorities .

Wear a mask, keep your distance, wash your hands, prefer well-ventilated areas. Stay healthy!

3.2 Instructions if task **completed**

1. Choice between initial prize (treatment 1 - covid-19 test, treatment 2 - bookshop voucher) OR cash (participants randomly allocated to treatment of {€20, €35, €50,€65}
2. Choice of lottery number, {1:30}

Well done for your patience!

Now it's time for the draw, choosing a blue lottery. But first you have a choice:

IF you win the draw, do you prefer to receive € 20 in cash (payable as usual) instead of the € 80 prize as mentioned before?

No, I want the gift ○
Yes, I want fluid ○
Okay, choose a lottery ticket

1	11	21
2	12	22
3	13	23
4	14	24
5	15	25
6	16	26
7	17	27
8	18	28
9	19	29
10	20	30

4. Lottery draw screen

13 was drawn!

You had chosen 24.

Unfortunately you did not win.

Do not restart the page, attempts are recorded
and there is no second chance.

You can close the window.

And remember, covid-19 is a real threat. Follow the instructions of the authorities .

Wear a mask, keep your distance, wash your hands, prefer well-ventilated areas. Stay healthy!

Experimental Instructions - Greek

Introduction

Καλώς ήλθατε στην έρευνά μας!
Ευχαριστούμε θερμά για την συμμετοχή σας!

Η συγκεκριμένη έρευνα πραγματοποιείται από το City, Πανεπιστήμιο του Λονδίνου και δεν αναμένεται να διαρκέσει πάνω από 15 λεπτά.

Η έρευνα είναι ανώνυμη, δηλαδή τα στοιχεία που θα συλλεχθούν δεν είναι συνδεδεμένα με καμία πληροφορία που να μπορεί να συσχετιστεί με εσάς προσωπικά. Σε αυτά τα ανώνυμα δεδομένα μπορεί να έχουν πρόσβαση και άλλοι ακαδημαϊκοί και να συνοψιστούν σε κάποιο άρθρο - γεγονός που αποτελεί πάγια τακτική στον συγκεκριμένο κλάδο.
Εάν έχετε κάποιο προβληματισμό, ερώτηση ή ανησυχία σχετικά με την συγκεκριμένη έρευνα, μπορείτε να επικοινωνήσετε με την υπεύθυνη της ερευνητικής ομάδας, κα[...].

Η παρούσα έρευνα έχει πάρει έγκριση από την Επιτροπή Ηθικής Δεοντολογίας Ερευνών του Πανεπιστημίου City του Λονδίνου.

Εάν είστε πάνω από 18 ετών, έχετε διαβάσει και κατανοήσει τα παραπάνω και επιθυμείτε να συμμετάσχετε, παρακαλώ πατήστε "Συμφωνώ" για να ξεκινήσετε την έρευνα.

Παρακαλώ εισάγετε τον κωδικό συμμετοχής σας (Participant ID) που λάβατε στην πρόσκληση για την έρευνα, ακριβώς όπως τον λάβατε.

Background Questions

- Έχετε κάνει τεστ για Covid-19 τις τελευταίες 5 ημέρες;
- Έχετε συμπτώματα αυτή τη στιγμή;
- Πόσο πιθανό πιστευετε οτι ειναι να έχετε COVID-19 αυτή τη στιγμή;
- Πόσο συχνά βγήκατε από το σπίτι σας τις τελευταίες 7 ημέρες;
- Έχετε νοσήσει από COVID-19 στο παρελθόν;
- Εαν ένα εμβόλιο κατά του COVID-19, το οποίο θα διατίθεται δωρεάν από τις τοπικές αρχές υγείας, είναι διαθέσιμο τον Ιανουάριο, θα θέλατε να το κάνετε;
- Πόσο πιθανό πιστεύετε πως είναι το να κολλήσετε COVID-19 μέσα στο επόμενο έτος;
- Έχετε έρθει σε επαφή κατά πρόσωπο με επιβεβαιωμένο κρούσμα Covid-19 τις τελευταίες 7 ημέρες;

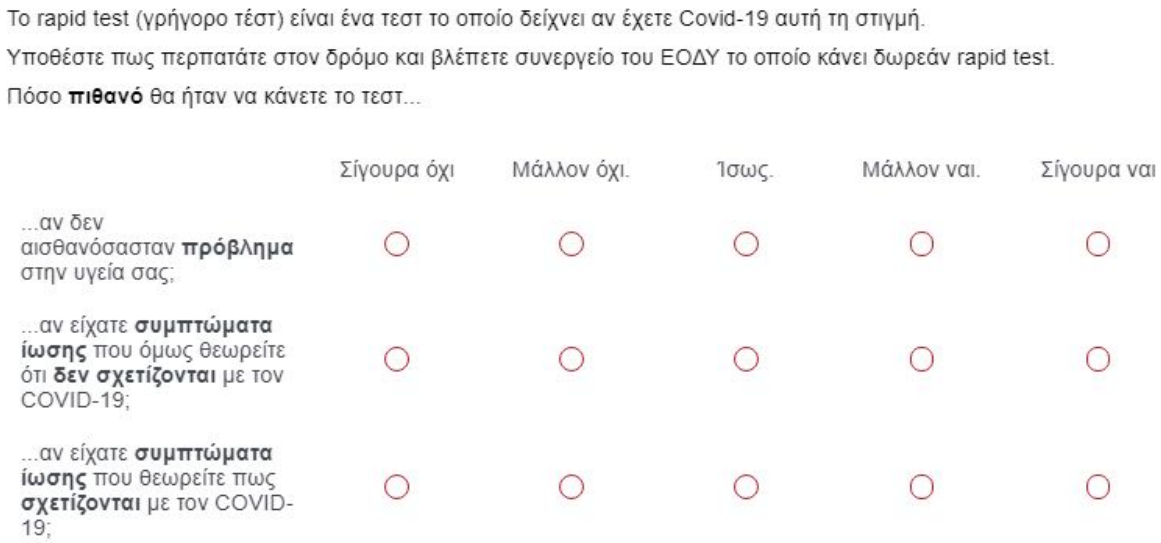
Willingness to take the test, conditional on **Symptoms**

Το rapid test (γρήγορο τεστ) είναι ένα τεστ το οποίο δείχνει αν έχετε Covid-19 αυτή τη στιγμή.
Υποθέστε πως περπατάτε στον δρόμο και βλέπετε συνεργείο του ΕΟΔΥ το οποίο κάνει δωρεάν rapid test.
Πόσο πιθανό θα ήταν να κάνετε το τεστ…

- ...αν δεν αισθανόσασταν κανένα **πρόβλημα** στην υγεία σας;
- ...αν είχατε **συμπτώματα ίωσης** που όμως θεωρείτε ότι **δεν σχετίζονται** με τον COVID-19;

- ...αν είχατε **συμπτώματα ίωσης** που θεωρείτε πως **σχετίζονται** με τον COVID-19;

Screen View


Το rapid test (γρήγορο τέστ) είναι ένα τεστ το οποίο δείχνει αν έχετε Covid-19 αυτή τη στιγμή.
Υποθέστε πως περπατάτε στον δρόμο και βλέπετε συνεργείο του ΕΟΔΥ το οποίο κάνει δωρεάν rapid test.
Πόσο **πιθανό** θα ήταν να κάνετε το τεστ...

	Σίγουρα όχι	Μάλλον όχι.	Ίσως.	Μάλλον ναι.	Σίγουρα ναι
...αν δεν αισθανόσασταν **πρόβλημα** στην υγεία σας;	○	○	○	○	○
...αν είχατε **συμπτώματα ίωσης** που όμως θεωρείτε ότι **δεν σχετίζονται** με τον COVID-19;	○	○	○	○	○
...αν είχατε **συμπτώματα ίωσης** που θεωρείτε πως **σχετίζονται** με τον COVID-19;	○	○	○	○	○

Willingness to Wait for the test - conditional on **Symptoms**

Στο παραπάνω σενάριο, στο οποίο περπατάτε στον δρόμο και βλέπετε συνεργείο του ΕΟΔΥ το οποίο κάνει δωρεάν Rapid τεστ, πόσο χρόνο θα περιμένατε στην ουρά μέχρι να έρθει η σειρά σας...

- ...αν δεν αισθανόσασταν κανένα **πρόβλημα** στην υγεία σας;
- ...αν είχατε **συμπτώματα ίωσης** που όμως θεωρείτε ότι **δεν σχετίζονται** με τον COVID-19;
- ...αν είχατε **συμπτώματα ίωσης** που θεωρείτε πως **σχετίζονται** με τον COVID-19;

Screen View

Στο παραπάνω σενάριο, στο οποίο περπατάτε στον δρόμο και βλέπετε συνεργείο του ΕΟΔΥ το οποίο κάνει δωρεάν Rapid τεστ, **πόσο χρόνο** θα **περιμένατε στην ουρά** μέχρι να έρθει η σειρά σας:...

	Δεν θα έκανα τεστ, ακόμα και αν δεν υπήρχε καθόλου ουρά/ αναμονή	Θα έκανα το τεστ μόνο αν δεν υπήρχε καθόλου ουρά/ αναμονή.	Λιγότερο από 5 λεπτά.	5-15 λεπτά	15-30 λεπτά	30-45 λεπτά	Μέχρι μία ώρα.	1-2 ώρες.	Πάνω από 2 ώρες.
...αν δεν αισθανόσασταν **πρόβλημα στην** υγεία σας;	○	○	○	○	○	○	○	○	○
...αν είχατε **συμπτώματα ίωσης** που όμως θεωρείτε ότι **δεν σχετίζονται** με τον COVID-19;	○	○	○	○	○	○	○	○	○
...αν είχατε **συμπτώματα ίωσης** που θεωρείτε πως **σχετίζονται** με τον COVID-19;	○	○	○	○	○	○	○	○	○

Willingness to Wait for the test - conditional on **Beliefs**

Στο παραπάνω σενάριο, στο οποίο περπατάτε στον δρόμο και βλέπετε συνεργείο του ΕΟΔΥ το οποίο κάνει δωρεάν Rapid τεστ, πόσο χρόνο θα περιμένατε στην ουρά μέχρι να έρθει η σειρά σας…

- Αν δεν είχατε **κανέναν λόγο** να πιστεύετε ότι έχετε COVID-19 εκείνη τη στιγμή;
- Αν είχατε **κάποιους λόγους** να πιστεύετε ότι μπορεί να έχετε COVID-19;
- Αν είχατε **σοβαρούς λόγους** να πιστεύετε ότι έχετε COVID-19;

Screen View

Στο παραπάνω σενάριο, στο οποίο περπατάτε στον δρόμο και βλέπετε συνεργείο του ΕΟΔΥ το οποίο κάνει δωρεάν Rapid τεστ, **πόσο χρόνο** θα **περιμένατε στην ουρά** μέχρι να έρθει η σειρά σας:...

	Δεν θα έκανα τεστ, ακόμα και αν δεν υπήρχε καθόλου ουρά/ αναμονή	Θα έκανα το τεστ μόνο αν δεν υπήρχε καθόλου ουρά/ αναμονή.	Λιγότερο από 5 λεπτά.	5-15 λεπτά	15-30 λεπτά	30-45 λεπτά	Μέχρι μία ώρα.	1-2 ώρες.	Πάνω από 2 ώρες.
Αν δεν είχατε **κανέναν λόγο** να πιστεύετε ότι έχετε COVID-19 εκείνη τη στιγμή	○	○	○	○	○	○	○	○	○
Αν είχατε **κάποιους λόγους** να πιστεύετε ότι μπορεί να έχατε COVID-19	○	○	○	○	○	○	○	○	○
Αν είχατε **σοβαρούς λόγους** να πιστεύετε ότι έχετε COVID-19	○	○	○	○	○	○	○	○	○

Risky Environments

(detailed answer options available in the Supplementary Materials - Coding file)

- Αυτήν την περίοδο, χρησιμοποιείτε τα Μέσα Μαζικής Μεταφοράς; Εάν ναι, πόσο συχνά;
- Στο πλαίσιο της εργασίας σας, πόσο συχνά αλληλεπιδράτε κατά πρόσωπο (όχι μέσω τηλεφώνου ή διαδικτυακά) με άτομα με τα οποία δεν διαμένετε στο ίδιο σπίτι;
- Στο πλαίσιο της εργασίας σας, με πόσα άτομα αλληλεπιδράτε κατά πρόσωπο;
- Από τα άτομα με τα οποία διαμένετε **στο ίδιο σπίτι**, υπάρχει κάποιος/α που να αλληλεπιδρά κατά πρόσωπο με άλλα άτομα στο πλαίσιο της δουλειάς τους;
- Ανήκετε εσείς σε ευπαθή ομάδα (όσον αφορά τον COVID-19);
- Ανήκει κάποιο από τα άτομα με τα οποία διαμένετε στο ίδιο σπίτι σε ευπαθή ομάδα (όσον αφορά τον COVID-19);

- Παρακαλούμε κάντε τικ στο παρακάτω κουτί αν μπορείτε να κάνετε δωρεάν τέστ COVID-19 στην εργασία σας.
- Αυτήν την περίοδο, βλέπετε μέλη της οικογένειάς σας με τα οποία <u>δεν</u> διαμένετε στο ίδιο σπίτι;
- Τα Χριστούγεννα, σχεδιάζετε να δείτε μέλη της οικογένειάς σας με τα οποία <u>δεν</u> διαμένετε στο ίδιο σπίτι;

Demographics

(detailed answer options available in the Supplementary Materials - Coding file)

- Ποιο είναι το φύλο σας;
- Ποια είναι η οικογενειακή σας κατάσταση;
- Μένετε με τον/την σύντροφό σας;
- Ποιο είναι το ανώτατο επίπεδο εκπαίδευσης **που έχετε λάβει ή ξεκινήσει**;
- Ποιο είναι το έτος γέννησης σας;
- Πως θα χαρακτηρίζατε την γενική κατάσταση της υγείας σας, ασχέτως covid-19;
- Υπάρχει κάποιος περιορισμός στις δραστηριότητές σας εξαιτίας προβλημάτων υγείας;
- Έχετε εργαστεί ποτέ;
- Ποιά είναι η τρέχουσα επαγγελματική σας κατάσταση;
- Αν έχετε συνταξιοδοτηθεί σε ποιόν τομέα ήταν η κύρια εργασία σας;
- Ποιο είναι το ύψος σας (σε εκατοστά);
- Αυτή η ερώτηση είναι καθαρά για στατιστικούς λόγους - θυμηθείτε: η έρευνα είναι ανώνυμη.
- Ποιο είναι το βάρος σας (σε κιλά);
- Αυτή η ερώτηση είναι καθαρά για στατιστικούς λόγους - θυμηθείτε: η έρευνα είναι ανώνυμη
- Ποιος είναι ο ταχυδρομικός σας κώδικας; (προαιρετικό)

Breaking lockdown rules - 3 treatments

1. Σκεφτείτε μια εφαρμογή που για την επόμενη εβδομάδα θα σας πλήρωνε 5 ευρώ για κάθε μέρα που δεν βγαίνετε από το σπίτι μεταξύ 11 και 12 το βράδυ, εκτός αν ήταν πραγματική έκτακτη ανάγκη.mΘα την βάζατε στο κινητό σας;
2. Σκεφτείτε μια εφαρμογή που για την επόμενη εβδομάδα θα σας πλήρωνε 5 ευρώ για κάθε μέρα που δεν συναντούσατε φίλους στο σπίτι σας ή σπίτι τους, εκτός αν ήταν πραγματική έκτακτη ανάγκη. Θα την βάζατε στο κινητό σας;
3. Σκεφτειτε μια εφαρμογή που για την επόμενη εβδομάδα θα σας πλήρωνε 5 ευρώ για κάθε μέρα που ΔΕΝ θα συναντούσατε φίλους ή συγγενεις (εκτός συγκατοίκων), στο σπίτι σας ή σπίτι τους, εκτός βέβαια αν ήταν πραγματική έκτακτη ανάγκη. Θα την βάζατε στο κινητό σας;

Screen View

Σκεφτείτε μια εφαρμογή που για την επόμενη εβδομάδα θα σας πλήρωνε 5 ευρώ για κάθε μέρα που δεν βγαίνετε από το σπίτι μεταξύ 11 και 12 το βράδυ, εκτός αν ήταν πραγματική έκτακτη ανάγκη.

Θα την βάζατε στο κινητό σας;

- Ναι
- Όχι

Experts question - 2 treatments

Αυτή η ερώτηση είναι ΠΡΟΑΙΡΕΤΙΚΉ. Μπορείτε να την παρακάμψετε αν θέλετε. Υπάρχει σωστή απάντηση και αν την βρείτε θα προστεθούν €0.30 επιπλέον στην αμοιβή σας.

1. Σε ένα ερωτηματολόγιο 10 ερωτήσεων σχετικά με την Ιαπωνική γεωγραφία, ο Δημήτρης απάντησε σωστά σε 4 ερωτήσεις.

Υπάρχουν δύο σύμβουλοι, ο Κ και ο Ζ.
Ο ένας είναι ειδικός στο θέμα και απάντησε σωστά σε 8 ερωτήσεις του παραπάνω ερωτηματολογίου.
Ο άλλος δεν έχει αντίστοιχη εξειδίκευση και απάντησε σωστά μόνο σε 5 ερωτήσεις.
Ο Δημήτρης δεν γνωρίζει ποιος από τους δύο συμβούλους είναι ο ειδικός στο θέμα και ποιος όχι.
Γνωρίζει όμως ότι με τον σύμβουλο Κ είχε 4 κοινές απαντήσεις, ενώ με τον σύμβουλο Ζ 6 κοινές.
Με βάση αυτές τις πληροφορίες και μόνο, ποιος από τους δύο συμβούλους, ο Κ ή ο Ζ, είναι πιο πιθανό να είναι ειδικός στο θέμα;

2. Σε ένα ερωτηματολόγιο 10 ερωτήσεων σχετικά με την Ιαπωνική γεωγραφία, ο Δημήτρης απάντησε σωστά σε 4 ερωτήσεις.

Υπάρχουν δύο σύμβουλοι, ο Κ και ο Ζ.
Ο ένας είναι ειδικός στο θέμα και απάντησε σωστά σε 8 ερωτήσεις του παραπάνω ερωτηματολογίου.
Ο άλλος δεν έχει αντίστοιχη εξειδίκευση και απάντησε σωστά μόνο σε 5 ερωτήσεις.
Ο Δημήτρης δεν γνωρίζει ποιος από τους δύο συμβούλους είναι ο ειδικός στο θέμα και ποιος όχι.
Γνωρίζει όμως ότι με τον σύμβουλο Κ είχε 3 κοινές απαντήσεις, ενώ με τον σύμβουλο Ζ 6 κοινές.
Με βάση αυτές τις πληροφορίες και μόνο, ποιος από τους δύο συμβούλους, ο Κ ή ο Ζ, είναι πιο πιθανό να είναι ειδικός στο θέμα;

Incentivised waiting task - 2 treatments

1. Covid-19 test treatment

Το υποχρεωτικό τμήμα της έρευνας ολοκληρώθηκε - ευχαριστούμε για τον χρόνο σας!
Οι απαντήσεις σας έχουν καταγραφεί και θα λάβετε την αμοιβή των €1,5 για τον χρόνο σας, συν €0,3 για την τελευταία ερώτηση, εφόσον την απαντήσατε σωστά.
Τώρα, έχετε την ΕΠΙΛΟΓΗ να λάβετε μέρος στην λοταρία μας, με πιθανότητα 1 στις 30 να κερδίσετε ένα voucher αξίας **€80** για ένα **μοριακό τεστ COVID-19 κατ’οίκον**.
Εφόσον κερδίσετε θα λάβετε το voucher εντός 36 ωρών!
Στην επόμενη οθόνη θα ακολουθήσει μια διαδικασία που θα διαρκέσει κάποια λεπτά. Θα ενημερωθείτε ακριβώς για την διάρκεια και τι πρέπει να κάνετε.

Screen View

Το υποχρεωτικό τμήμα της έρευνας ολοκληρώθηκε - ευχαριστούμε για τον χρόνο σας!

Οι απαντήσεις σας έχουν καταγραφεί και θα λάβετε την αμοιβή των €1,5 για τον χρόνο σας, συν €0,3 για την τελευταία ερώτηση, εφόσον την απαντήσατε σωστά.

Τώρα, έχετε την ΕΠΙΛΟΓΗ να λάβετε μέρος στην λοταρία μας, με πιθανότητα 1 στις 30 να κερδίσετε ένα voucher αξίας **€80** για ένα **μοριακό τεστ COVID-19 κατ’οίκον**.
Εφόσον κερδίσετε θα λάβετε το voucher εντός 36 ωρών!

Στην επόμενη οθόνη θα ακολουθήσει μια διαδικασία που θα διαρκέσει κάποια λεπτά. Θα ενημερωθείτε ακριβώς για την διάρκεια και τι πρέπει να κάνετε.

2. Bookshop voucher treatment

Το υποχρεωτικό τμήμα της έρευνας ολοκληρώθηκε - ευχαριστούμε για τον χρόνο σας!
Τώρα, έχετε την ΕΠΙΛΟΓΗ να λάβετε μέρος στην λοταρία μας, με πιθανότητα 1 στις 30 να κερδίσετε ένα voucher αξίας €80 από τα καταστήματα Public.
Εφόσον κερδίσετε θα λάβετε το voucher εντός 36 ωρών!
Στην επόμενη οθόνη θα ακολουθήσει μια διαδικασία που θα διαρκέσει κάποια λεπτά. Θα ενημερωθείτε ακριβώς για την διάρκεια και τι πρέπει να κάνετε.

Screen View

Το υποχρεωτικό τμήμα της έρευνας ολοκληρώθηκε - ευχαριστούμε για τον χρόνο σας!

Τώρα, έχετε την ΕΠΙΛΟΓΗ να λάβετε μέρος στην λοταρία μας, με πιθανότητα 1 στις 30 να κερδίσετε ένα voucher αξίας **€80** από τα **καταστήματα Public**.
Εφόσον κερδίσετε θα λάβετε το voucher εντός 36 ωρών!

Στην επόμενη οθόνη θα ακολουθήσει μια διαδικασία που θα διαρκέσει κάποια λεπτά. Θα ενημερωθείτε ακριβώς για την διάρκεια και τι πρέπει να κάνετε.

Incentivised waiting task - Implementation

1. Instructions at Qualtrics end screen

Παρακαλούμε αντιγράψτε τον κωδικό συμμετοχή σας (η-μεηλ σας) για να τον χρησιμοποιήσετε στην λοταρία!
Για λογους ασφαλείας είναι νέος κωδικός και όχι το μεηλ σας
Ο κωδικός σας είναι:
Κλικ εδώ για να προχωρήσετε στην διαδικασία:

http://georgana.net/sotiris/task/atten/

Τώρα μπορείτε να κλείσετε αυτό το παράθυρο.

Παρακαλούμε αντιγράψτε τον κωδικό συμμετοχή σας (η-μεηλ σας) για να τον χρησιμοποιήσετε στην λοταρία!

Για λογους ασφαλείας είναι νέος κωδικός και όχι το μεηλ σας

Ο κωδικός σας είναι: EXPERIMENT_ID κλικ εδω για αντιγραφη του κωδικου

Κλικ εδώ για να προχωρήσετε στην διαδικασία:

http://georgana.net/sotiris/task/atten/

Τώρα μπορείτε να κλείσετε αυτό το παράθυρο.

2. Instructions at task starting screen

Ποιός είναι ο κωδικός σας;

Χωρίς κωδικό, ΔΕΝ ΠΛΗΡΩΝΕΣΤΕ για αυτό το στάδιο.

Συνολικά η διαδικασία θα κρατήσει XXX δευτερόλεπτα.

Το XXX θα αποκαλυφθεί όταν ξεκινήσετε. Αν ολοκληρώσετε την διαδικασία ως το τέλος, θα μπορείτε να συμμετάσχετε στην λοτταρία.

Όποτε θέλετε κλείνετε το παράθυρο, αλλά δεν υπάρχει δεύτερη ευκαιρία συμμετοχής
(μην πατήσετε ανανεωση/refresh, ακυρώνεται η συμμετοχή).

Σε άγνωστες χρονικές στιγμές θα εμφανίζεται ένα κουμπί.

Έχετε 4 δεύτερα να πατήσετε το κουμπί όταν εμφανίζεται.

Το κουμπί θα εμφανιστεί εδώ

Σε αυτή την μορφή Κουμπί

Εντάξει;

Εκκίνηση!

3. Instructions during waiting task

Wait time = 300 seconds

Ο κωδικός σας είναι:

Συνολικά η διαδικασία θα κρατήσει 300 δευτερόλεπτα.

Έχετε 4 δεύτερα να πατήσετε το κουμπί όταν εμφανίζεται.

Μένουν 179

Εντάξει, περιμένετε να εμφανιστεί πάλι.

Wait time = 600 seconds

Ο κωδικός σας είναι:

Συνολικά η διαδικασία θα κρατήσει 600 δευτερόλεπτα.

Έχετε 4 δεύτερα να πατήσετε το κουμπί όταν εμφανίζεται.

Μένουν 596

Το κουμπί θα εμφανιστεί εδώ

4. Instructions if task **completed**

Μπράβο για την υπομονή σας!

Τώρα ήρθε η ώρα για την κλήρωση, διαλέγοντας έναν μπλε λαχνό. Αλλά πρώτα έχετε μια επιλογή:

ΑΝ κερδίσετε στην κλήρωση, προτιμάτε αντί για το δώρο αξίας 80 ευρώ όπως αναφέρθηκε πριν, να λάβετε 35 ευρώ σε ρευστό (πληρωτέα όπως συνήθως);

Οχι, θέλω το δώρο ○
Ναι, θέλω ρευστό ○
Εντάξει, διαλέξτε λαχνό

1	11	21
2	12	22
3	13	23
4	14	24
5	15	25
6	16	26
7	17	27
8	18	28
9	19	29
10	20	30

5. Lottery draw screen

Κληρώθηκε 10!

Είχατε διαλέξει 26.

Δυστυχώς δεν κερδίσατε.

Μην επανεκκινήσετε την σελίδα, οι απόπειρες καταγράφονται και δεν υπάρχει δεύτερη ευκαιρία.

Μπορείτε να κλείσετε το παράθυρο.

Και να θυμάστε, ο covid-19 είναι πραγματική απειλή. Ακολουθείτε [τις οδηγίες των αρχών].

Φοράτε μάσκα, κρατάτε αποστάσεις, πλένετε τα χέρια σας, προτιμάτε καλά αεριζόμενους χώρους. Μείνετε υγιείς!

M.2 Experimental instructions: Prosociality

Experimental Instructions - English

Introduction

Welcome to our survey!
Thank you very much for your participation!

What is the purpose of this research?
This research is carried out by the City University of London and concerns volunteering and charity . By taking part in this research you contribute to a better understanding of people's behavior in relation to these activities.

What should I do?
First, you will answer a series of short questions about your demographics. Next, you will be asked for your views on charity and volunteering. You will then be given the opportunity to take a short test if you wish. Then you will answer a series of questions about quarantine . Finally, you will be given the opportunity to win 200 Euros by taking part in a lottery.
In total, the search takes about 20 minutes.

Is it mandatory to participate?
Your participation in this research is voluntary and you have the opportunity to leave at any point, closing the navigation window and informing the researchers.

What will happen next?
This research is anonymous, ie the data collected will not be linked to any information that may be related to you personally. This anonymous data can be accessed by other academics and summarized in an article - a fact that is a regular practice in this field.

What can I do if I have a problem?
If you have any concerns, questions or concerns about this research, you can contact the professor and head of the research team, [contacts]

If you are over 18, have read and understood the above and would like to participate, please click "Agree" to start the survey.

Socio-economic questions

What is your year of birth?

What is your gender?

Do you live with your partner?

What is the highest level of education you have completed?

How would you describe your health?

Is there any restriction to your activities due to health problems?

Have you ever worked?

What is your current labour status?

What is/was your profession?

In your job do (did) you supervise or manage any personnel?

Which sector did you retire from?

EU-SILC and SHARE questions

What year did you/are planning to retire in?

Did you get an early or late retirement?

- (range between 15 years before expected and 15 years after expected)

Do you provide or receive financial assistance to your relatives? Please indicate the monthly amount.

- (range between -1000 and 1000 Euros per month)
-

How many hours per week (if any) do you spend helping your children/grandchildren with childcare or home care?

- 0
- 1-10
- 10+
- I do not have children/grandchildren

Before the quarantine, how often did you meet with relatives and how often with friends?

How often do you contact your relatives and how often your friends (on the phone or internet)?

In the last 12 months, have you volunteered for an informal activity / event?

- Yes
- No

If not, why?

- Lack of time
- Lack of interest
- Other reason

During the last 12 months, have you volunteered in any formal activity?

- Yes
- No

If not, why?

- Lack of time
- Lack of interest
- Other reason

During the last 12 months, have you volunteered in any political activity?

- Yes
- No

If not, why?

- Lack of time
- Lack of interest
- Other reason

How often have you done voluntary/charity work the last 12 months?

- Almost every day.
- Almost every week.
- Almost every month .
- Less often than once per month.
- Never.

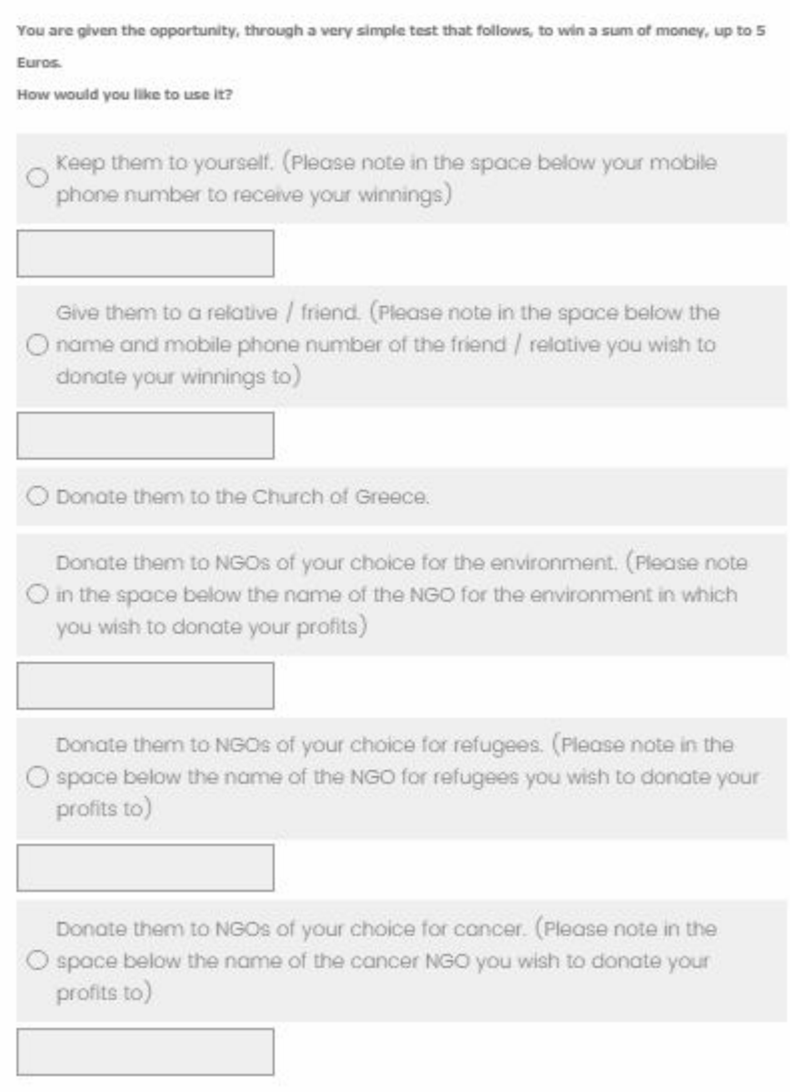
Incentivised measures

Real-effort task - Recipient choice instructions

How would you like to spend the money (up to 5 euros) earnt by the real effort task?

- ➢ Keep it for yourself.
- ➢ Give it to a relative/friend - please type the name and phone number of the person you would like to take your winnings.
- ➢ Donate it to the Church of Greece.
- ➢ Donate it to a charity (of your choice) for the environment - please type the name of the charity you would like to take your winnings..
- ➢ Donate it to a charity (of your choice) for refugees - please type the name of the charity you would like to take your winnings.
- ➢ Donate it to a charity (of your choice) for cancer - please type the name of the charity you would like to take your winnings.

Real-effort task - Recipient choice instructions - Screen view


You are given the opportunity, through a very simple test that follows, to win a sum of money, up to 5 Euros.

How would you like to use it?

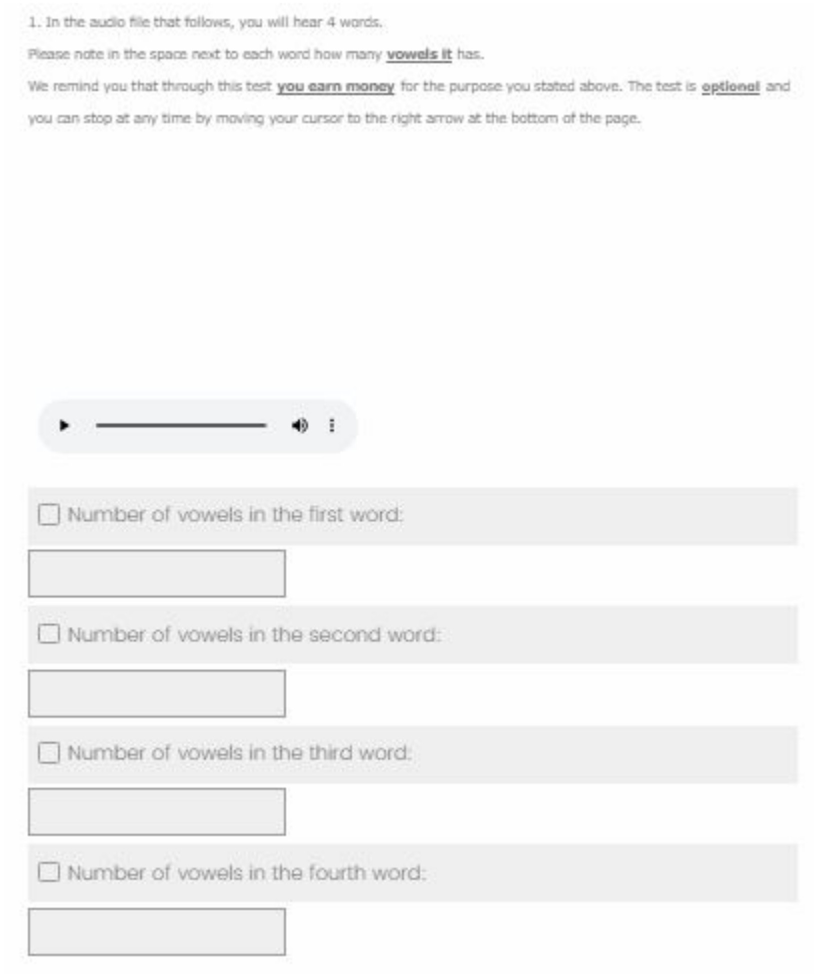
- Keep them to yourself. (Please note in the space below your mobile phone number to receive your winnings)
- Give them to a relative / friend. (Please note in the space below the name and mobile phone number of the friend / relative you wish to donate your winnings to)
- Donate them to the Church of Greece.
- Donate them to NGOs of your choice for the environment. (Please note in the space below the name of the NGO for the environment in which you wish to donate your profits)
- Donate them to NGOs of your choice for refugees. (Please note in the space below the name of the NGO for refugees you wish to donate your profits to)
- Donate them to NGOs of your choice for cancer. (Please note in the space below the name of the cancer NGO you wish to donate your profits to)

Real-effort task instructions

In the audio file that follows you will listen to 4 words.
Please note in the space following each word how many vowels the word has.
We remind you that through this task you win money for the purpose you have set above. The task is voluntary and you can stop at any point by pressing the next button at the end of the page.

- Number of vowels in the first word: ____________________.
- Number of vowels in the second word: ____________________.
- Number of vowels in the third word: ____________________.
- Number of vowels in the forth word: ____________________.

Real-effort task - Screen view



Lottery instructions

And now it's time to take part in the lottery that gives you the opportunity to win the amount of 200 Euros. In case you win, how would you like to spend it?

- Keep it for yourself.
- Give it to a relative/friend - please type the name and phone number of the person you would like to take your winnings.
- Donate it to the Church of Greece.
- Donate it to a charity (of your choice) for the environment - please type the name of the charity you would like to take your winnings..

- ➢ Donate it to a charity (of your choice) for refugees - please type the name of the charity you would like to take your winnings.
- ➢ Donate it to a charity (of your choice) for cancer - please type the name of the charity you would like to take your winnings.

Lottery instructions - Screen view

And now it's time to take part in the lottery that gives you the opportunity to win the amount of 200 Euros.

If so, how would you like to use this amount?

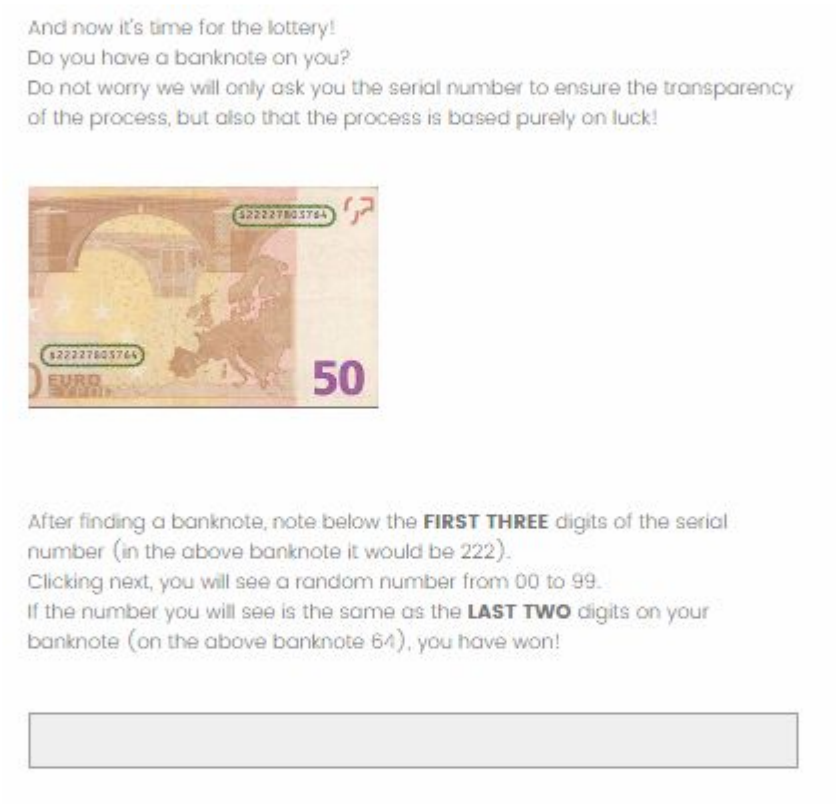
	0 20 40 60 80 100 120 140 160 180 200	
Keep them to yourself. (Please note in the space below your mobile phone number to receive your winnings)		0
Donate them to a relative / friend (Please note the mobile phone number and first name of the person you wish to donate your winnings to).		0
Donate them to the Church of Greece.		0
Donate them to NGOs of your choice for the environment. (Please note in the space below the name of the NGO for the environment in which you wish to donate your profits)		0
Donate to NGOs of your choice for refugees. (Please note in the space below the name of the refugee NGO you wish to donate your profits to)		0
Donate them to NGOs of your choice for cancer. (Please note in the space below the name of the cancer NGO you wish to donate your profits to)		0
I do not wish to take part in the lottery.		0
Total:		**0**

Lottery administration

And now it's time for the lottery!
Do you have a banknote on you?
Do not worry we will only ask you the serial number to ensure the transparency of the process, but also that the process is based purely on luck!
After finding a banknote, note below the **FIRST THREE** digits of the serial number (in the above banknote it would be 222).
Clicking next, you will see a random number from 00 to 99.
If the number you will see is the same as the **LAST TWO** digits on your banknote (on the above banknote 64), you have won!

And the number is ... [number] !!
Please note below ALL the DIGITS of the serial number of the banknote you have in your hands.
If the last two are the same as the above number , congratulations, you have won !
We will get in touch with you as soon as possible for the next steps!
IMPORTANT : Do not use the banknote, as you will need to show it as proof!

Lottery administration - screen view



And the number is ... 38 !!

Please note below **ALL the VOTES** of the serial number of the banknote you have in your hands.
If the **last two** are the **same as the above number** , congratulations, **you have won** !
We will get in touch with you as soon as possible for the next steps!
IMPORTANT : *Do not use the banknote, as you will need to show it as proof!*

Thank you very much for taking part in our research!
- For any additional information about this research, please do not hesitate to contact the lead researcher:
[Contact Information of the Responsible Researcher]

Experimental Instructions - GREEK

Καλώς ήλθατε στην έρευνά μας!
Ευχαριστούμε θερμά για την συμμετοχή σας!

Ποιος είναι ο σκοπός της συγκεκριμένης έρευνας;
Η συγκεκριμένη έρευνα πραγματοποιείται από το Πανεπιστήμιο City του Λονδίνου και αφορά τον εθελοντισμό και τη φιλανθρωπία. Λαμβάνοντας μέρος σε αυτήν την έρευνα συνεισφέρετε στην καλύτερη κατανόηση της συμπεριφοράς των ανθρώπων σχετικά με τις δραστηριότητες αυτές.

Τι πρέπει να κάνω;
Αρχικά, θα απαντήσετε σε μια σειρά σύντομων ερωτήσεων σχετικά τα δημογραφικά σας χαρακτηριστικά. Στην συνέχεια, θα ερωτηθείτε για τις απόψεις σας σχετικά με τη φιλανθρωπία και τον εθελοντισμό. Κατόπιν,θα σας δοθεί η δυνατότητα να κάνετε μια σύντομη δοκιμασία, εφόσον το επιθυμείτε. Στη συνέχεια θα απαντήσετε σε μια σειρά

ερωτήσεων σχετικά με την καραντίνα. Τέλος, θα σας δοθεί η δυνατότητα να κερδίσετε 200 Ευρώ, λαμβάνοντας μέρος σε μια λοταρία.
Συνολικά η έρευνα διαρκεί περίπου 20 λεπτά.

Είναι υποχρεωτικό το να συμμετάσχω;
Η συμμετοχή σας στην συγκεκριμένη έρευνα είναι εθελοντική και έχετε την δυνατότητα να αποχωρήσετε σε οποιοδήποτε σημείο, κλείνοντας το παράθυρο περιήγησης και ενημερώνοντας τους ερευνητές.

Τι θα συμβεί μετά;
Η συγκεκριμένη έρευνα είναι ανώνυμη, δηλαδή τα στοιχεία που θα συλλεχθούν δεν είναι συνδεδεμένα με καμία πληροφορία που να μπορεί να συσχετιστεί με εσάς προσωπικά. Σε αυτά τα ανώνυμα δεδομένα μπορεί να έχουν πρόσβαση και άλλοι ακαδημαϊκοί και να συνοψιστούν σε κάποιο άρθρο - γεγονός που αποτελεί πάγια τακτική στον συγκεκριμένο κλάδο.

Τι μπορώ να κάνω εάν έχω κάποιο πρόβλημα;
Εάν έχετε κάποιο προβληματισμό, ερώτηση ή ανησυχία σχετικά με την συγκεκριμένη έρευνα, μπορείτε να επικοινωνήσετε με τον καθηγητή και υπεύθυνο της ερευνητικής ομάδας [...]

Εάν είστε πάνω από 18 ετών, έχετε διαβάσει και κατανοήσει τα παραπάνω και επιθυμείτε να συμμετάσχετε, παρακαλώ πατήστε "Συμφωνώ" για να ξεκινήσετε την έρευνα.

Ποιο είναι το έτος γέννησης σας;
Ποιο είναι το φύλο σας;
Ποια είναι η επίσημη οικογενειακή σας κατάσταση;
Μένετε με τον/την σύντροφό σας;
Ποιο είναι το ανώτατο επίπεδο εκπαίδευσης που έχετε λάβει;
Πως θα χαρακτηρίζατε την κατάσταση της υγεία σας;
Υπάρχει κάποιος περιορισμός στις δραστηριότητές σας εξαιτίας προβλημάτων υγείας;
Έχετε εργαστεί ποτέ;
Ποιά είναι η τρέχουσα επαγγελματική σας κατάσταση;
Ποιο ήταν/είναι το κύριο επάγγελμά σας;
Στην εργασία σας, έχετε/είχατε προσωπικό υπό την επίβλεψή σας;
Αν έχετε συνταξιοδοτηθεί σε ποιόν τομέα ήταν η κύρια εργασία σας;

Ποιο έτος συνταξιοδοτηθήκατε / σχεδιάζετε να συνταξιοδοτηθείτε;
- ➢ (έτος συνταξιοδότησης)

Συνταξιοδοτηθήκατε νωρίτερα ή αργότερα από το προβλεπόμενο;
- ➢ (εύρος κυμαινόμενο μεταξύ 15 χρόνια πριν το προβλεπόμενο και 15 χρόνια μετά το προβλεπόμενο)

Παρέχετε ή λαμβάνετε οικονομική βοήθεια από συγγενείς σας;
Αν ναι, παρακαλουμε μετακινήστε την παρακάτω μπάρα στο αντίστοιχο <u>μηνιαίο</u> ποσό
Το αρνητικό πρόσημο αντιστοιχεί σε χρήματα που δίνεται και το θετικό σε χρήματα που λαμβάνετε.
- ➢ (εύρος κυμαινόμενο μεταξύ -1000 και 1000 Ευρώ)

Πόσες ώρες την εβδομάδα περνάτε παρέχοντας βοήθεια στα παιδιά/εγγόνια/άλλους συγγενείς σας φροντίζοντας τα ίδια ή το σπίτι τους?
- ➢ 0
- ➢ 1-10
- ➢ 10+
- ➢ Δεν έχω παιδιά/εγγόνια

Υπό κανονικές συνθήκες (π.χ. πριν την καραντίνα / τους τελευταίους 12 μήνες), πόσο συχνά συναντιέστε με συγγενείς και πόσο συχνά με φίλους;:
Πόσο συχνά επικοινωνείτε (τηλεφωνικά ή / και μέσω ίντερνετ) με συγγενείς και πόσο συχνά με φίλους;

Τους τελευταίους 12 μήνες, λάβατε μέρος εθελοντικά σε κάποια ανεπίσημη δραστηριότητα/εκδήλωση;
- Ναι.
- Όχι.

Αν όχι, γιατί δεν συμμετείχατε;
- 'Έλλειψη χρόνου.
- 'Ελλειψη ενδιαφέροντος.
- Άλλος λόγος.

Τους τελευταίους 12 μήνες, λάβατε μέρος εθελοντικά σε κάποια επίσημη δραστηριότητα/εκδήλωση;
- Ναι.
- Όχι.

Αν όχι, γιατί δεν συμμετείχατε;
- 'Έλλειψη χρόνου.
- 'Ελλειψη ενδιαφέροντος.
- Άλλος λόγος.

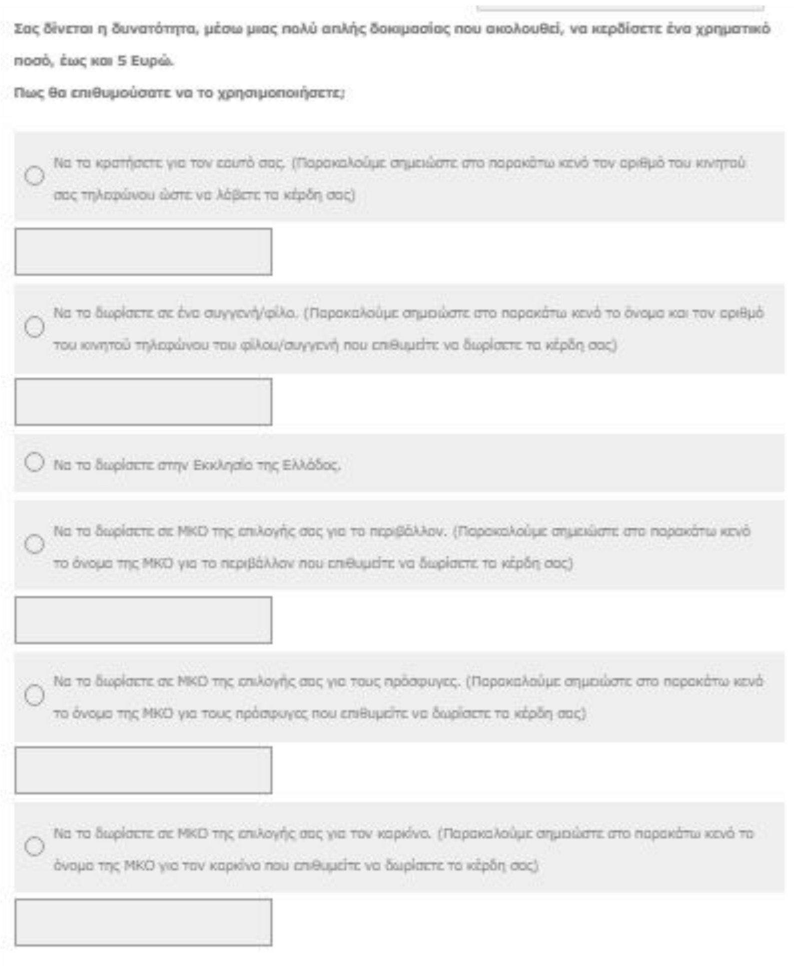
Τους τελευταίους 12 μήνες, λάβατε μέρος εθελοντικά σε κάποια πολιτική δραστηριότητα/εκδήλωση;
- Ναι.
- Όχι.

Αν όχι, γιατί δεν συμμετείχατε;
- 'Έλλειψη χρόνου.
- 'Ελλειψη ενδιαφέροντος.
- Άλλος λόγος.

Πόσο συχνά έχετε κάνει εθελοντική/κοινωφελή εργασία τους τελευταίους 12 μήνες;
- Σχεδόν κάθε μέρα.
- Σχεδόν κάθε βδομάδα.
- Σχεδόν κάθε μήνα.
- Πιο σπάνια από μια φορά τον μήνα.
- Ποτέ.

Σας δίνεται η δυνατότητα, μέσω μιας πολύ απλής δοκιμασίας που ακολουθεί, να κερδίσετε ένα χρηματικό ποσό, έως και 5 Ευρώ.

Πως θα επιθυμούσατε να το χρησιμοποιήσετε;
- Να τα κρατήσετε για τον εαυτό σας. (Παρακαλούμε σημειώστε στο παρακάτω κενό τον αριθμό του κινητού σας τηλεφώνου ώστε να λάβετε τα κέρδη σας)..
- Να τα δωρίσετε σε ένα συγγενή/φίλο. (Παρακαλούμε σημειώστε στο παρακάτω κενό το όνομα και τον αριθμό του κινητού τηλεφώνου του φίλου/συγγενή που επιθυμείτε να δωρίσετε τα κέρδη σας)
- Να τα δωρίσετε στην εκκλησία της Ελλάδας..
- Να τα δωρίσετε σε ΜΚΟ της επιλογής σας για το περιβάλλον. (Παρακαλούμε σημειώστε στο παρακάτω κενό το όνομα της ΜΚΟ για το περιβάλλον που επιθυμείτε να δωρίσετε τα κέρδη σας) .
- Να τα δωρίσετε σε ΜΚΟ της επιλογής σας για τους πρόσφυγες. (Παρακαλούμε σημειώστε στο παρακάτω κενό το όνομα της ΜΚΟ για τους πρόσφυγες που επιθυμείτε να δωρίσετε τα κέρδη σας).
- Να τα δωρίσετε σε ΜΚΟ της επιλογής σας για τον καρκίνο. (Παρακαλούμε σημειώστε στο παρακάτω κενό το όνομα της ΜΚΟ για τον καρκίνο που επιθυμείτε να δωρίσετε τα κέρδη σας).



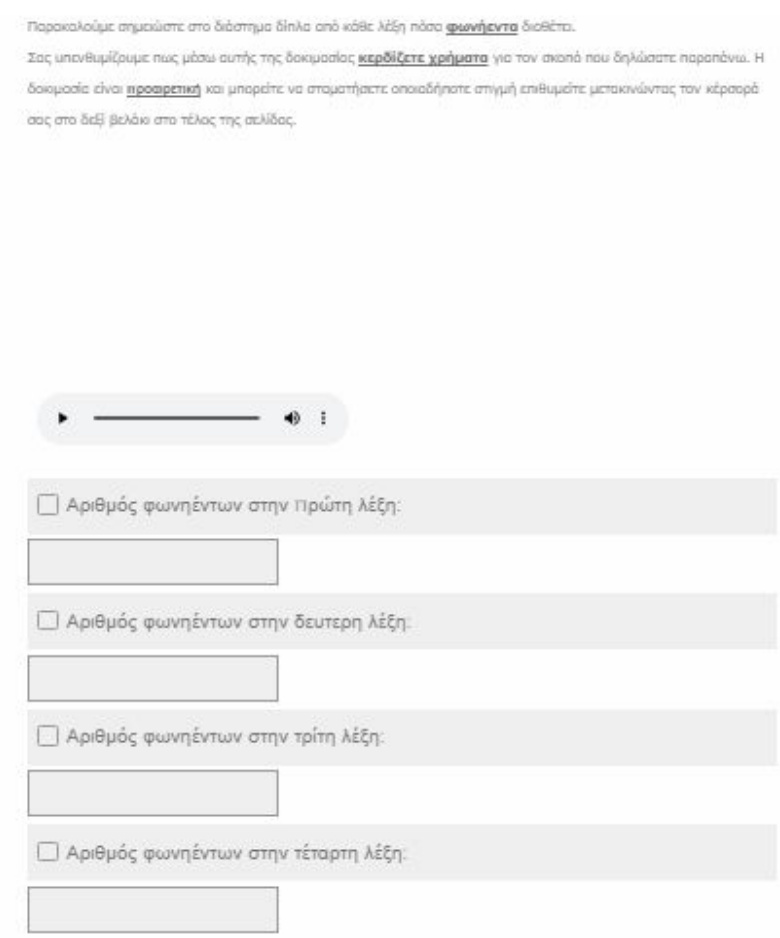
Σας δίνεται η δυνατότητα, μέσω μιας πολύ απλής δοκιμασίας που ακολουθεί, να κερδίσετε ένα χρηματικό ποσό, έως και 5 Ευρώ.

Πως θα επιθυμούσατε να το χρησιμοποιήσετε;

- Να το κρατήσετε για τον εαυτό σας. (Παρακαλούμε σημειώστε στο παρακάτω κενό τον αριθμό του κινητού σας τηλεφώνου ώστε να λάβετε τα κέρδη σας)
- Να το δωρίσετε σε ένα συγγενή/φίλο. (Παρακαλούμε σημειώστε στο παρακάτω κενό το όνομα και τον αριθμό του κινητού τηλεφώνου του φίλου/συγγενή που επιθυμείτε να δωρίσετε τα κέρδη σας)
- Να το δωρίσετε στην Εκκλησία της Ελλάδος.
- Να το δωρίσετε σε ΜΚΟ της επιλογής σας για το περιβάλλον. (Παρακαλούμε σημειώστε στο παρακάτω κενό το όνομα της ΜΚΟ για το περιβάλλον που επιθυμείτε να δωρίσετε τα κέρδη σας)
- Να το δωρίσετε σε ΜΚΟ της επιλογής σας για τους πρόσφυγες. (Παρακαλούμε σημειώστε στο παρακάτω κενό το όνομα της ΜΚΟ για τους πρόσφυγες που επιθυμείτε να δωρίσετε τα κέρδη σας)
- Να το δωρίσετε σε ΜΚΟ της επιλογής σας για τον καρκίνο. (Παρακαλούμε σημειώστε στο παρακάτω κενό το όνομα της ΜΚΟ για τον καρκίνο που επιθυμείτε να δωρίσετε τα κέρδη σας)

Στο ηχητικό αρχείο που ακολουθεί, θα ακούσετε 4 λέξεις.
Παρακαλούμε σημειώστε στο διάστημα δίπλα από κάθε λέξη πόσα φωνήεντα διαθέτει.
Σας υπενθυμίζουμε πως μέσω αυτής της δοκιμασίας κερδίζετε χρήματα για τον σκοπό που δηλώσατε παραπάνω. Η δοκιμασία είναι προαιρετική και μπορείτε να σταματήσετε οποιαδήποτε στιγμή επιθυμείτε μετακινώντας τον κέρσορά σας στο δεξί βελάκι στο τέλος της σελίδας.

- Αριθμός φωνηέντων στην πρώτη λέξη: ___________________.
- Αριθμός φωνηέντων στην δεύτερη λέξη: ___________________.
- Αριθμός φωνηέντων στην τρίτη λέξη: ___________________.
- Αριθμός φωνηέντων στην τέταρτη λέξη: ___________________.



Παρακαλούμε σημειώστε στο διάστημα δίπλα από κάθε λέξη πόσα **φωνήεντα** διαθέτει.

Σας υπενθυμίζουμε πως μέσω αυτής της δοκιμασίας **κερδίζετε χρήματα** για τον σκοπό που δηλώσατε παραπάνω. Η δοκιμασία είναι **προαιρετική** και μπορείτε να σταματήσετε οποιαδήποτε στιγμή επιθυμείτε μετακινώντας τον κέρσορά σας στο δεξί βελάκι στο τέλος της σελίδας.

- [ ] Αριθμός φωνηέντων στην Πρώτη λέξη:
- [ ] Αριθμός φωνηέντων στην δευτερη λέξη:
- [ ] Αριθμός φωνηέντων στην τρίτη λέξη:
- [ ] Αριθμός φωνηέντων στην τέταρτη λέξη:

Και τώρα ήρθε η ώρα να λάβετε μέρος στην λοταρία που σας δίνει την δυνατότητα να κερδίσετε το ποσό των 200 Ευρώ.
Σε περίπτωση που κερδίσετε, πως θα επιθυμούσατε να χρησιμοποιήσετε το ποσό αυτό;

- Να τα κρατήσετε για τον εαυτό σας. (Παρακαλούμε σημειώστε στο παρακάτω κενό τον αριθμό του κινητού σας τηλεφώνου ώστε να λάβετε τα κέρδη σας)
- Να τα δωρίσετε σε ένα συγγενή/φίλο. (Παρακαλούμε σημειώστε στο παρακάτω κενό το όνομα και τον αριθμό του κινητού τηλεφώνου του φίλου/συγγενή που επιθυμείτε να δωρίσετε τα κέρδη σας)
- Να τα δωρίσετε στην εκκλησία της Ελλάδας..
- Να τα δωρίσετε σε ΜΚΟ της επιλογής σας για το περιβάλλον. (Παρακαλούμε σημειώστε στο παρακάτω κενό το όνομα της ΜΚΟ για το περιβάλλον που επιθυμείτε να δωρίσετε τα κέρδη σας) .
- Να τα δωρίσετε σε ΜΚΟ της επιλογής σας για τους πρόσφυγες. (Παρακαλούμε σημειώστε στο παρακάτω κενό το όνομα της ΜΚΟ για τους πρόσφυγες που επιθυμείτε να δωρίσετε τα κέρδη σας).
- Να τα δωρίσετε σε ΜΚΟ της επιλογής σας για τον καρκίνο. (Παρακαλούμε σημειώστε στο παρακάτω κενό το όνομα της ΜΚΟ για τον καρκίνο που επιθυμείτε να δωρίσετε τα κέρδη σας).

Σε περίπτωση που κερδίσετε, πως θα επιθυμούσατε να χρησιμοποιήσετε το ποσό αυτό;

0 20 40 60 80 100 120 140 160 180 200

Να τα κρατήσετε για τον εαυτό σας. (Παρακαλούμε σημειώστε στο παρακάτω κενό τον αριθμό του κινητού σας τηλεφώνου ώστε να λάβετε τα κέρδη σας) 0

Να τα δωρίσετε σε ένα συγγενή/φίλο (Παρακαλούμε σημειώστε τον αριθμό του κινητού τηλεφώνου και το μικρό όνομα του ατόμου που επιθυμείτε να δωρίσετε τα κέρδη σας). 0

Να τα δωρίσετε στην Εκκλησία της Ελλάδος. 0

Να τα δωρίσετε σε ΜΚΟ της επιλογής σας για το περιβάλλον. (Παρακαλούμε σημειώστε στο παρακάτω κενό το όνομα της ΜΚΟ για το περιβάλλον που επιθυμείτε να δωρίσετε τα κέρδη σας) 0

Να τα δωρίσετε σε ΜΚΟ της επιλογής σας για τους πρόσφυγες.(Παρακαλούμε σημειώστε στο παρακάτω κενό το όνομα της ΜΚΟ για τους πρόσφυγες που επιθυμείτε να δωρίσετε τα κέρδη σας) 0

Να τα δωρίσετε σε ΜΚΟ της επιλογής σας για τον καρκίνο. (Παρακαλούμε σημειώστε στο παρακάτω κενό το όνομα της ΜΚΟ για τον καρκίνο που επιθυμείτε να δωρίσετε τα κέρδη σας) 0

Δεν επιθυμώ να λάβω μέρος στην λοταρία. 0

Σύνολο: 0

Και τώρα ήρθε η ώρα για την λοταρία!

Έχετε πάνω σας ένα χαρτονόμισμα?

Μην ανησυχείτε θα σας ρωτήσουμε μόνο τον σειριακό του αριθμό ώστε να εξασφαλίσουμε την διαφάνεια της διαδικασίας, αλλά και το οτι η διαδικασία στηρίζεται καθαρά στην τύχη!
Αφού βρείτε ένα χαρτονόμισμα, σημειώστε παρακάτω τα **ΤΡΊΑ ΠΡΏΤΑ** ψηφία του σειριακού αριθμού (στο παραπάνω χαρτονόμισμα θα ήταν 222).

Πατώντας επόμενο, θα δείτε έναν τυχαίο αριθμό από το 00 εως και το 99.

Αν ο αριθμός που θα δείτε είναι ίδιος με τα **ΔΥΟ ΤΕΛΕΥΤΑΙΑ** ψηφία στο χαρτονόμισμά σας (στο παραπάνω χαρτονόμισμα 64), έχετε κερδίσει!

Και ο αριθμός είναι...14!!

Παρακαλώ σημειώστε παρακάτω τα ΌΛΑ ΤΑ ΨΗΦΊΑ του σειριακού αριθμού του χαρτονομίσματος που έχετε στα χέρια σας.

Εαν τα δύο τελευταία είναι ίδια με τον παραπάνω αριθμό, συγχαρητήρια, έχετε κερδίσει!

Θα έρθουμε σε επαφή μαζί σας το συντομότερο για τα επόμενα βήματα!

ΣΗΜΑΝΤΙΚΌ: Μην χρησιμοποιήσετε το χαρτονόμισμα, καθώς θα χρειαστεί να το δείξετε ως απόδειξη!

Ευχαριστούμε θερμά που λάβατε μέρος στην έρευνά μας!

- Για οποιαδήποτε επιπλέον πληροφορία για αυτήν την έρευνα, παρακαλούμε μην διστάσετε να επικοινωνήσετε με τον υπεύθυνο ερευνητή:

Στοιχεία Επικοινωνίας Υπεύθυνου Ερευνητή:
[...]

Bibliography

Adda, Jérôme (2016). "Economic activity and the spread of viral diseases: Evidence from high frequency data". In: *The Quarterly Journal of Economics* 131.2, pp. 891–941.

Alvarez, Fernando E, David Argente, and Francesco Lippi (2020). *A simple planning problem for Covid-19 lockdown*. Tech. rep. National Bureau of Economic Research.

Anand, Paul, Jaya Krishnakumar, and Ngoc Bich Tran (2011). "Measuring welfare: Latent variable models for happiness and capabilities in the presence of unobservable heterogeneity". In: *Journal of public economics* 95.3-4, pp. 205–215.

Andersen, Steffen et al. (2008). "Lost in state space: are preferences stable?" In: *International Economic Review* 49.3, pp. 1091–1112.

Anderson, Lisa R and Jennifer M Mellor (2009). "Are risk preferences stable? Comparing an experimental measure with a validated survey-based measure". In: *Journal of Risk and Uncertainty* 39.2, pp. 137–160.

Arellano, Manuel and Stephen Bond (1991). "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations". In: *The review of economic studies* 58.2, pp. 277–297.

Atkeson, Andrew (2020). *How deadly is COVID-19? Understanding the difficulties with estimation of its fatality rate*. Tech. rep. National Bureau of Economic Research.

Baert, Stijn and Sunčica Vujić (2018). "Does it pay to care? Volunteering and employment opportunities". In: *Journal of Population Economics* 31.3, pp. 819–836.

Battistin, Erich et al. (2009). "The retirement consumption puzzle: evidence from a regression discontinuity approach". In: *American Economic Review* 99.5, pp. 2209–26.

Bauer, Daniel J and Patrick J Curran (2004). "The integration of continuous and discrete latent variable models: Potential problems and promising opportunities." In: *Psychological methods* 9.1, p. 3.

BBC (2020). *Covid: Mass testing in Liverpool sees 'remarkable decline' in cases.* Available at: https://www.bbc.com/news/uk-england-merseyside-55044488 Accessed 19 December 2020.

Bellemare, Charles, Sabine Kröger, and Arthur Van Soest (2008). "Measuring inequity aversion in a heterogeneous population using experimental decisions and subjective probabilities". In: *Econometrica* 76.4, pp. 815–839.

Bénabou, Roland and Jean Tirole (2006). "Incentives and prosocial behavior". In: *American economic review* 96.5, pp. 1652–1678.

Binder, Martin and Alex Coad (2011). "From Average Joe's happiness to Miserable Jane and Cheerful John: using quantile regressions to analyze the full subjective well-being distribution". In: *Journal of Economic Behavior & Organization* 79.3, pp. 275–290.

— (2013). "Life satisfaction and self-employment: a matching approach". In: *Small business economics* 40.4, pp. 1009–1033.

Biswas-Diener, Robert and ED Diener (2006). "The subjective well-being of the homeless, and lessons for happiness". In: *Social Indicators Research* 76.2, pp. 185–205.

Blanchflower, David G (2004). *Self-employment: More may not be better*. Tech. rep. National Bureau of Economic Research.

Bloomberg (2020). *Seoul's full cafes, apple store lines and show mass testing success.* Available at: https://www.bloomberg.com/news/articles/2020-04-18/seoul-s-full-cafes-apple-store-lines-show-mass-testing-success Accessed 19 December 2020.

Blundell, Richard and Stephen Bond (1998). "Initial conditions and moment restrictions in dynamic panel data models". In: *Journal of econometrics* 87.1, pp. 115–143.

Bond, Stephen R (2002). "Dynamic panel data models: a guide to micro data methods and practice". In: *Portuguese economic journal* 1.2, pp. 141–162.

Bowers, Alex J and Ryan Sprott (2012). "Why tenth graders fail to finish high school: A dropout typology latent class analysis". In: *Journal of Education for Students Placed at Risk (JESPAR)* 17.3, pp. 129–148.

Brañas-Garza, Pablo, Valerio Capraro, and Ericka Rascon-Ramirez (2018). "Gender differences in altruism on Mechanical Turk: Expectations and actual behaviour". In: *Economics Letters* 170, pp. 19–23.

Brauner, Jan M. et al. (2020). "Inferring the effectiveness of government interventions against COVID-19". In: *Science*. ISSN: 0036-8075. DOI: 10.1126/science.abd9338. eprint: https://science.sciencemag.org/content/early/2020/12/15/science.abd9338.full.pdf. URL: https://science.sciencemag.org/content/early/2020/12/15/science.abd9338.

Brief, Arthur P and Stephan J Motowidlo (1986). "Prosocial organizational behaviors". In: *Academy of management Review* 11.4, pp. 710–725.

Brown, Sarah et al. (2014). "Modelling financial satisfaction across life stages: A latent class approach". In: *Journal of Economic Psychology* 45, pp. 117–127.

Cadegiani, Flavio A et al. (2021). "Clinical diagnosis of COVID-19: A prompt, feasible, and sensitive diagnostic tool for COVID-19 based on a 1,757-patient cohort (The AndroCoV Clinical Scoring for COVID-19 diagnosis)." In: *medRxiv*, pp. 2020–12.

Carpenter, Jeffrey, Cristina Connolly, and Caitlin Knowles Myers (2008). "Altruistic behavior in a representative dictator experiment". In: *Experimental Economics* 11.3, pp. 282–298.

Carstensen, Laura L and Susan Turk Charles (1998). "Emotion in the second half of life". In: *Current Directions in Psychological Science* 7.5, pp. 144–149.

Celeux, Gilles and Gilda Soromenho (1996). "An entropy criterion for assessing the number of clusters in a mixture model". In: *Journal of classification* 13.2, pp. 195–212.

Charness, Gary and Marie-Claire Villeval (2009). "Cooperation and competition in intergenerational experiments in the field and the laboratory". In: *American Economic Review* 99.3, pp. 956–78.

Chen, Yan and Sherry Xin Li (2009). "Group identity and social preferences". In: *American Economic Review* 99.1, pp. 431–57.

Clark, Andrew et al. (2005). "Heterogeneity in reported well-being: evidence from twelve European countries". In: *The Economic Journal* 115.502, pp. C118–C132.

Coe, Norma B and Gema Zamarro (2011). "Retirement effects on health in Europe". In: *Journal of health economics* 30.1, pp. 77–86.

Collins, Linda M and Stephanie T Lanza (2009). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. Vol. 718. John Wiley & Sons.

Cooper, David J and John H Kagel (2016). "Other-regarding preferences". In: *The handbook of experimental economics* 2, p. 217.

Cozzi, Guido, Noemi Mantovan, and Robert M Sauer (2017). "Does it pay to work for free? Negative selection and the wage returns to volunteer experience". In: *Oxford Bulletin of Economics and Statistics* 79.6, pp. 1018–1045.

Dattner, Itai et al. (2020). "The role of children in the spread of COVID-19: Using household data from Bnei Brak, Israel, to estimate the relative susceptibility and infectivity of children". In: *medRxiv*.

Dawson, Chris, Michail Veliziotis, and Benjamin Hopkins (2017). "Temporary employment, job satisfaction and subjective well-being". In: *Economic and Industrial Democracy* 38.1, pp. 69–98.

DellaVigna, Stefano, John A List, and Ulrike Malmendier (2012). "Testing for altruism and social pressure in charitable giving". In: *The quarterly journal of economics* 127.1, pp. 1–56.

Dempster, Arthur P, Nan M Laird, and Donald B Rubin (1977). "Maximum likelihood from incomplete data via the EM algorithm". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 39.1, pp. 1–22.

Dergiades, Theologos, Costas Milas, and Theodore Panagiotidis (2020). "Effectiveness of government policies in response to the COVID-19 outbreak". In: *Available at SSRN 3602004*.

Diener, Edward, Richard E Lucas, Shigehiro Oishi, et al. (2002). "Subjective well-being: The science of happiness and life satisfaction". In: *Handbook of positive psychology* 2, pp. 63–73.

Dohmen, Thomas et al. (2008). "Representative trust and reciprocity: Prevalence and determinants". In: *Economic Inquiry* 46.1, pp. 84–90.

Dolan, Paul and Laura Kudrna (2016). "Sentimental hedonism: pleasure, purpose, and public policy". In: *Handbook of eudaimonic well-being*. Springer, pp. 437–452.

Dolan, Paul and Robert Metcalfe (2012). "Measuring subjective wellbeing: Recommendations on measures for use by national governments". In: *Journal of social policy* 41.2, pp. 409–427.

Dolan, Paul and Aki Tsuchiya (2011). "Determining the parameters in a social welfare function using stated preference data: an application to health". In: *Applied Economics* 43.18, pp. 2241–2250.

Dush, Claire M Kamp and Paul R Amato (2005). "Consequences of relationship status and quality for subjective well-being". In: *Journal of Social and Personal Relationships* 22.5, pp. 607–627.

ECDC (2021). *Data on testing for Covid-19 by week and country.* Tech. rep. Available at: https://www.ecdc.europa.eu/en/publications-data/covid-19-testing.

Elimian, Kelly Osezele et al. (2020). "Patient characteristics associated with COVID-19 positivity and fatality in Nigeria: Retrospective cohort study". In: *BMJ open* 10.12, e044079.

Erlinghagen, Marcel (2010). "Volunteering after retirement: evidence from German panel data". In: *European Societies* 12.5, pp. 603–625.

Fehr, Ernst and Klaus M Schmidt (1999). "A theory of fairness, competition, and cooperation". In: *The quarterly journal of economics* 114.3, pp. 817–868.

Fernandez-Blanco, Victor, Luis Orea, and Juan Prieto-Rodriguez (2009). "Analyzing consumers heterogeneity and self-reported tastes: An approach consistent with the consumer?s decision making process". In: *Journal of Economic Psychology* 30.4, pp. 622–633.

Fitzpatrick, Maria D and Timothy J Moore (2018). "The mortality effects of retirement: Evidence from Social Security eligibility at age 62". In: *Journal of Public Economics* 157, pp. 121–137.

Gelman, Andrew and Guido Imbens (2019). “Why high-order polynomials should not be used in regression discontinuity designs”. In: *Journal of Business & Economic Statistics* 37.3, pp. 447–456.

Georganas, Sotiris, Mirco Tonin, and Michael Vlassopoulos (2015). “Peer pressure and productivity: The role of observing and being observed”. In: *Journal of Economic Behavior & Organization* 117, pp. 223–232.

Gilligan, Michael J, Benjamin J Pasquale, and Cyrus Samii (2014). “Civil war and social cohesion: Lab-in-the-field evidence from Nepal”. In: *American Journal of Political Science* 58.3, pp. 604–619.

Giovanis, Eleftherios (2014). “Relationship between well-being and recycling rates: evidence from life satisfaction approach in Britain”. In: *Journal of Environmental Economics and Policy* 3.2, pp. 201–214.

Golman, Russell, David Hagmann, and George Loewenstein (2017). “Information avoidance”. In: *Journal of Economic Literature* 55.1, pp. 96–135.

Green, Leonard and Joel Myerson (2004). “A discounting framework for choice with delayed and probabilistic rewards.” In: *Psychological bulletin* 130.5, p. 769.

Greene, Sharon K et al. (2021). “Nowcasting for Real-Time COVID-19 Tracking in New York City: An Evaluation Using Reportable Disease Data From Early in the Pandemic”. In: *JMIR Public Health and Surveillance* 7.1, e25538.

Hagenaars, Jacques A and Allan L McCutcheon (2002). *Applied latent class analysis*. Cambridge University Press.

Haug, Nils et al. (2020). “Ranking the effectiveness of worldwide COVID-19 government interventions”. In: *Nature Human behaviour* 4.12, pp. 1303–1312.

Heffetz, Ori and Matthew Rabin (2013). “Conclusions regarding cross-group differences in happiness depend on difficulty of reaching respondents”. In: *American Economic Review* 103.7, pp. 3001–21.

Hetschko, Clemens and Malte Preuss (2020). “Income in jeopardy: How losing employment affects the willingness to take risks”. In: *Journal of Economic Psychology* 79, p. 102175.

Holm, Håkan and Paul Nystedt (2005). “Intra-generational trust—a semi-experimental study of trust among different generations”. In: *Journal of Economic Behavior & Organization* 58.3, pp. 403–419.

Hsiang, Solomon et al. (2020). “The effect of large-scale anti-contagion policies on the COVID-19 pandemic”. In: *Nature* 584.7820, pp. 262–267.

Huang, Guan-Hua and Karen Bandeen-Roche (2004). “Building an identifiable latent class model with covariate effects on underlying and measured variables”. In: *Psychometrika* 69.1, pp. 5–32.

Kachelmeier, Steven J and Mohamed Shehata (1992). “Examining risk preferences under high monetary incentives: Experimental evidence from the People’s Republic of China”. In: *The American Economic Review*, pp. 1120–1141.

Katz, Eliakim and Jacob Rosenberg (2005). “An economic interpretation of institutional volunteering”. In: *European Journal of Political Economy* 21.2, pp. 429–443.

Kelvin, Alyson A and Scott Halperin (2020). “COVID-19 in children: the link in the transmission chain”. In: *The Lancet Infectious Diseases* 20.6, pp. 633–634.

Kettner, Sara Elisa and Israel Waichman (2016). “Old age and prosocial behavior: Social preferences or experimental confounds?” In: *Journal of Economic Psychology* 53, pp. 118–130.

Knabe, Andreas et al. (2010). “Dissatisfied with life but having a good day: time-use and well-being of the unemployed”. In: *The Economic Journal* 120.547, pp. 867–889.

Kristoffersen, Ingebjørg (2018). "Great expectations: Education and subjective well-being". In: *Journal of Economic Psychology* 66, pp. 64–78.
Laffan, Kate (2018). "Every breath you take, every move you make: Visits to the outdoors and physical activity help to explain the relationship between air pollution and subjective wellbeing". In: *Ecological Economics* 147, pp. 96–113.
Lamu, Admassu N and Jan Abel Olsen (2016). "The relative importance of health, income and social relations for subjective well-being: An integrative analysis". In: *Social Science & Medicine* 152, pp. 176–185.
Lang, Natasja DJ van et al. (2006). "Latent class analysis of anxiety and depressive symptoms of the Youth Self-Report in a general population sample of young adolescents". In: *Behaviour research and therapy* 44.6, pp. 849–860.
Larney, Andrea, Amanda Rotella, and Pat Barclay (2019). "Stake size effects in ultimatum game and dictator game offers: A meta-analysis". In: *Organizational Behavior and Human Decision Processes* 151, pp. 61–72.
Lee, David S and Thomas Lemieux (2010). "Regression discontinuity designs in economics". In: *Journal of economic literature* 48.2, pp. 281–355.
Lenzenweger, Mark F (2004). "Consideration of the challenges, complications, and pitfalls of taxometric analysis." In: *Journal of Abnormal Psychology* 113.1, p. 10.
Lin, Ting Hsiang and C Mitchell Dayton (1997). "Model selection information criteria for non-nested latent class models". In: *Journal of Educational and Behavioral Statistics* 22.3, pp. 249–264.
Linzer, Drew A, Jeffrey B Lewis, et al. (2011). "poLCA: An R package for polytomous variable latent class analysis". In: *Journal of statistical software* 42.10, pp. 1–29.
List, John A (2004). "Young, selfish and male: Field evidence of social preferences". In: *The Economic Journal* 114.492, pp. 121–149.
Macintyre, Sally et al. (1998). "Do housing tenure and car access predict health because they are simply markers of income or self esteem? A Scottish study." In: *Journal of Epidemiology & Community Health* 52.10, pp. 657–664.
Mahase, Elisabeth (2020). "Covid-19: Mass testing in Slovakia may have helped cut infections". In: *BMJ* 371:m4761. DOI: `https://doi.org/10.1136/bmj.m4761`.
Manski, Charles F (1993). "Identification of endogenous social effects: The reflection problem". In: *The review of economic studies* 60.3, pp. 531–542.
Manski, Charles F and Francesca Molinari (2021). "Estimating the COVID-19 infection rate: Anatomy of an inference problem". In: *Journal of Econometrics* 220.1, pp. 181–192.
Masyn, Katherine E (2013). "25 latent class analysis and finite mixture modeling". In: *The Oxford handbook of quantitative methods*, p. 551.
McLachlan, Geoffrey J and Thriyambakam Krishnan (2007). *The EM algorithm and extensions*. Vol. 382. John Wiley & Sons.
Menchik, Paul L and Burton A Weisbrod (1987). "Volunteer labor supply". In: *Journal of Public Economics* 32.2, pp. 159–183.
Middelburg, Rutger A and Frits R Rosendaal (2020). "COVID-19: How to make between-country comparisons". In: *International Journal of Infectious Diseases* 96, pp. 477–481.
Mina, Michael J., Roy Parker, and Daniel B. Larremore (2020). "Rethinking Covid-19 Test Sensitivity — A Strategy for Containment". In: *New England Journal of Medicine* 383.22, e120. DOI: `10.1056/NEJMp2025631`. eprint: `https://doi.org/10.1056/NEJMp2025631`. URL: `https://doi.org/10.1056/NEJMp2025631`.
Moreau, Nicolas and Elena Stancanelli (2015). "Household consumption at retirement: A regression discontinuity study on French data". In: *Annals of Economics and Statistics/Annales d'Économie et de Statistique* 117/118, pp. 253–276.

Müller, Tobias and Mujaheed Shaikh (2018). "Your retirement and my health behavior: Evidence on retirement externalities from a fuzzy regression discontinuity design". In: *Journal of health economics* 57, pp. 45–59.
Mutchler, Jan E, Jeffrey A Burr, and Francis G Caro (2003). "From paid worker to volunteer: Leaving the paid workforce and volunteering in later life". In: *Social forces* 81.4, pp. 1267–1293.
Nagin, Daniel S (1999). "Analyzing developmental trajectories: a semiparametric, group-based approach." In: *Psychological methods* 4.2, p. 139.
Nickell, Stephen (1981). "Biases in dynamic models with fixed effects". In: *Econometrica: Journal of the econometric society*, pp. 1417–1426.
Norwich, AMf and IB Turksen (1984). "A model for the measurement of membership and the consequences of its empirical implementation". In: *Fuzzy sets and systems* 12.1, pp. 1–25.
OECD (2017). *Pensions at a Glance 2017: OECD and G20 indicators*. Tech. rep. OECD Publishing, Paris.
Oster, Emily, Ira Shoulson, and E Dorsey (2013). "Optimal expectations and limited medical testing: Evidence from Huntington disease". In: *American Economic Review* 103.2, pp. 804–30.
Oswald, Frank et al. (2003). "Housing and life satisfaction of older adults in two rural regions in Germany". In: *Research on aging* 25.2, pp. 122–143.
Patzelt, Holger and Dean A Shepherd (2011). "Negative emotions of an entrepreneurial career: Self-employment and regulatory coping behaviors". In: *Journal of Business venturing* 26.2, pp. 226–238.
Pavelka, Martin et al. (2020). "The effectiveness of population-wide, rapid antigen test based screening in reducing SARS-CoV-2 infection prevalence in Slovakia". In: *medRxiv*.
Prelec, Drazen and George Loewenstein (1991). "Decision making over time and under uncertainty: A common approach". In: *Management science* 37.7, pp. 770–786.
Qiu, Haiyan et al. (2020). "Clinical and epidemiological features of 36 children with coronavirus disease 2019 (COVID-19) in Zhejiang, China: an observational cohort study". In: *The Lancet Infectious Diseases* 20.6, pp. 689–696.
Rabin, Matthew (1993). "Incorporating fairness into game theory and economics". In: *The American economic review*, pp. 1281–1302.
REACT (2020). *REACT-1: Summary of sample extraction and fieldwork dates, and response rates, by round*. Tech. rep. Retreived on 2021-01-30. URL: https://www.imperial.ac.uk/media/imperial-college/institute-of-global-health-innovation/REACT1-Fieldwork-info-table-for-Imperial-website_with-Round-7.pdf.
Riley, Steven et al. (2020). "Community prevalence of SARS-CoV-2 virus in England during May 2020: REACT study". In: *medRxiv*.
Roodman, David (2009). "How to do xtabond2: An introduction to difference and system GMM in Stata". In: *The stata journal* 9.1, pp. 86–136.
Rumpf, Hans-Jürgen et al. (2014). "Occurence of internet addiction in a general population sample: a latent class analysis". In: *European addiction research* 20.4, pp. 159–166.
Sauer, Robert M (2015). "Does it pay for women to volunteer?" In: *International Economic Review* 56.2, pp. 537–564.
Schildberg-Hörisch, Hannah (2018). "Are risk preferences stable?" In: *Journal of Economic Perspectives* 32.2, pp. 135–54.
Selten, Richard (1975). "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games". In: *International Journal of Game Theory* 4, pp. 25–55.

Shelter (2018). *A Vision for Social Housing.* Tech. rep. Shelter UK. URL: https://england.shelter.org.uk/support_us/campaigns/a_vision_for_social_housing.

Shen, Chen et al. (2021). "Unraveling the flaws of estimates of the infection fatality rate for COVID-19". In: *Journal of Travel Medicine*.

Sherman, Arie and Tal Shavit (2012). "How the lifecycle hypothesis explains volunteering during retirement". In: *Ageing and Society* 32.8, p. 1360.

Stancanelli, Elena and Arthur Van Soest (2012). "Retirement and home production: A regression discontinuity approach". In: *American Economic Review* 102.3, pp. 600–605.

Steptoe, Andrew, Angus Deaton, and Arthur A Stone (2015). "Subjective wellbeing, health, and ageing". In: *The Lancet* 385.9968, pp. 640–648.

Stiglitz, Joseph, Amartya Sen, Jean-Paul Fitoussi, et al. (2009). "The measurement of economic performance and social progress revisited". In: *Reflections and overview. Commission on the measurement of economic performance and social progress, Paris*.

Stone, Arthur A and Christopher Ed Mackie (2013). *Subjective well-being: Measuring happiness, suffering, and other dimensions of experience.* National Academies Press.

Sugden, Robert (1993). "Welfare, resources, and capabilities: a review of inequality reexamined by Amartya Sen". In: *Journal of Economic literature* 31.4, pp. 1947–1962.

Sutter, Matthias and Martin G Kocher (2007). "Trust and trustworthiness across different age groups". In: *Games and Economic behavior* 59.2, pp. 364–382.

Sze, Jocelyn A et al. (2012). "Greater emotional empathy and prosocial behavior in late life." In: *Emotion* 12.5, p. 1129.

Thornton, Rebecca L. (2008). "The Demand for, and Impact of, Learning HIV Status". In: *American Economic Review* 98.5, pp. 1829–63. DOI: 10.1257/aer.98.5.1829. URL: https://www.aeaweb.org/articles?id=10.1257/aer.98.5.1829.

Van Goethem, Anne AJ et al. (2014). "Socialising adolescent volunteering: How important are parents and friends? Age dependent effects of parents and friends on adolescents' volunteering behaviours". In: *Journal of Applied Developmental Psychology* 35.2, pp. 94–101.

Van Koten, Silvester, Andreas Ortmann, and Vitezslav Babicky (2013). "Fairness in risky environments: Theory and evidence". In: *Games* 4.2, pp. 208–242.

Vermunt, Jeroen K and Jay Magidson (2003). "Latent class models for classification". In: *Computational Statistics & Data Analysis* 41.3-4, pp. 531–537.

Wang, Jian et al. (2020). "Are patient-regarding preferences stable? Evidence from a laboratory experiment with physicians and medical students from different countries". In: *European Economic Review*, p. 103411.

Weber, Bethany J and Gretchen B Chapman (2005). "Playing for peanuts: Why is risk seeking more common for low-stakes gambles?" In: *Organizational Behavior and Human Decision Processes* 97.1, pp. 31–46.

Yamashita, Takashi et al. (2019). "Underlying motivations of volunteering across life stages: A study of volunteers in nonprofit organizations in Nevada". In: *Journal of Applied Gerontology* 38.2, pp. 207–231.

Zumbro, Timo (2014). "The relationship between homeownership and life satisfaction in Germany". In: *Housing Studies* 29.3, pp. 319–338.