COVID-19 vaccination intentions and subsequent uptake:

An analysis of the role of marginalisation in society using British longitudinal data

By Silvia Mendolia* and Ian Walker**

Abstract

COVID-19 vaccine hesitancy has previously been modelled using data on intentions – expressed prior to vaccine availability. Once vaccines became widely available, it became possible to model hesitancy using actual vaccination uptake data. This paper estimates the determinants of the joint distribution of COVID-19 vaccination intentions (declared before the release of any vaccine) and actual vaccination take-up (when it was widely available across the age distribution). We use high quality longitudinal data (UK Household Longitudinal Study) collected during the pandemic in the UK, merged to a wide variety of individual characteristics collected prior to the COVID-19 pandemic. Our estimation draws on pre-Covid values of variables for a sample that includes 10,073 observations from the September 2021 wave.

The contribution of this paper is to model hesitancy and uptake *jointly*. The work shows that people who might be regarded as marginalised in society (measured, before the pandemic began) are less likely to say that they *intend* to be vaccinated and they go on to also be more likely to *actually* remain unvaccinated. It also shows that there is a large positive correlation between the unobservable determinants of intention and of uptake. This high positive correlation has an important implication - that information campaigns can be reasonably well profiled to target specific groups on the basis of *intention* data alone. We also show that changing one's mind is not correlated with observable data. This is consistent with two explanations. Firstly, the new information available on the arrival of vaccines, that they are safe and effective, may be more optimistic than was originally assumed. Secondly, individuals may have been more pessimistic about the effects associated with infection before vaccines became available.

Keywords: vaccination; COVID-19; marginalisation

*Silvia Mendolia, University of Turin: silvia.mendolia@unito.it (corresponding author)

**Ian Walker, Lancaster University

1. Introduction

Vaccine hesitancy is defined as "delay in acceptance or refusal of vaccines despite availability of vaccine services" (SAGE Vaccine Hesitancy Working Group, 2015). It is a complex phenomenon, which varies across times, vaccines, and diseases. It is one of the major threats to achieving effective immunity when herd immunity requires a particularly high of immunisation rate. Much of the existing literature has focussed on intentions to get the vaccine, rather than actual vaccination – unsurprisingly so, because of the difficulty of untangling the supply of vaccines from the demand for them when vaccines are in short supply, or even not yet available at all. We use data that tracks individuals over time, up to a point where supply constraints are no longer present.

This allows us to make two principal contributions. First and foremost, we are able to decompose hesitancy into: the intention to remain unvaccinated, even when vaccines become available; and the refusal to vaccinate when they do. We exploit this data to jointly model intentions **not** to vaccinate and the **failure to** vaccinate - using an empirical specification that allows for the unobservable determinants of each response to be correlated . We find that people who, for idiosyncratic reasons, have weak intentions are prone to also be less likely, than their observable characteristics suggest, to vaccinate when it is possible to do so. Our estimation method allows for intentions and actual choices to differ. Thus, we place no constraints on the two dependent variables: we allow individuals to change their minds over time – in either direction. The covariance between the unobservables in the two equations is high (over 0.75).

Our second contribution comes from the richness and nature of our data. While previous work has considered a patchwork of covariates, we have a particularly comprehensive dataset that effectively embraces all existing work. Since these covariates might be correlated with each other, having data that allows us to control for such a wide range of factors is likely to mitigate the risk of parameter instability associated with omitted confounders. Moreover, our data comes from a long-running representative panel study which provided a sampling frame for further questions specific to COVID-19 to be asked throughout the pandemic, until the point where safe and effective vaccines had become widely available. Individuals were interviewed monthly, or even bi-monthly, to provide observations during the COVID-19 pandemic. Importantly, unlike other studies, the variables that shape the decisions that we model are predetermined – that is, they are recorded prior to the advent of COVID-19. This is important because it implies that the coefficients on the explanatory variables do not suffer from endogeneity bias arising from the advent of COVID-19 itself. The advantage of our approach

is that we can decompose decision-making and we can use the resulting parameters to provide estimates of the marginal effects of each characteristic on the probabilities of being in each cell of our data - defined by likelihood to not intend to vaccinate **and** of ultimately being un-vaccinated. Moreover, we allow for the possibility of switching to being pro-vaccine once one is available.

We cannot control for selection on unobservables - it is possible that there are some *unobservable* characteristics that influence vaccine hesitancy despite the richness of our data. However, we use recently developed tests (Oster, 2019) to show that it would take large, likely implausible, levels of selection bias to drive our results to zero. While caution is needed in interpreting our findings as causal, one of our most important contributions is to support policymakers in targeting individuals who, on the basis of their *observable* traits, and causality is of second order importance.

The results suggest that vaccine intentions and ultimate uptake are well aligned. The raw data shows that a large majority of those who thought that they were likely, or very likely, to vaccinate, expressed prior to any vaccine being available, did actually get vaccinated in the subsequent months when the vaccine had become widely available. This is not only due to the strong positive correlation between the unobservables in each equation – we also find that the coefficients on the explanatory variables in each equation are remarkably similar. Indeed, only one or two explanatory variables appear to be statistically significant in their marginal effects associated with *changing one's mind*, either way.

We find well-determined effects of individual characteristics, consistent with other studies. In particular, people who are young, or with low levels of political and civic participation, or who lack trust in the democratic institutions, are from ethnic minorities, or from the bottom quartile of the income distribution, put themselves at significantly higher risks of infection. Our work shows that these people, to the extent that they are marginalised in society (before the pandemic), are more likely to be hesitant about the vaccine and so represent a higher risk of to themselves (and others) - with effect sizes, relative to the mean for the omitted category, that are large compared to the mean probability.

We find a high degree of consistency between intentions and subsequent actions. After vaccines became available in the UK, and became known to be effective and safe, we would expect actual take-up to increase over time relative to uptake intentions. Indeed, in our data, only around 35% of hesitant individuals persisted in their intention to remain unvaccinated.

Our results reflect the sign patterns from previous studies. Our finding on the coefficients on intentions closely match those in earlier work and our estimates of actual uptake, provides reassurance about the relevance of earlier work. Our finding that there are no significant determinants of *switching*, from not intending to vaccinate to taking-up a vaccine when available, suggests that initial reservations based on concerns about ineffectiveness and fears of side-effects were unduly pessimistic once vaccines became available. The high correlation between unobservables suggests that attitudes are otherwise quite persistent. All of this points to the legitimacy of profiling take-up based on intentions - rather than waiting for actual decisions being made when reliable supply becomes available – so allowing getting jabs into as many arms as possible, as fast as possible, to reduce mortality rates.

Section 2 provides a brief overview of existing work that we have used to drive our choice of explanatory variables. Section 3 describes our data. Section 4 explains the methodology. Section 5 explains the results, while Section 6 provides discussion around them.

2. Literature

Individuals who are sceptical about COVID-19 vaccinations often have various concerns: about side effects, limitations of personal freedom of choice, and beliefs that COVID-19 is not a serious threat to the young and healthy. Furthermore, the rollout of the vaccine was organised by sequential risk classes until it was available for all. Delaying commitment allows individuals to retain option value given the uncertainty about the effectiveness and safety of vaccine(s), given that vaccination is irreversible. Such a second-mover advantage motive has also been reflected in the literature (for example, Kennedy et al. 2021 and Mouter et al., 2022).

The World Health Organization (WHO) has stressed the importance of ensuring that communities have access to reliable public health information and has warned against the so-called "infodemic" – an overload of information, especially inaccurate messaging, about various aspects of the pandemic, including vaccine efficacy and safety (WHO, 2020). Recent studies have shown the importance of elaborating and framing messages with specific information about the vaccine, to emphasize safety and so reduce hesitancy (Diament et al., 2022). Ashworth et al (2021) show that different types of public messages can differentially increase vaccine intentions, but the strongest effects come from those that focus on the personal health benefits from COVID-19 vaccination. Such an emphasis is normally a reflection of free-rider considerations - but in the case of COVID-19 it seems that none of the available vaccines

have shown a strong record on *preventing* infection, and the major effect appears to be largely a *personal* benefit associated with improved outcomes conditional on infection.

The literature on attitudes towards COVID-19 vaccination has considered many factors which may affect hesitancy, including socio-demographic characteristics, perceived vaccine efficacy and safety, compliance with wider protective behaviours, and perceptions of the severity of COVID-19 symptoms (see for example Caserotti et al., 2021; Dabla-Norris, 2021; Edwards et al, 2020; among others). However, the literature has focussed on intentions to get the vaccine, rather than actual vaccination – unsurprisingly so, because of the impossibility of untangling the supply of vaccines from the demand for them when supply is currently rationed, but is expected to become available soon. Cascini et al (2021) review the existing literature on COVID-19 vaccination hesitancy and show differences in hesitancy across countries, but some sub-groups of population seem to be consistently more hesitant. These are, for example: women, people from low-income groups, those living in rural areas, and those belonging to an ethnic minority. Furthermore, younger individuals tend to be more hesitant, as do individuals with low levels of education (Galasso et al., 2021; Paul et al., 2021; Robertson et al, 2021, among others).

Recent literature has analysed the observable traits of vaccine hesitant individuals in greater detail, including socio-economic status, political preferences, perception of authority and cultural values (Ward et al., 2020; Chauduri et al., 2022; Fortunato and Lombini, 2022). Similarly, Murphy et al (2021) show that vaccine hesitant individuals in UK and Ireland exhibit a series of common attitudes including: being more likely to have anti-immigrant views, being less likely to acquire information about the pandemic from traditional sources, show low levels of trust in science, and profess higher levels of self-interest. A related topic is the relationship between important economic preferences and social responsibility and behaviours during the COVID-19 crisis. For example, Muller and Rau (2021) analyse how risk tolerance, time preferences, trust, and honesty predict compliance with restrictive measures during the pandemic.

Finally, there is similar, but less extensive, descriptive information on vaccination (both intentions to vaccinate and actual vaccination) in ONS (2021). This uses a Covid-19 module added to the regular Opinions and Lifestyle Surveys (OLS) run by the ONS. The OLS is the general-purpose survey that was used by government for collecting information on urgent issues, prior to COVID-19. In March 2020 it was rededicated to COVID-19, expanded to a sample of around 5,000 individuals with an average response rate of close to 70%, and it

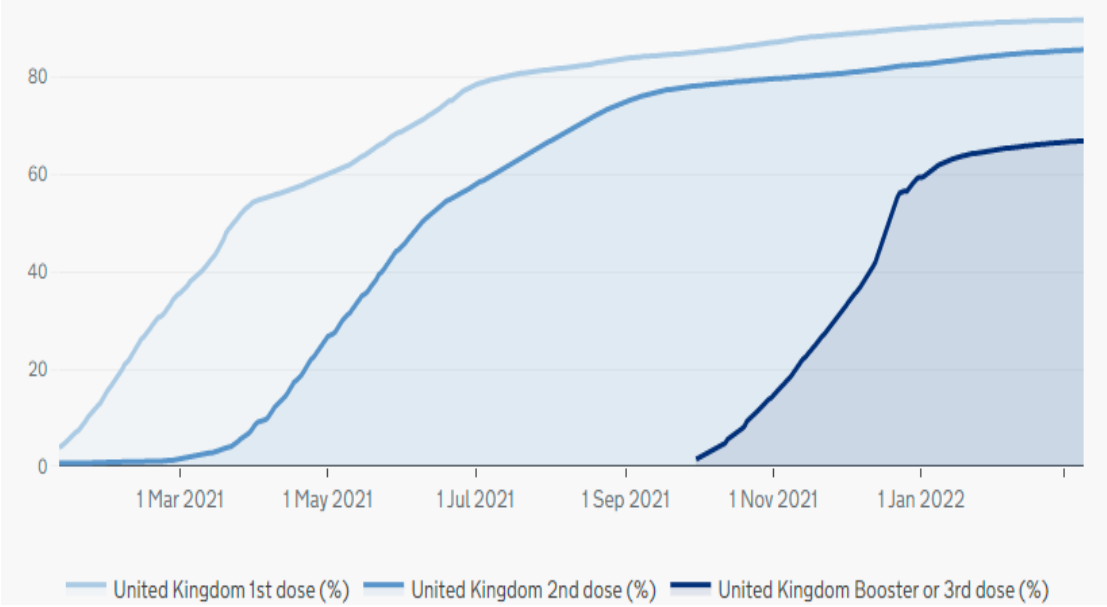
became weekly until August 2021. ONS (2021) provides descriptions of vaccination non-takeup (not taking up an offer of a vaccine) and intentions (by January 2020). The OLS data consists of pooled cross sections of data rather than a panel so, unlike here, it cannot model the joint distribution of the unobservable determinants of intentions to vaccinate and actual vaccination. Moreover, the data on some variables are contemporaneous (for example, employment status) and these variables may then have been affected, at least in part, by COVID-19. Nonetheless, the ONS study presents tabulations that resonate with results here.

3. Context, Data and Descriptive Statistics

The UK became the first country in Europe to grant emergency use authorisation for a COVID-19 vaccine when the Medicines and Healthcare products Regulatory Authority gave approval for use of the Pfizer-BioNTech vaccine in adults on 2 December 2020. The AstraZeneca vaccine was subsequently approved for use in adults on 30 December 2020. The UK National Health Service (NHS) started administering vaccinations in England on 8th December 2020. The first phase of the vaccination campaign aimed at offering a vaccination to the elderly and individuals with underlying health conditions. Apart from this latter priority, the rollout was rigidly age-related. The next phase of the vaccine rollout started on 13th April 2021 and by September 2021 all individuals aged 18 and over were able to book a vaccination (NHS, 2021) and, in our data, 96% reported that they had received a vaccination. Figure 1 shows the cumulative vaccination rate, from official administrative statistics, for the 1st and subsequent doses in the UK. This phase of the vaccine rollout program initially aimed at offering the first dose to all individuals older than 18 years old by end July 2021 (NHS, 2021).

However, the NHS established a new target of offering all adults a first dose, and two thirds of adults their second dose, by July 2021. This target was met on 18th July. In addition, a decision was made to accelerate the rollout of second doses, by reducing the interval between doses from 12 weeks to 8 weeks, so that all individuals older than 40 years old who received a first dose by mid-May were offered their second dose by 19th July. Vaccinations were free to everyone and were made available through many hospitals, most community health practices (GPs), and in many places in the community – large and small, such as individual pharmacies, sports centres, and major convention centres.

Figure 1 UK vaccine rollout (%)



Source: ONS, 2022

3.1 Understanding Society (the UK Household Longitudinal Study)

We use data from the UK Household Longitudinal Study (UKHLS), also known as *Understanding Society* (University of Essex, 2022). UKHLS surveyed approximately 40,000 households living in the United Kingdom in Wave 1. Data collection started in 2009–2010 for Wave 1, and eleven waves are currently available.

The University of Essex Ethics Committee has approved all data collection on Understanding Society main study and innovation panel waves, including asking consent for all data linkages (except to health records). Understanding Society data (including biological samples) is available to all bona fide researchers (whether in universities, government departments, charities or commercial companies), for research that can demonstrate public benefit and fits within the Study's ethical framework (see Understanding Society, 2022).

The last pre-COVID-19 main wave is wave 10 and it consists of individuals surveyed during the period 2018–2019.The survey contains a wide range of questions on social, economic, and behavioural issues asked to individuals living in the 12 Government Office Regions across the UK (nine English regions, Scotland, Wales, and Northern Ireland). In April 2020, respondents to the UKHLS were invited to take part in the first wave of a new COVID-19 survey, which consisted of questions on the impact of the pandemic on the well-being of individuals, families, and wider communities, including information about caring responsibilities and family life, employment and financial situation, financial well-being, home schooling, and mental well-being. Participants were asked to complete one survey per month

until July 2020, followed by a survey every two months from September 2020 so as to track changes in their circumstances and environments. There were 17,452 individuals who completed a full post-COVID-19 survey in April 2020 (ISER, 2020), 12,035 in November 2020, and 12,818 in September 2021.

Our estimation sample includes 10,073 observations from the September 2021 data. The missing 2,745 were either missing entirely or missing crucial covariates. Our analysis is based on Wave 10 of the main UKHLS data, which yields pre-COVID background information, together with four of the nine waves of the UKHLS COVID-19 special survey (November 2020, and January, March, and September 2021), which we exploit for their vaccination receipt information in each wave as the rollout progresses. By the September 2021 survey, all individuals aged 16+ will have been vaccinated already or have been offered one, and either have taken-up that offer, or not.

Information about COVID-19 vaccination was collected several times in the COVID-19 UKHLS survey. First, before the vaccine actually became available, in November 2020, individuals were asked: "Imagine that a vaccine against COVID-19 was available for anyone who wanted it. How likely or unlikely would you be to take the vaccine?" They could answer on a scale 1 (very likely) to 4 (very unlikely). Next, after the actual vaccine roll-out had started, individuals were asked in subsequent waves until September 2021 about their actual take up of the vaccine, and could indicate whether they had received one or two doses, whether they had not received any doses yet but were booked for an appointment, and whether they had not received any doses at all. These questions were asked repeatedly in January, March and September 2021. Given the timing of the vaccine roll-out in the UK, all individuals over 18 years of age were offered the possibility of receiving (at least) one dose of COVID-19 vaccine by mid-July 2021. Therefore, we can safely assume that all individuals interviewed in September 2021, would have received the opportunity of at least one vaccine jab.

3.2 Descriptive statistics

Table 1 presents descriptive statistics of the variables of interest, broken down by intention to be vaccinated and actually being vaccinated. There are around 4% of individuals who have not received even a single dose of vaccine by September 2021. This varies significantly by age groups and Table 1 shows that individuals who had not received any dose of the vaccine at the last available wave of the survey (September 2021) are, on average, younger, less educated, and from lower income households in the last pre-COVID-19 wave.

They are also more likely to have engaged in risky health behaviours (especially smoking) and to belong to an ethnic minority. There are no important differences in vaccination take-up or intentions by personality.

Table 1 shows that there are significant differences between individuals who are vaccinated and those who are not (see column 3 and 5 for tests' results). At the same time, the characteristics of the non-vaccinated are very similar to those that say they are unlikely to vaccinate – and this is true for the characteristics that are drawn from the pre COVID wave, as well as those that come from the additional information that was collected in the COVID survey waves. The same is true for the characteristics of those that are vaccinated compared to those that say they are likely to vaccinate. This confirms that intentions are a good predictor of actual take up. However, note that, the non-vaccination rate in the estimation sample is around 3%, while the unlikely-to-vaccinate rate is around 15%. This suggests that most individuals who were sceptical before the vaccine availability, eventually decided to receive it.

Table 2 shows a cross-tabulation of vaccinated and non-vaccinated individuals in September 2021 by hesitancy status in November 2020. The majority of individuals (65.3%) who were hesitant before the vaccine was released, decided to get vaccinated. Only a little over one third (34.7%) of those who said they were very unlikely to get the vaccine persisted in their initial intention.

In fact, of the 325 non-vaccinated individuals, around 15% were initially likely or very likely to take-up the vaccine, while 27% declared they were unlikely and 58% were very unlikely to vaccinate.

Clearly, knowing if the non-vaccinated individuals are employed (and in which sector) can make a difference to the urgency and salience of any intervention because they may then pose a risk to others. We report descriptive statistics of the occupation of hesitant and unvaccinated individuals (in Appendix Table A1). The information about occupation is derived from the last available pre-Covid wave of the regular survey. A sizeable minority (34.6%) of those that are unlikely to vaccinate do not work, and 39% of non-vaccinated individuals do not work. The rest is spread across various occupations with 18% working in unskilled occupations, with around 5-6% working in the health sector. So, these unlikely to vaccinate and those that remain unvaccinated represent a significant risk to other people outside their homes.

Table 1: Descriptive statistics (estimation sample)

	Non-vaccinated individuals	Vaccinated individuals	P-value test of mean difference	Unlikely to takeup vaccine	Likely to takeup vaccine	P-value test of mean difference
Female*	0.68	0.58	0.000	0.67	0.57	0.000
Key worker	0.27	0.24	0.271	0.30	0.23	0.000
Ethnic minority*	0.31	0.08	0.000	0.24	0.07	0.000
Married/living with partner	0.53	0.69	0.000	0.60	0.71	0.000
Smoker*	0.19	0.08	0.000	0.14	0.07	0.000
Heavy drinker*	0.28	0.31	0.242	0.31	0.31	0.774
House owner*	0.66	0.84	0.000	0.70	0.86	0.000
Household gross income (GBP) (sd)*	3568.05 (2627.84)	4443.33 (2931.94)	0.000	3996.46 (2637.43)	4486.70 (2967.26)	0.000
Household had COVID-19	0.03	0.02	0.177	0.03	0.02	0.000
Ever tested COVID-19 positive	0.04	0.02	0.019	0.03	0.02	0.000
Long term illness*	0.28	0.35	0.005	0.31	0.36	0.000
Age						
<20	0.03	0.02	0.000	0.02	0.02	0.000
20-39	0.43	0.16	0.000	0.32	0.15	0.000
40-59.	0.34	0.39	0.878	0.47	0.37	0.000
60-79	0.15	0.36	0.000	0.19	0.44	0.000
>80	0.01	0.03	0.020	0.01	0.04	0.000
*Employment status**						
Employed	0.52	0.50	0.880	0.60	0.48	0.000
Self-employed	0.11	0.08	0.017	0.09	0.08	0.022
Unemployed	0.07	0.02	0.000	0.04	0.02	0.000
Out of labour force	0.31	0.40	0.001	0.27	0.42	0.000
*Education**						
Degree or other higher qual	0.42	0.51	0.005	0.41	0.52	
A levels	0.25	0.20	0.036	0.25	0.20	0.000
GCSE	0.20	0.18	0.394	0.22	0.17	0.000
Lower qualification	0.06	0.07	0.688	0.07	0.07	0.760
No qualification	0.07	0.04	0.017	0.05	0.04	0.102
Big-5 Personality[*+] *(sd)*						
Agreeableness	5.6 (1.1)	5.6 (0.98)	0.871	5.6 (1.7)	5.6 (0.9)	0.113
Conscientiousness	5.5 (1.1)	5.6 (1.0)	0.268	5.5 (1.0)	5.6 (1.0)	0.252
Extraversion	4.7 (1.3)	4.5 (1.3)	0.069	4.6 (1.3)	4.5 (1.3)	0.132
Neuroticism	3.6 (1.4)	3.6 (1.4)	0.521	3.6 (1.4)	3.5 (1.4)	0.115
Open to experience	4.7 (1.4)	4.6 (1.2)	0.157	4.5 (1.3)	4.6 (1.2)	0.070
*Unconventional news sources**						
Personal experience	0.02	0.007	0.055	0.016	0.006	0.000
Friends	0.01	0.007	0.520	0.009	0.007	0.269
Word of mouth	0.06	0.02	0.000	0.03	0.01	0.000
N	325	9,748		1,473	8,600	

Note: * = variable taken from *Understanding Society* main wave 10 in 2019. + Personality traits are scores from 1 to 7, where 7 is best. Unlikely includes very unlikely, and likely includes very likely. News sources add to 1 and mainstream news is the conventional source.

Table 2 –Vaccination decisions by hesitancy status

Likelihood of vaccine take-up (November 2020)	Vaccinated (%) (September 2021)	Non-Vaccinated (%) (September 2021)	Total (N)
Very likely	99.7	0.3	100 (5,992)
Likely	98.7	1.3	100 (2,608)
Unlikely	90.8	9.2	100 (925)
Very Unlikely	65.3	34.7	100 (548)
Total (N)	(9,748)	(325)	100 (10,073)

The following figures highlight out two of the dependent variables broken down by some of the variables that have figured in the existing literature. Figure 2 shows that the age profile of non-vaccinated individuals for males, females, and all - which shows that older individuals are less likely to be non-vaccinated.

Figure 2 – Age profile of non-vaccinated and vaccinated individuals (%)

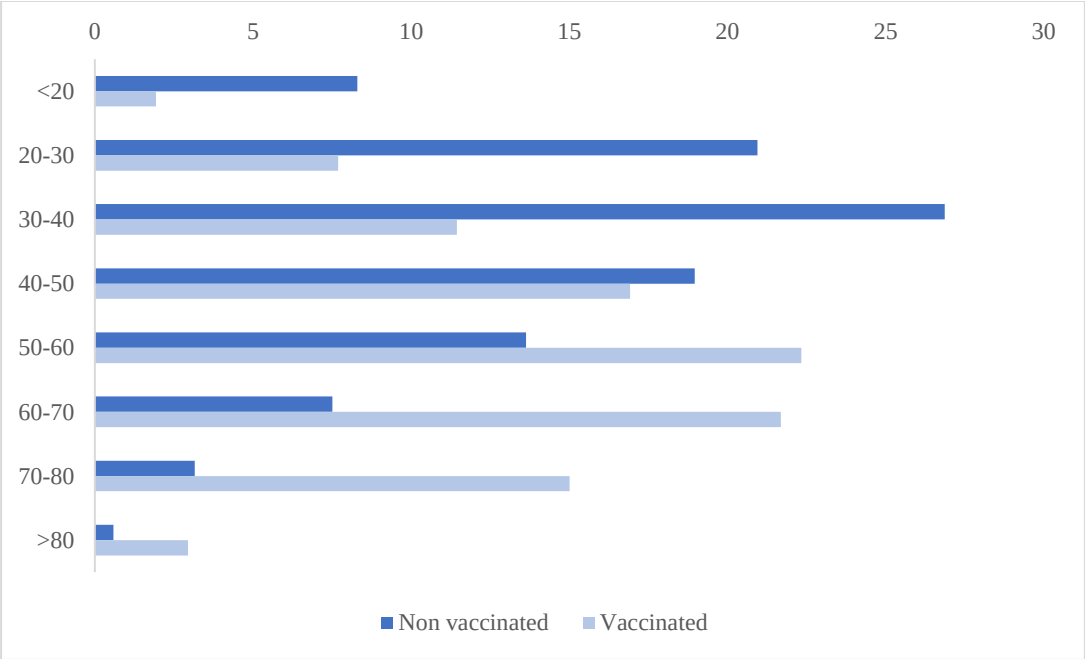


Figure 3a shows the time path of the two dependent variables across the development of the pandemic and the increasing availability of vaccines. There was very little supply available until January 2021- so there were few vaccinated and even fewer with an appointment but waiting for their jab – but 2% of the sample claimed to be very unlikely to get vaccinated, and a further 4% claimed to be unlikely to vaccinate. By September 2021 96% were vaccinated, and only 3% were still saying that they were unlikely or very unlikely to vaccinate. Figure 3b shows that the proportion of non-vaccinated individuals is always higher among ethnic minorities.

Figure 3a Hesitancy Across the COVID-19 Pandemic – General population

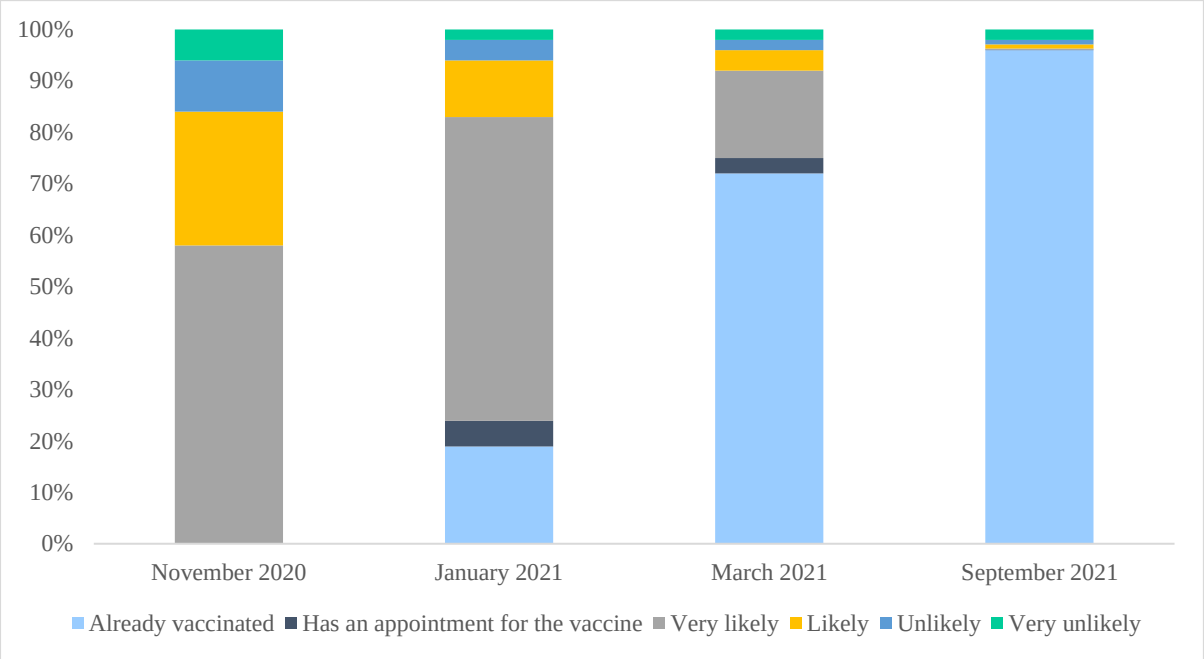
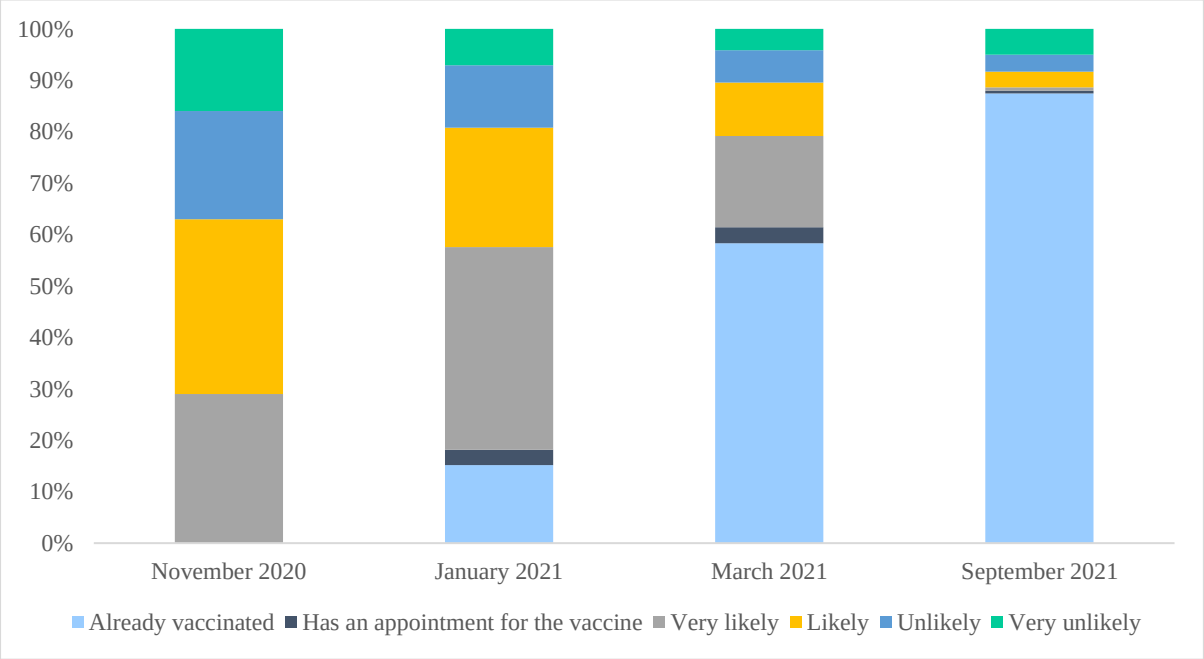


Figure 3b Hesitancy Across the COVID-19 Pandemic – Ethnic minorities



What we see in Figures 3a and 3b is reflected in Figure 4 which shows the non-vaccination rate by region and ethnicity. The biggest differentials in the vaccination rates by ethnicity occur in London and the South East, where the minority rate in approximately 13% compared to the population rate of just 2 or 3%. There is a similar distinction by age in Figure 5 which shows the non-vaccination rate by age and ethnicity – at each age the minority non-vaccination rate is around three times the population rate. These figures indicate worrying levels of non-

compliance in specific groups despite the realisation that the vaccines represented essentially private benefits because of their effectiveness in reducing the severity of the effects. This is important because even relatively small clusters of individuals with high non- vaccination rates can be a source of outbreaks that lead to further waves of cases and deaths.

Figure 4 Non-vaccination rate by region and ethnicity

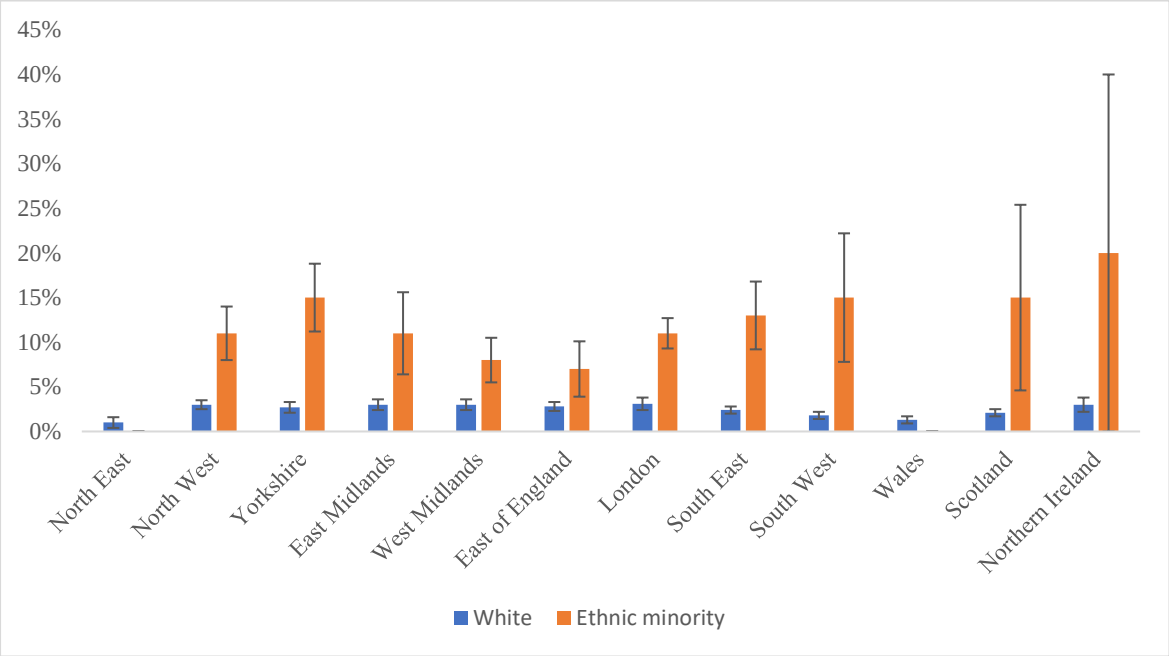
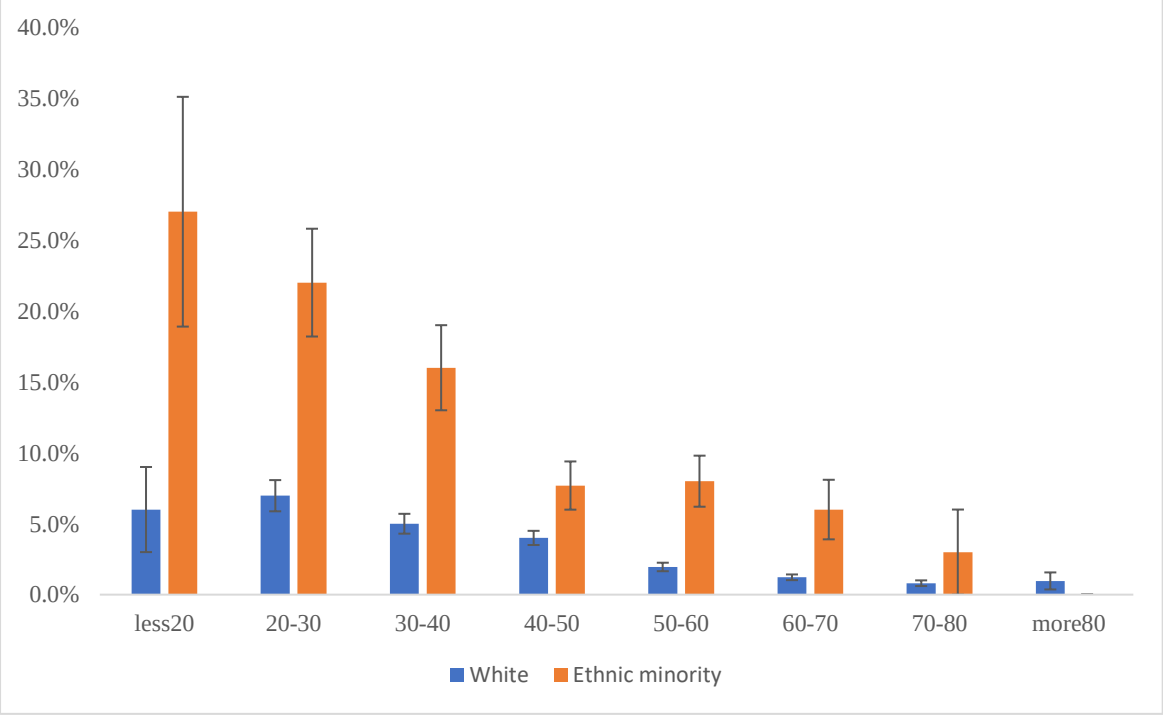


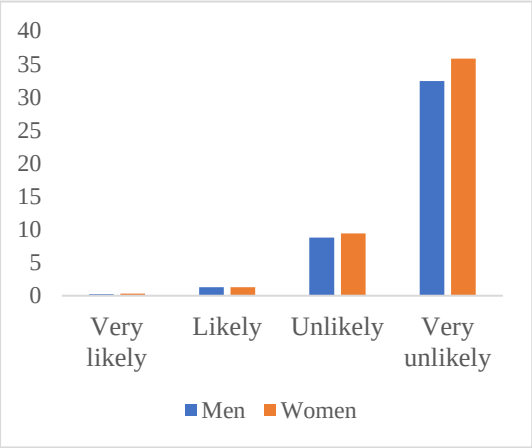
Figure 5 Non-vaccination rate by age and ethnicity



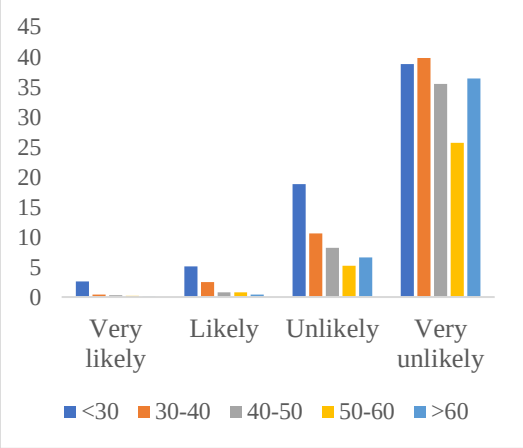
Below, we present the data on vaccination intentions in November 2020 and non-vaccination rate in 2021. Overall, the vast majority of individuals who said they were likely to take the vaccine in November 2020 confirmed their intentions and got vaccinated in 2021. Among those who said they were ***unlikely*** to get vaccinated, around 9% did not change their mind and ended up not taking the vaccine (Figure 6a). This proportion is similar across gender but is higher for younger individual and those from ethnic minorities (see Figure 6b and 6c).

Figure 6 – Non-Vaccination rates by intentions

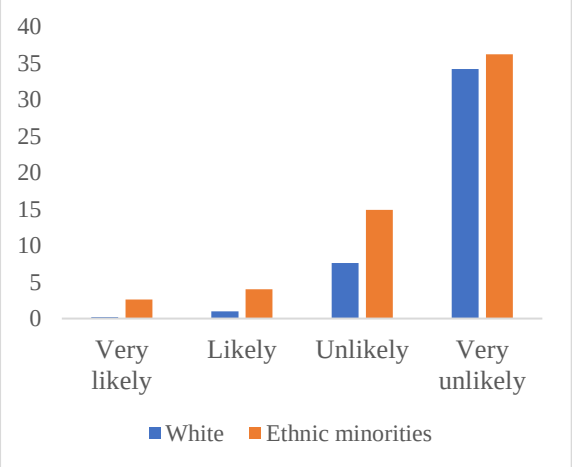
a) By gender



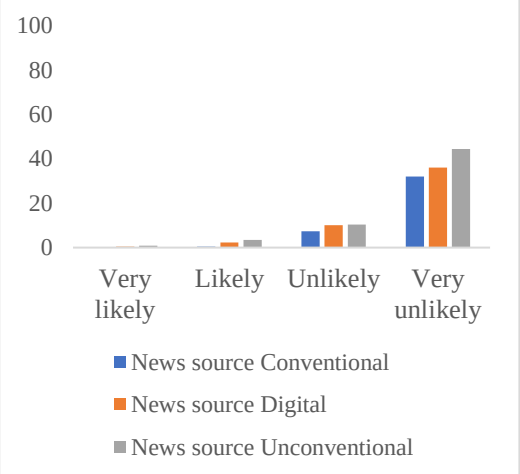
b) By age



c) By ethnic minority status



d) By news source



In the group of individuals who declared to be very unlikely to take a vaccine in November 2020, around 35% of individuals did not take-up offers of vaccination in 2021, and this proportion is stable across gender and ethnic minority status. As above, this figure is higher for young individuals (see Figure 6b). News source also seem to play a role. Among those who usually get their news from unconventional sources (personal experience, friends, and word of mouth) and had previously declared to be unlikely to take the vaccine, almost 50% did not

change their mind and did not get vaccinated (see Figure 6d). This descriptive evidence shows the importance of communicating public health messages in an effective and inclusive way, to reach individuals who may be sceptical or resistant to important scientific evidence.

4. Methodology

We estimate a bivariate probit model, where we jointly estimate the determinants of two probabilities. The first being that of **not** taking-up the offer of a vaccine when one has become available (coded as either very unlikely or unlikely, compared to likely or very likely). This is denoted Y_{1i} and is asked in September 2021. The second being the absence of an intention to vaccinate when one is not yet available. This is denoted by Y_{2i} and is asked in November 2020, before the COVID-19 vaccine became available to the British population. Thus, we estimate the parameters $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ and the covariance between the residuals of each equation, ρ, of the following model, which controls for observable confounders $\boldsymbol{X}_i$:

$$Y_{1i} = \mathbf{1}\left[\alpha + \boldsymbol{\beta}'_1 \mathbf{X}_i + \varepsilon_{1i} \geq 0\right]$$
$$Y_{2i} = \mathbf{1}\left[\alpha + \boldsymbol{\beta}'_2 \mathbf{X}_i + \varepsilon_{2i} \geq 0\right]$$

where $\begin{pmatrix}\varepsilon_{1i}\\ \varepsilon_{2i}\end{pmatrix} \sim N\left[\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}\right]$

where the equations use the index function format that indicates that the dependent variable is 1 if the condition in square brackets is satisfied and 0 otherwise. The estimates are invariant to whichever response is treated as "first". We estimate two versions of the model. The first one is a parsimonious one, where the vector X only includes age, gender, education, ethnicity and region of residence. In the second one, we add several control variables, including detailed characteristics on individual employment, marital status, health behaviours, trust in government institutions, etc. (the detailed list is reported in Section 5). In the Appendix, we provide estimates that treat Y_2 as an *ordered* dependent variable, taking the values 1, 2, 3 or 4 to indicate intensity We show that the effects of this generalisation on the estimates of the determinants of the probability of being vaccinated, ie Y_1, remains insignificantly different from that reported in the main text.

It is important to realise that the adoption of a bivariate model does **not** imply that the estimates of either equation would be different, in probability limit, to corresponding estimates from two independent probit equations (i.e. imposing the restriction that $\rho = 0$ on the model). In particular, relaxing this restriction does not allay any concerns regarding selection bias associated with endogenous explanatory variables. For reassurance on this point, we rely on the richness of our data, and the that allows us to control for a wide array of factors that have been noted across the literature. Moreover, these are measured pre-COVID so there is less

chance of these variables being contaminated by the pandemic itself. In addition, we follow the normal practice of examining the robustness of the coefficients to adding further controls that *might* reveal potential selection bias associated with remaining unobservables. Oster (2019) rightly cautions against relying on such an approach, since there is no guarantee that adding covariates would necessarily reduce the covariance between the error term and the explanatory variables of interest. Fortunately, Oster (2019) is one of a number of recent papers that provide a partial solution to this problem. She proposes that by placing a restriction on the size of this unknowable covariance one can estimate a corresponding bound on the estimate of the coefficient of interest. Equivalently, one might calculate the size of this unknowable covariance that would generate a coefficient of zero (or any other value). In the first case, one might imagine that your dataset has been collected with your research in mind – so that the variables that it contains are better predictors of the dependent variable than those variables that were not collected. That is, one might assume that the covariance between the included variable and the error term should be no smaller than the covariance between the excluded variable and the error. Placing a restriction that the ratio of these two covariances is 1 might seem reasonable and leads to a new estimate of the coefficient on the observable variable that can also thought to be reasonable. In the next section we implement this Oster approach.

One final methodological problem is attrition. There are 2604 observations (20%) missing from the November 2020 file, which records intentions to vaccinate, but present in the September 2021 which records actual vaccination status. This is a non-trivial attrition rate that has the potential to bias estimates. We use an Inverse Probability Weighting approach whereby we estimate a missingness probit as a function of the covariates from our econometric model of intentions and actual vaccination, and then use the inverse of the resulting probabilities as weights in the econometric model to address the potential bias from the missingness not being at random

5. Results

We are interested in understanding the main correlates of vaccine hesitancy and in investigating the relationship between pre-COVID-19 social marginalisation, and decisions about vaccination. We begin by estimating a parsimonious model of the joint probability of being sceptical about getting the COVID-19 vaccine and actual take up of at least one dose of vaccine before September 2021. This model includes a limited set of covariates: gender, age, ethnicity, region of residence, and education.

The second specification adds a set of further controls to get a complete picture of the determinants of vaccine hesitancy but, at the same time, being wary of adding bad controls. Thus, we add: marital status; whether the individual or any family members ever tested positive for COVID-19; and a series of pre-COVID-19 characteristics, including: where the individual gets most of her/his news (mainstream media; personal experience; friends and relatives; word of mouth); labour force status; keyworker status; health status; risky behaviours (smoking and heavy drinking); quartiles of household income; and an indicator of civic engagement. This latter indicator is based on the following questions: "On the whole, are you very satisfied, fairly satisfied, a little dissatisfied or very dissatisfied with the way democracy works in this country"; "Public officials don't care much about what people like me think"; "People like me don't have any say in what the government does"; "I would be seriously neglecting my duty as a citizen if I didn't vote"; "I feel a sense of satisfaction when I vote". Answers to the last four questions are on a scale 1 to 5, from "strongly agree" to "strongly disagree". We combine the answers above and create an indicator of civic engagement and sense of belonging of the society using factor analysis (see Appendix Table A2 and A3 for details about the creation of the indicator) and then classify individuals with a score in the bottom decile of the distribution of this index as having very low levels of civic engagement.

Risks can be multiplicative and Figures 4 and 5 above suggest that there may be some merit in including interactions between variables such as ethnicity and age and gender. Others also come to mind – such as key worker, gender and living with an elderly person. However, our cell sizes would be rather too small to have the power required. This, our estimates need to be interpreted as the effect of one variable evaluated at the mean of other variables.

In the Appendix we report further results that use contemporaneous (i.e. during the pandemic) data. In particular, we add household composition and employment status. These covariates are more likely to suffer from endogeneity because they are measured contemporaneously to the outcomes, but we believe they can offer some useful insights in the main findings and therefore these results are reported in the Appendix (see Table A8-A9). We also report results from an additional specification, where age is included as a series of binary variables (rather than linear and quadratic variable) (Table A12). These results are very similar to the main ones presented in the rest of the paper, and the coefficients of the age bands confirm that, as expected, young individuals are more likely to be against vaccination, while the elderly are more likely to comply and get vaccinated.

The estimated parameters capture the relationship between observable characteristics and the decision to take-up an available vaccination and the intention to do so. One purpose of this study is to help policymakers to profile individuals with tailored interventions - such as information regarding vaccination availability or effectiveness. For this reason, even if the main results may be partially driven by unobservable characteristics, it is still important that individuals with specific observable traits are targeted for interventions. We cannot, in this data, address the selection on unobservables problem. That is, the effects of the X's might, in part, be due to a correlation between the X's and the unobservable determinants of hesitancy. Thus, we compare estimates the coefficients on the variables in the parsimonious model with those in the more comprehensive model, in the hope that more controls will reduce the problem.

We also use recently developed tests (Oster, 2019) to assess the *potential* role of selection on unobservables and we present some reassuring results in Appendix Table A10. In order to use this test, we estimate a linear probability model (LPM) where the outcome variable is a binary variable equal to 1 if an individual was hesitant in November 2020 and not vaccinated in September 2021. The results are presented in Table A10, where we report estimates of the parameter δ, developed in Oster (2019), which indicates the level of selection on unobserved variables (assuming proportionality to the level of selection on observed variables), required to drive the treatment effect to zero. The assumptions behind the calculation of δ can be varied. In particular, it is possible to vary the assumed value of *R-max,* defined as the *R*-squared from a hypothetical regression of the outcome on treatment and both observed and unobserved controls. We follow Oster (2019) and set *R-max* equal to 1.3 times the *R*-squared from a regression of the outcome on the treatment and observed control variables. All of the estimates of the δ parameter associated with Model 1 and 2 are above 1, consistent with an 'acceptable' level of selection based on the rule-of-thumb suggested in Oster (2019). These results provide evidence supporting the credibility of our main estimates, showing that it would take large, likely implausible (the value of δ) levels of selection bias to drive our results to zero and therefore selection on unobservable is unlikely to overturn our main conclusions.

The weighted estimates to address attrition are not materially different from the unweighted ones, so we present the unweighted results throughout the paper while the weighted results are available in Appendix Table A11.

We include exogenous covariates in the first, most parsimonious, model. In the second model, we increase the set of independent variables, in order to understand vaccine hesitancy

better and see if, and how, the estimated results from the basic model change when we include additional covariates. This makes an important difference – looking only at the significant effects we find that the sizes of the marginal effects are approximately halved (see Appendix Table A4 for the underlying estimates).

We select the main results for visualisation in Figures 7 and 8. The marginal effects themselves can be founds in Appendix Table A13a and Table A13b. Several variables which may be thought to be important are notable for being insignificant, such as: where one gets one's news; being a keyworker; prior health; having previously had COVID; personality traits, and the regional case and mortality rates. This suggests that the attention given to these factors in earlier work reflects the absence of other controls, that are available to us and are correlated with these.

A feature of our research is we allow for the possibility that individuals might change their minds, in either direction, when vaccines become available. Thus, we present results for the marginal effects of all variables on the probability of switching opinion (in any direction) in the Appendix (Table A13b). The first column of Table A13b presents results on the marginal effect of relevant variables on the probability of *switching* **from** being non-hesitant to being unvaccinated, while the last column shows the marginal effects for the probability of *switching* from hesitant **to vaccinated.** These show that there are no statistically significant covariates that are also quantitatively important for switching - in either direction. In particular, no observable characteristics seem relevant to predicting the probability that hesitant individuals will switch to opting to vaccinate (see the last column), apart from a very small effect of age squared (showing that middle-aged individuals become more likely to switch). These results show that waiting for intentions to be turned into decisions has little benefit – little new is learned because the factors that drive changes of mind are idiosyncratic. That is, interventions to encourage take-up when a vaccine becomes available can be effectively designed on the basis of data on intentions alone.

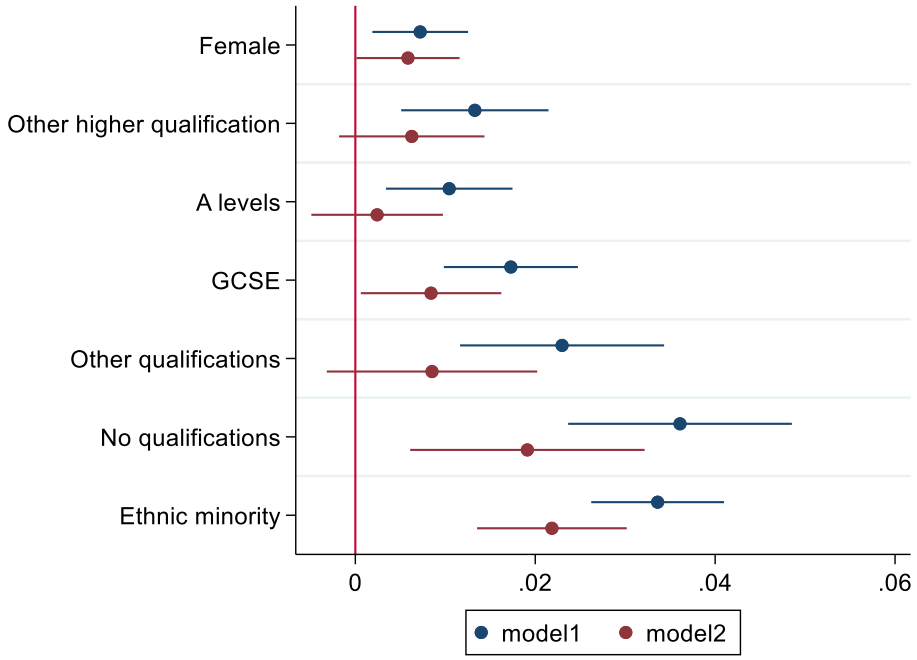
In Figures 7 and 8 below, point estimates and 95% confidence intervals of the marginal effects from Table A13a (derived from the estimates in Appendix Table A4) for selected variables of interest. Figure 7 shows that females are around 1 percentage point more likely to be *both* unwilling to be vaccinated *and* to have not actually received at least one dose of COVID-19 vaccine, in both models 1 and 2, compared to males. The non-vaccination rate in the data is around 3% and an effect size of 1% is a large effect relative to the mean in the data. This implies that the female probability of not intending to vaccinate and then not accepting a

vaccination when available is around three times the probability for men (with the characteristics of the omitted category) which is less than ½% . In both models the effects are, according to the confidence intervals, only marginally significant. Figure 7 also shows that individuals who have no qualifications are around 2 percentage points (3.5 pp for model 1), more likely to be as hesitant than individuals with a degree. And those who belong to an ethnic minority are 3.3 percentage points (2 pp in the case of model 2) more likely to not intend to vax and to be unvaxed, relative to white. We are also interested in analysing the relationship between marginalisation in society and vaccine hesitancy. But it is possible that such individuals became less trustful towards the government and its institutions in the course of the pandemic. For this reason, we include a series of indicators of socio-economic status and participation in society *at the last pre-COVID19 wave,* including income quartiles, employment status, and home ownership. We also consider the possibility that vaccine hesitancy is related to risk-loving behaviours and therefore we control for smoking and drinking status. Figure 8 shows the distribution of the probabilities of effect of income quartile relative to the top quintile case. Lower quartiles have positive effects on hesitancy with the highest having a statistically significant 2% effect – again, large relative to the mean probability.

Figure 8 also shows the association (relative to the baseline individual whose probability of being unlikely to vaccinate and be unvaccinated is 1.1%) of new variables with the probability of being hesitant with respect to the vaccine. Interestingly, the most noticeable effects come from being a smoker or self-employed (around + 1.5 p.p), belonging to the lowest income quartile (+2 p.p.) and showing low participation and low sense of belonging to the society (+ 1.3 p.p.). Further, with respect to results from the basic model, gender is no longer significant, while the effects of low education and belonging to an ethnic minority are slightly reduced (+2 and +2.4 p.p.). Again, the sizes of these effects are large, considering that the mean of the non-vaccination rate in the estimation sample is around 3%.

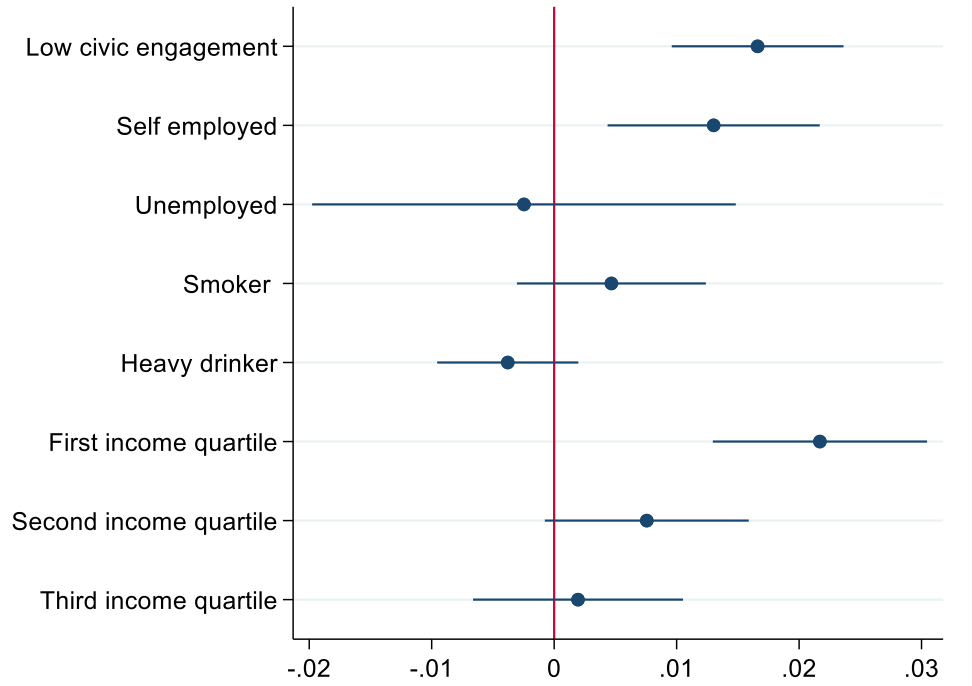
These results are consistent with previous evidence on factors contributing to vaccine hesitancy, even if we extend existing findings by looking at actual take-up of vaccination. For example, Galasso et al (2021) show that specific interventions aimed at providing information on reasons to get the vaccine (e.g protecting others, protecting health, and protecting the economy) increase vaccination intentions and vaccination rates by 2 to 4 percentage points.

Figure 7 Bivariate Probit Estimates of Selected Marginal Effects of Being Very Unlikely / Unlikely to Vaccinate and Not Getting Vaccinated



Note: Zero corresponds to an individual who is white, male and has a degree which, at the average age, implies a probability of being unvaccinated and being unlikely to get vaccinated of 0.45% (for Model 1). Confidence intervals are 95%.

Figure 8 Bivariate Probit Estimates of Selected Marginal Effects of Being Very Unlikely / Unlikely to Vaccinate and Not Getting Vaccinated (Model 2)



Note: Zero corresponds to an individual who is white, male, has a degree, is employed, in the top income quartile, not a heavy drinker, and is not a smoker, which, at the average age, implies a probability of being unvaccinated and being unlikely to get vaccinated of 1.1% (in Model 2). Confidence intervals are 95%.

Similarly, Dabla-Norris et al (2021) show that women are around 4 p.p. less likely to vaccinate and that trust in the capacity of the healthcare system to respond to COVID-19 increases the probability of vaccination by up to 5 percentage points. Lastly, Edwards et al (2021) show that women were around 2 p.p to be hesitant, those who had confidence in their state or territory government or in their hospitals and health system were less likely to be resistant to vaccination (-3.4 and -4.4 percentage points respectively).

In the less parsimonious version of the two models presented in Table A13a, and whose estimates are visualised alongside those of model 1 in Figure 7 and in Figure 8, where we also control for previous experience of COVID-19 in the individual's household (we know whether the individual or any of her/his family members ever tested positive) and local infection and death rates in the month of the interview. None of these variables significantly affect hesitancy.

As a sensitivity analysis of the main results, we limit the sample to individuals over 50, in order to understand whether the effect sizes of the main determinants of vaccine hesitancy vary with age. The main results are unchanged and vaccine hesitancy continues to be higher for individuals with low socio-economic status and low levels of civic engagement (see Table A5).

We estimate a further version of the model, which includes additional covariates measured during the pandemic and, in particular, includes the presence of children or elderly in the family, and employment status. Results are reported in the Appendix (Table A8-A9) and confirm the main findings. Not surprisingly, individuals living with elderly relatives are less likely to be hesitant or non-vaccinated. The other results are very similar to the ones presented in the previous tables for Model 1 and 2, and the main variables identified above remain significant predictors of hesitancy and not-vaccination. We also estimate a version of the model which decomposes the binary nature of the likelihood to take the vaccine into finer categories. The results of this ordered probit equation (estimated using *bioprobit* routine in Stata) for Y_2 are reported in the Appendix Table A6 and A7 and are consistent with previous findings. Individuals from ethnic minorities, low socio-economic status and pre-COVID 19 low civic engagement were less likely to be willing vaccinate and less likely to actually take the vaccine.

6. Conclusion

We use a bivariate probit model to jointly model intentions to vaccinate against COVID-19 and subsequent actual vaccination using high quality longitudinal data from the UK collected before and during the pandemic. Most of the existing literature uses data on

intentions to be vaccinated, rather than actual take up and our results are interesting and innovative with respect to previous studies, as we model the joint probability of being hesitant, before the vaccines were released, and of not accepting any doses once the vaccine became generally available.

The empirical specification allows for the unobservable determinants of each response to be correlated with each other – for example, people who have weak intentions to vaccinate, for idiosyncratic reasons, are prone to also be less likely, than their observable characteristics suggest, to vaccinate when it is possible to do so. The correlation between the unobservable determinants of these two decisions is high, indicating that individuals are likely to follow through with their initial intentions, although we do find that a significant proportion of individuals did switch from their initial hesitancy.

The main results of this analysis show that women, ethnic minorities, individuals with low education, and those with low income are significantly more likely to be hesitant towards the COVID-19 vaccine and to fail to take it up. Similar effects are found for those who were self-employed and poorly engaged with society before the pandemic started.

These results show that it is possible to plan pro-vax interventions even prior to actual availability of vaccines before people form negative opinions and scepticism. Indeed, we show that there are no characteristics of individuals that can be used to target sceptical individuals to persuade them to change their mind once opinions have been formed. Further, some sub-groups of individuals should be targeted for more intensive interventions, possibly through GPs or community leaders.

References

Ashworth,M., Thunström, L., Cherry, T., Newbold, S., Finnoff. D. (2021) Emphasize personal health benefits to boost COVID-19 vaccination rates. PNAS 118 (32), e2108225118; https://doi.org/10.1073/pnas.2108225118.

Campos-Mercade, P. et al. (2021), Monetary Incentives Increase COVID-19 Vaccinations, Science, 12; 374(6569):879-882.

Cascini, F., Pantovic, A., Al-Ajlouni Y., Failla, G., Ricciardi, W. (2021) Attitudes, acceptance and hesitancy among the general population worldwide to receive the COVID-19 vaccines and their contributing factors: A systematic review, EClinicalMedicine 40- 101113: 1-14.

Caserotti, M., Girardi, P., Rubaltelli, E., Tasso, A., Lotto, L., Gavaruzzi, T. (2021) Associations of COVID-19 risk perception with vaccine hesitancy over time for Italian residents. Social Science and Medicine 272 113688.

Chauduri, K., Chakrabarti, A., Chandan, JS, Bandyopadhyay, S. (2022) COVID-19 vaccine hesitancy in the UK: a longitudinal household cross-sectional study. BMC Public Health 2:104 https://doi.org/10.1186/s12889-021-12472-3.

Dabla-Norris, E., Khan, H., Lima, F., Sollaci, A. (2021) Who Doesn't Want to be Vaccinated? Determinants of Vaccine Hesitancy During COVID-19. IMF working paper 21/130.

Diament, S., Kaya, A., Magenheim, E. (2022) Frames that matter: Increasing the willingness to get the Covid-19 vaccines. Social Science and Medicine 292, 114562.

Edwards, B., Biddle, N., Gray, M., Sollis, K (2021) COVID-19 vaccine hesitancy and resistance: Correlates in a nationally representative longitudinal survey of the Australian population. PLoS ONE 16(3): e0248892. https://doi.org/10.1371/journal.pone.0248892.

Fortunato, P., Lombini, A. (2022) COVID-19 Vaccination, Political Partisanship, and Moral Values. IRENE Working Paper 22-09.

Galasso, V., Profeta, P., Foucault, M., Pons, V. (2021) COVID-19 Vaccine's Gender Paradox, medRxiv, doi: https://doi.org/10.1101/2021.03.26.21254380.

Kennedy, E., Daoust, JF, Vikse J., Nelson V. (2021) "Until I Know It's Safe for Me": The Role of Timing in COVID-19 Vaccine Decision-Making and Vaccine Hesitancy. Vaccines, 9(12), 1417.

Muller, S., Rau H (2021) Economic preferences and compliance in the social stress test of the COVID-19 crisis. Journal of Public Economics 194: 104322.

Murphy, J, Vallières, F., Bentall, RP., Shevlin, M. McBride, O., Hartman, T., McKay, R., Bennett, K., Mason, L., Gibson-Miller, J., Levita, L., Martinez, A., Stocks, T., Karatzias, T., Hyland, P (2021), Psychological characteristics associated with COVID-19 vaccine hesitancy and resistance in Ireland and the United Kingdom. Nature Communication https://doi.org/10.1038/s41467-020-20226-9.

Niek Mouter, Annamarie de Ruijter, Ardine de Wit, Mattijs S Lambooij , Maarten van Wijhe, Jobvan Exel e , Roselinde Kessels (2022), "Please, you go first!" preferences for a COVID-19 vaccine among adults in the Netherlands. Social Science and Medicine 292, 114626.

NHS (2021), COVID-19 Vaccination Statistics. Week ending Sunday 26th September 2021. https://www.england.nhs.uk/statistics/wp-content/uploads/sites/2/2021/09/COVID-19-weekly-announced-vaccinations-30-September-2021.pdf

ONS (2021) "Coronavirus and vaccine hesitancy, Great Britain: 9 August 2021", https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandwellbeing/bulletins/coronavirusandvaccinehesitancygreatbritain/9august2021.

ONS (2022) Coronavirus (COVID-19) latest insights: Vaccines. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19latestinsights/vaccines.

ONS (2022a) Deaths involving COVID-19 by vaccination status, England: deaths occurring between 1 January 2021 and 31 January 2022. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsinvolvingcovid19byvaccinationstatusengland/deathsoccurringbetween1january2021and31january2022

Oster, E. (2019). Unobservable selection and coefficient stability: theory and evidence, Journal of Business & Economic Statistics, Vol. 37, pp. 187–204

Paul, E., Steptoe, A., Fancourt, D. (2021) Attitudes towards vaccines and intention to vaccinate against COVID-19: Implications for public health communications. The Lancet Regional Health – Europe https://doi.org/10.1016/j.lanepe.2020.100012

Robertson, E., Reeve, K., Niedzwiedz, C., Moore, J.., Blake, M., Green, M., Katikireddi, S., Benzeval, M. (2021), Predictors of COVID-19 vaccine hesitancy in the UK household longitudinal study. Brain Behavior and Immunity 94: 41-50.

Sage Vaccine Hesitancy Working Group (2015), Diagnosing the determinants of vaccine hesitancy in specific subgroups: The Guide to Tailoring Immunization Programmes (TIP), Butler and MacDonald. Journal of Vaccine.

Understanding Society (2022). Understanding Society Data Access Strategy. Available at:

https://www.understandingsociety.ac.uk/sites/default/files/downloads/general/ukhls_data_access_strategy.pdf

University of Essex, Institute for Social and Economic Research. (2022), Understanding Society: Waves 1-11, 2009-2020 and Harmonised BHPS: Waves 1-18, 1991-2009. [data collection]. 15th Edition. UK Data Service. SN: 6614, http://doi.org/10.5255/UKDA-SN-6614-16.

University College London, UCL Institute of Education, Centre for Longitudinal Studies. (2021). *COVID-19 Survey in Five National Longitudinal Cohort Studies: Millennium Cohort Study, Next Steps, 1970 British Cohort Study and 1958 National Child Development Study, 2020-2021*. [data collection]. *3rd Edition.* UK Data Service. SN: 8658, DOI: 10.5255/UKDA-SN-8658-3

Ward, J., Alleaume, C., Peretti-Watel, P., the COCONEL Group (2020), The French public's attitudes to a future COVID-19 vaccine: The politicization of a public health issue. Social Science and Medicine 265: 113414.

WHO (2020) Tackling COVID-19 Misinformation. A social media toolkit for healthcare practitioners. Available online at: file:///C:/Users/smendoli/Downloads/VACCINE-MISINFORMATION-TOOLKIT_desktop1.pdf