



City Research Online

City St George's, University of London

Citation: Mu, Yi (2022). Essays on policy evaluation: assessing policy impact on mental well-being among vulnerable populations using quasi-experimental methods. (Unpublished Doctoral thesis, City, University of London)

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://city-ref-test.eprints-hosting.org/id/eprint/30460/>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

**Essays on policy evaluation: assessing
policy impact on mental well-being among
vulnerable populations using
quasi-experimental methods**

by

Yi Mu

A thesis submitted to the Department of Economics

In conformity with the requirements for

The degree of Doctor of Philosophy (PhD)

Department of Economics

City, University of London

Supervisors: Dr. Alice Mesnard, Dr. Agne Suziedelyte

December 2022

Abstract

This thesis investigates the impact of policies on people's mental and physical well-being. It highlights the impact on vulnerable people and explores the heterogeneity of impacts for those from different backgrounds.

Chapter 1 explores the impact of displacement due to land expropriation on physical and mental health using data on household heads and their spouses from the Longitudinal Survey on Rural Urban Migration in China. The results show that the displaced are more likely to be sick or injured and have higher medical expenditure. At the same time, impacts on the standardised mental health factor and self-rated health status are statistically insignificant. This chapter investigates this further using the time since displacement and finds that the impact on mental health is significantly negative for the recently displaced people, while the impact on physical health is significantly negative for the long-term displaced. In order to understand these results, this chapter also explores effects on other outcomes and finds that displaced people have more years of education, fewer children and are less likely to have medical insurance. Those having more years of education are more likely to care about their health and thus spend more on it. Having fewer or no children can explain the higher possibility of being sick or injured. People without a settled family life are more prone to dangerous jobs and unstable lifestyles.

Chapter 2 explores the heterogeneous impacts of forced displacement due to the Three Gorges Dam project (TGDP) on people's mental well-being in China. Using the TGDP survey data on both residents and relocatees, this chapter estimates the average treatment effect on the treated and finds that depression levels increased among the displaced. The impacts for other standardised mental health factors, including anxiety, low mood and low self-esteem, are statistically insignificant. This chapter also finds heterogeneous impacts on mental health considering differences in

people's educational level, marital status, moving distance and whether they were displaced individually or with their village. The illiterate individuals became more anxious than those having general education. The divorced were not as severely affected as the married for depression and low mood but were worse for anxiety and low self-esteem. People who moved out of town became much more depressed than those who moved within the town. Finally, people who moved with the village became more depressed than those who moved with their households alone. These results suggest that heterogeneity of the impacts exists, and thus it should be considered when making policy decisions.

With an event study approach, Chapter 3 explores the dynamics of the heterogeneous impact of lifting shielding by ethnicity for the clinically extremely vulnerable people in the COVID-19 pandemic. There is no evidence suggesting that the high-risk individuals from a BAME background were more likely to have COVID-19 symptoms once the shielding requirement was relaxed than those from a non-BAME background. However, the impact on the diagnosis of mental ill-health is significantly higher for the high-risk BAME than their non-BAME counterparts after the shielding was relaxed in July 2020. Furthermore, they felt more at risk of catching COVID-19 in November 2020 when the daily infections increased, and the shielding was not required. This negative impact continued in January 2021, when the shielding requirement was re-implemented. In terms of behaviour, when the COVID-19 pandemic deteriorated in November 2020, its impact on the use of trains by the high-risk BAME decreased significantly, showing that the high-risk BAME population were more concerned about catching the virus than the high-risk non-BAME population. This could explain the insignificant difference between high-risk BAME and non-BAME in having COVID-19 symptoms; that is, although the high-risk BAME may live in an environment that increases the chance of getting infected than the high-risk non-BAME, their cautiousness lowered their exposure to the virus.

Acknowledgements

First of all, I would like to thank my supervisors, Dr. Alice Mesnard and Dr. Agne Suziedelyte, for their invaluable advice throughout my PhD. They helped me at each step of the journey, from generating research ideas to polishing the completed drafts. Their encouragement helped me overcome the obstacles I faced over these years.

I am grateful to the Department of Economics at City, University of London for the PhD studentship, without which I would not be able to imagine doing a PhD, and for their financial support on my training.

I thank Dr. Fabrice Defever and Prof. Joseph Pearlman for the talk and encouragement in the first year of my PhD. I am grateful for the referee comments from Dr. Edward Pinchbeck and Dr. Giulia Faggio.

I thank Dr. Christopher McDowell for his advice on the Three Gorges Dam project. I thank Dr. Juan Xi for sharing the Three Gorges Dam migrants survey data.

I thank the participants of the PSE Summer School in Migration Economics workshop, EUDN workshop, SPEN PhD seminars and City, University of London internal seminar for their comments and discussion.

Thanks to Michael Gechter, my master's essay supervisor, whose guidance and encouragement had strengthened my eagerness to do research. From him, I first learnt the challenges and excitement of research.

I thank Ivy Sabuga for being a kind, supportive and excellent company on everything we went through together during PhD. I enjoyed talking with Vania Esady, who

was always warm-hearted and provided helpful suggestions. I thank Iris Lopes Rafe-gas, Salma Ben Lalouna and Wasana Kalansooriya for being supportive and always willing to talk.

I thank Jia Xiang, Mengli Sha and Yuta Watabe for their friendship that guided me out of a tunnel.

Special thanks to Xu Zhang, my old friend, for our long chats over these years that helped me take my mind off my research whenever I felt stuck.

I thank my parents for their support and endless trust in my ability to complete the PhD. I thank my brother for always helping me check literature in Chinese and taking care of our parents while I was not around. I thank my parents-in-law for their love, support, and always cheering me up throughout this journey.

Finally, I sincerely thank my husband for being a patient listener, proofreader, knowledgeable discussant, cheerleader, housekeeper, and many more. One thing for sure is that I would not even imagine starting the PhD without his encouragement, and I would not believe that I could complete the PhD without his calm but determined personality.

Contents

List of Figures	1
List of Tables	2
Introduction	7
1 Development-induced displacement and health in China	14
1.1 Introduction	14
1.2 Background Information	19
1.2.1 Urbanisation and land expropriation	19
1.2.2 How does the government make decisions on where to expropriate?	20
1.2.3 <i>Hukou</i>	21
1.2.4 Process of forced displacement and a change of <i>Hukou</i> status	23
1.3 Data	24
1.4 Empirical Framework	28
1.5 Results	31
1.5.1 Main results	31
1.5.2 Supplementary results: time since displacement	34
1.5.3 Mechanisms	36
1.5.4 Heterogeneous effects: living in urban areas versus rural areas	40
1.6 Robustness	42
1.6.1 Matching	42

1.6.2	Treatment effects over time with different time intervals	44
1.7	Conclusion	48
	Appendix 1A	51
	Appendix 1B.a	53
	Appendix 1B.b	56
	Appendix 1C	59
	Appendix 1D	65
	Appendix 1E	77
2	Impact of forced displacement on mental well-being	81
2.1	Introduction	81
2.2	Background Information	85
2.3	Data	86
2.4	Empirical Framework	90
2.4.1	Average effect	92
2.4.2	Heterogeneous impact	96
2.5	Results	97
2.5.1	Average impact	97
2.5.2	Heterogeneous impact	105
2.6	Robustness	115
2.7	Conclusion	119
	Appendix 2A	121
	Appendix 2B	122
	Appendix 2C	128
	Appendix 2D	136
3	Impact of shielding policies during COVID-19 pandemic on BAME versus non-BAME high-risk population	144
3.1	Introduction	144

3.2	Background	149
3.2.1	How are the “high-risk” individuals defined?	149
3.2.2	How have shielding policies been changing over time?	149
3.2.3	Lockdown and COVID-19 pandemic	150
3.3	Data	151
3.4	Methods	157
3.4.1	Overview	157
3.4.2	Estimating the Intent-to-Treat Effects of Shielding Policies Change	158
3.5	Results and Discussion	162
3.5.1	Visuals: BAME versus non-BAME	162
3.5.2	Main Results	164
3.5.3	Supplementary results	169
3.5.4	Does the vaccine help?	173
3.5.5	Ethnic groups (Black and Asian)	176
3.6	Robustness Check	177
3.7	Conclusion	178
	Appendix 3A	180
	Appendix 3B	188

List of Figures

1.1	China Urbanization history	20
1.2	Matched sample versus unmatched sample, 2008	43
1.3	Matched sample versus unmatched sample, 2009	44
2.1	Matched versus unmatched sample	116
3.1	Average daily new COVID-19 infections and deaths	151
3.2	BAME versus non-BAME - Main Outcomes	163
3.3	Event Study Estimates	165
3.4	BAME versus non-BAME - Use of Transportation	171
3A.1	Mortality across different ethnic groups over waves of COVID-19 pandemic in the UK	180

List of Tables

1.1	Scoring coefficients from principal component factor analysis	25
1.2	Sample balance	26
1.3	Health Outcomes	29
1.4	Weight Categories	33
1.5	Self-reported Health Status	33
1.6	Health outcomes with time interval since displacement	37
1.7	Weight categories with time interval since displacement	38
1.8	Medical Insurance	39
1.9	Most recent disease or injury in last 3 months	41
1.10	Education and children	42
1A.1	Correlation among outcome variables	51
1A.2	Mental Health Selection effect Estimation	52
1B.a.1	Health Outcomes - Urban movers	53
1B.a.2	Weight Categories - Urban movers	54
1B.a.3	Self-reported Health Status - Urban movers	55
1B.b.1	Health Outcomes - Rural stayers	56
1B.b.2	Weight Categories - Rural stayers	57
1B.b.3	Self-reported Health Status - Rural stayers	58
1C.1	Matching Quality	59
1C.2	Health Outcomes - Matching	60
1C.3	Weight Categories - Matching	61
1C.4	Self-reported Health Status - Matching	62

1C.5	Medical Insurance - Matching	63
1C.6	Education and children - Matching	64
1D.1	Health Outcomes - Time Interval version 1	65
1D.2	Weight Categories - Time Interval version 1	66
1D.3	Self-reported Health Status - Time Interval version 1	67
1D.4	Health Outcomes - Time Interval version 2	68
1D.5	Weight Categories - Time Interval version 2	69
1D.6	Self-reported Health Status - Time Interval version 2	70
1D.7	Health Outcomes - Time Interval version 3	71
1D.8	Weight Categories - Time Interval version 3	72
1D.9	Self-reported Health Status - Time Interval version 3	73
1D.10	Health Outcomes - Time Interval version 4	74
1D.11	Weight Categories - Time Interval version 4	75
1D.12	Self-reported Health Status - Time Interval version 4	76
1E.1	Unadjusted p-value versus Westfall-Young adjusted p-value	77
1E.2	Health outcomes with time interval since displacement - two-interval version	77
2.1	Correlations between missing mental health information and indi- vidual characteristics	88
2.2	Summary Statistics - Standardized Mental Health Factors	88
2.3	Summary Statistics - Covariates	89
2.4	Summary Statistics - Characteristics for heterogeneity analysis	91
2.5	Sample Balance	92
2.6	Intention to Treat Effect - Depression	98
2.7	Intention to Treat Effect - Anxiety, Low Mood and Low Self-esteem	100
2.8	Treatment Effect on the Treated - Depression	103
2.9	Treatment Effect on the Treated - Anxiety, Low Mood and Low Self-esteem	104

2.10	Anticipation	105
2.11	Gender	106
2.12	Education	108
2.13	Marital Status	110
2.14	Moving Distance	112
2.15	Moving Model	114
2.16	Sample Balance with Matched Sample	117
2A.1	Principal Component Factor Analysis - Regression Score	121
2B.1	Intention to Treat Effect - Anxiety	122
2B.2	Intention to Treat Effect - Low Mood	123
2B.3	Intention to Treat Effect - Low Self-esteem	124
2B.4	Treatment Effect on the Treat - Anxiety	125
2B.5	Treatment Effect on the Treat - Low Mood	126
2B.6	Treatment Effect on the Treated - Low Self-esteem	127
2C.1	Intention to Treat Effect - Depression (matched sample)	128
2C.2	Intention to Treat Effect - Anxiety (matched sample)	129
2C.3	Intention to Treat Effect - Low Mood (matched sample)	130
2C.4	Intention to Treat Effect - Low Self-esteem (matched sample)	131
2C.5	Treatment Effect on the Treated - Depression (matched sample)	132
2C.6	Treatment Effect on the Treated - Anxiety (matched sample)	133
2C.7	Treatment Effect on the Treated - Low Mood (matched sample)	134
2C.8	Treatment Effect on the Treated - Low Self-esteem (matched sample)	135
2D.1	Unadjusted p-value versus Westfall-Young adjusted p-value - full sample	136
2D.2	Unadjusted p-value versus Westfall-Young adjusted p-value - re- duced sample	137
3.1	Summary statistics for individual characteristics	153
3.2	Summary statistics for outcomes of interest	154

3.3	Scoring coefficients from principal component factor analysis . . .	156
3.4	Summary statistics for sub-group individual characteristics	161
3.5	Event Study of Shielding Requirement Lift and Health Outcomes .	166
3.6	Face-to-face contact with families and friends who do not live with them	170
3.7	Use of Transportation	172
3.8	Vaccination	175
3A.1	Event Study of Shielding Requirement Lift and Health Outcomes .	181
3A.2	Transportation - BAME (complete)	182
3A.3	Event Study of Shielding Requirement Lift and Health Outcomes - Black	183
3A.4	Event Study of Shielding Requirement Lift and Health Outcomes - Asian	184
3A.5	Face-to-face contact with families and friends who do not live with them - Asian and Black	185
3A.6	Transportation - Asian and Black	186
3A.7	Vaccination - Asian and Black	187
3B.1	Event Study of Shielding Requirement Lift and Health Outcomes with Standard Errors Clustered at PSU Level	189
3B.2	Transportation with Standard Errors Clustered at PSU Level . . .	190
3B.3	Unadjusted p-value versus Westfall-Young adjusted p-value	191

List of abbreviations

ATT	average treatment effect on the treated
BAME	Black, Asian and Minority Ethnic
BMI	body mass index
DDIV	instrumented difference-in-differences
FE	fixed effects
ITT	average intention-to-treat effect
MHS	Migrant Household Survey
NHS	National Health Service
OLS	Ordinary Least Square
PCR	polymerase chain reaction
RHS	Rural Household Survey
RUMiC	Longitudinal Survey on Rural Urban Migration in China
TGD	Three Gorges Dam
TGDP	Three Gorges Dam Project
UHS	Urban Household Survey

Introduction

No matter what our position in society may be, we are all influenced by policies, from substantial legislation such as increasing the nationwide minimum wage to more targeted measures like providing subsidies to local schools. However, how great is the impact of policy change? Who benefits from or is left at a disadvantage by them? Is the policy objective achieved? Moreover, does the policy have any unexpected consequences? Over recent years, it has become widely accepted that policy evaluation is vital to the evidence-based policymaking process. However, it has been quite common that policies are implemented without evaluation, and sometimes they are not even piloted before adoption. This is true in developing countries like China, but it is also the case in developed countries like the UK. Many programmes or policies require vast resources, money and labour, but the real impact is often never known. Moreover, in many cases, they failed to consider the side effects of these policies. Given this concern, this thesis investigates three policies that were never evaluated primarily from a rigorous quantitative perspective. By providing such evaluation of these policies, this thesis demonstrates how policy evaluation can provide insights into the policymaking process; and also, how policies can appear misleading without evaluation. Furthermore, the evaluation also offers insights into how these policies might be better designed should they need to be implemented again in the future.

This thesis focuses on three quantitative policy evaluations. The first policy (Chapter 1) is migration due to land expropriation. As a fast speed developing country, urban areas have expanded rapidly during the last three decades in China.

A significant amount of rural land has been converted into skyscrapers and surfaced streets. The people affected have had to leave their homes, give up their arable land and begin life afresh as urban residents; their social network was usually torn apart as many people had to relocate far from their place of origin. Such a forced move has brought challenging experiences to the migrants. Moreover, city expansion is still ongoing. Thus it is essential to understand how such a policy affects the migrants so that they might be eventually supported in effective ways.

The second policy (Chapter 2) is forced displacement due to the Three Gorges Dam project. As a means of enlarging the potential to generate clean energy, for the past half-century, China has been constructing hydropower stations and dams. In recent decades, such projects have become more numerous to meet high energy demands. However, these projects require mass migration since they entail the deliberate flooding of vast areas of dry land. Anecdotal evidence shows that residents are usually dissatisfied at the prospect of having to leave their homes and land. Such projects are still being actively planned, so more people are expected to be displaced. Therefore, understanding the impact of such projects on the well-being of those displaced is essential. The Three Gorges Dam project is one of the most significant hydroelectric projects in the world, and it requires a displacement of more than one million people. Therefore, evaluating this case study of forced displacement should provide helpful information that can guide similar undertakings in the future.

The third policy (Chapter 3) is shielding during the COVID-19 pandemic in the UK. This policy required patients who were extremely vulnerable in a clinical sense to shield themselves at home to prevent them from catching coronavirus. Shielding was necessary in the UK when the number of new cases surged. At the outset, no vaccine was available, and subsequently, those at high risk had to wait to be vaccinated. As every pandemic experiences a period when no medical protection is available, and the COVID-19 pandemic will not be the last, understanding shielding policy is important for the future. It is crucial to focus on shielding; however, it is also vital to consider

the impact of lifting these precautions. We must ensure that lifting shielding does not increase the risk of catching COVID-19 or other adverse effects for vulnerable patients.

Chapter 1 explores the impact of land expropriation accompanied by a change in household registration (*Hukou*) on people’s mental and physical well-being. City expansion through land expropriation has been a prevalent phenomenon in China over the past decades. The experience of being expropriated from their home and being forced to become urban residents is regarded as an exogenous shock that is beyond the control of the people involved. Although there are more opportunities for economic improvement in cities, becoming urban residents may not be an experience that is beneficial to every aspect of these people’s lives. Despite the long-lasting debate on land expropriation and resident resettlement, most evaluations were conducted based on observation or qualitative evidence but not rigorous quantitative methodologies. As projects such as city expansion are still ongoing, realising how this process may affect people’s well-being is very important to policymakers. They could make policy adjustments when they understand the issues better. Exploiting the exogenous shock – land expropriation – as the treatment in a quasi-experiment, this chapter first seeks to determine the typical impact that obtaining an urban *Hukou* through land expropriation has on mental and physical well-being: those concerned are more likely to experience sickness or injured and incur higher medical expenses, while the impact on mental health, self-rated health status and weight categories are not statistically significant. To understand how different subgroups respond to the policy, this study explores the heterogeneity based on how long people have been displaced. Considering the time since obtaining urban *Hukou* through land expropriation, the study finds a significant short-term negative impact of the treatment on mental health, which eventually faded over time. However, the opposite is found for physical health, suggesting that the new urban lifestyle has a negative effect. Further exploration of the mechanisms suggests possible explanations for these results. For example, the higher

education level of the treatment group could explain their higher medical expenditure (including reimbursed or reduced amount), as people from a higher educational background usually pay more attention to their health. In addition, having fewer or no children in the treatment group could explain their higher probability of getting sick or injured, as in the absence of settled family life, they are more inclined to undertake dangerous jobs and lead an unstable lifestyle.

Chapter 2 investigates the impact of forced displacement due to the Three Gorges Dam construction on people's mental well-being. Most studies evaluating this project so far have been qualitative and without the support of quantitative data. Given the negative impact shown in the existing literature and the fact that many more people will be affected by dam projects in the future, understanding how forced displacement and resettlement affect people should provide valuable insights to policymakers. This chapter first identifies how forced displacement affects people in terms of depression. Nevertheless, there is no evidence to suggest it affects anxiety levels, low mood or low self-esteem. The statistically insignificant average impact on these conditions could conceal a significant effect on subgroups. I further examine the heterogeneous impact regarding educational backgrounds, marital status, whether displaced out of town, and whether displaced with the village. The results confirm that the impacts on depression, anxiety, low mood and low self-esteem differ among subgroups. Illiterate individuals were more prone to anxiety than those who had studied for at least a few months in elementary schools. Regarding depression and low mood, the divorced were not as severely affected as the married but became more anxious and had lower self-esteem. Out-of-town movers were more depressed than within-town movers. People moving with the village were more depressed than those who moved alone. These results suggest that a lack of education and unstable family life (e.g. divorce) contribute to anxiety, while long-distance moving and moving with the village contribute to depression.

Chapter 3 shifts the focus to the current COVID-19 pandemic. It explores one

of the policies in the UK – shielding. The study investigates whether people with different ethnic backgrounds (BAME and non-BAME) were affected differently by lifting the shielding policies. When shielding ended, policymakers expected people to return to their jobs. However, not all positions require a physical presence in the workplace. Since the COVID-19 pandemic was still ongoing, it was risky to meet people physically, especially for the high-risk. This is why the present study focuses on high-risk BAME individuals. It also fills a gap in the COVID-19 literature: studies either focus on BAME or high-risk, but not a combination of the two. This study employs an event study approach to capture the dynamic impact of shielding policies on people’s mental and physical well-being. The findings suggest that, first of all, after shielding was relaxed in July 2020, the impact on the diagnosis of mental ill-health is significantly higher for the high-risk BAME than their non-BAME counterparts. Second, the high-risk BAME individuals felt more at risk of catching COVID-19 than their non-BAME counterparts in November 2020 when the daily infections increased, and the shielding was not required. This significant distinction continued in January 2021, when shielding was re-implemented. However, no evidence suggests that the high-risk from a BAME background were more likely to have COVID-19 symptoms than the high-risk from a non-BAME background once the shielding requirement was relaxed. In terms of behaviour, the high-risk BAME significantly decreased their use of trains and minimised face-to-face contact with families and friends to a significantly greater extent than their non-BAME counterparts. This suggests a more acute concern about catching the virus in the high-risk BAME population. This concern could explain the insignificant difference between high-risk BAME and non-BAME in having COVID-19 symptoms; that is to say, although the high-risk BAME may live in an environment that increases the chance of getting infected compared with the high-risk non-BAME, their greater caution gave them fewer opportunities to be exposed to the virus. A further investigation of the BAME groups suggests heterogeneity between Black and Asian populations, reflecting differences in culture

and lifestyle.

Overall, this thesis demonstrates how policies can affect various subpopulations differently. It highlights the negative impacts of the policies or programmes for vulnerable groups compared with their counterparts. It also emphasises the necessity of heterogeneity analysis in policy evaluation to generate better-tailored solutions to the issues at hand.

Chapter 1

Development-induced displacement and health in China

1.1 Introduction

Since China's reform and opening-up policy in 1978, the speed of urbanisation has been accelerating. The process of urbanisation requires a large amount of land in rural areas to be converted for construction, part of which becomes urban areas afterwards. People who lived in those areas were, therefore, forced to move out of their houses and give up their land. Since their land becomes part of the urban area, their household registration status, Hukou, would be converted from rural to urban. Although the aim of development seems achieved by seeing China's rapid growth, the side effects of these projects deserve attention: the affected population in the project would normally face leaving their farmland, their home, their social connections and their livelihood. The effect of urbanisation on economic factors has been widely explored ([Tolley, 1987](#); [Tang, 1997](#); [Liu et al., 2004](#); [Wang, 2014](#)). However, little is known about the impact of development on the well-being of the people. In China, although the displaced people receive certain kinds of compensation for their loss, the stress of relocation and fear of an uncertain future cannot be overlooked. How does

land expropriation affect people's physical and mental health? How do these effects evolve over time? And what are the mechanisms that explain such effects? This chapter aims to answer these questions and quantify the impact. A large literature in refugee studies has qualitatively established that forced displacement adversely affects health ([Eisenman et al., 2003](#); [Fussell and Lowe, 2014](#); [Fazel et al., 2005](#)). [Porter and Haslam \(2005\)](#) carried out a synthesis of over 56 research reports on the relationship between forced displacement and mental health. Despite a large number of studies, the majority focused on disaster-induced displacements, such as those caused by war and natural disasters, and very few examined development-induced displacement. Although people in China who were displaced and changed their household registration (*Hukou*) due to land expropriation share similar experiences with refugees - in that both have to move out from their homes involuntarily, seek new residential places for living, build new social networks since the old ones may have broken, lose careers, face uncertainty, and so on - they differ from refugees in many aspects. For example, they obtain compensation for giving up their land; as they are rural residents and get expropriated due to urbanisation, some of them may also obtain urban *Hukou*, which enables them to access education and health care in urban areas. In short, development-induced displaced people may suffer from the experience but also benefit from it. Such opposite effects make the impact on these people more complicated to analyse than the impact on refugees.

Among the research on development-induced displacement, the literature shows various conclusions associated with different circumstances. For example, [Kasteler et al. \(1968\)](#) examined the aged population who were forcibly displaced within their local area because of highway construction and found an adverse effect on people's adjustment related to their health. [Randell \(2016\)](#) explored displacement as a result of the Belo Monte Dam construction in Brazil. Considering the heterogeneity of the destinations of those displaced households, he found that the effect on their subjective well-being was positive for those who received a higher compensation and those who

stayed closer to their origin or their acquaintances. In contrast, households who moved to urban areas experienced an adverse effect from displacement due to the extended distance from families and friends.

In terms of development project-induced displacement in China, much of the literature focuses on the Three Gorges Dam project (Hwang et al., 2007; Hwang et al., 2010; Cao et al., 2012; Xi et al., 2013). Cao et al. (2012) conducted difference-in-differences estimation to identify the effect of the Three Gorges Dam Project-Induced displacement on the displaced. They found that forced displacement adversely affected depression and self-rated health, and the mechanism was through “secondary stressors”, including “social integration”, “socioeconomic status”, and “community material resources”. Xi et al. (2013) explores heterogeneous effects on different age groups and finds that all age groups were negatively affected by displacement; the effect sizes for different age groups are similar.

In terms of the effect of displacement due to land expropriation, Wang et al. (2020) explored the labour market effect and found that those who obtained urban *Hukou* after displacement were more successful in the labour market than those who stayed in rural areas. Tani (2017), on the other hand, explored the impact on psychological well-being using the Rural Household Survey and found that obtaining urban *Hukou* would strengthen psychological well-being. However, both studies include endogenous control variables such as marital status, education level, and number of children, which potentially confound the effect of getting urban *Hukou* through expropriation. For example, Tani (2017) concludes that the impact of a change of *Hukou* status from rural to urban due to land expropriation on people’s well-being is positive and significant, while the present study finds no significant impact by excluding those endogenous control variables from the model and including the treated who have moved to urban areas in the sample.

As this chapter focuses on people’s well-being, two essential elements are mental and physical well-being. Therefore, these two elements are the main outcomes. For

mental well-being, this chapter examines a mental well-being index constructed from 12 questions of subjective well-being. This chapter examines four primary outcomes for physical health that reflect physical well-being, including self-reported health status, BMI, whether been sick or injured in the last three months and medical expenditure. Self-reported health status reflects how people feel about their physical well-being, while BMI, whether being sick or injured and medical expenditure measure their physical conditions from different angles. We hypothesize that people who were forced to move will be affected negatively due to a loss of familiar environment and neighbourhood and a change of social security system from rural to urban.¹ Their lack of experience living in urban areas, language barrier, and lower level of education would be obstacles for them to live a comfortable life. Therefore, their mental well-being could be negatively affected. Medical expenditure is expected to be affected the most for outcomes related to physical well-being due to increased consumption levels from rural to urban areas. Other physical well-being measures will also be adversely affected because people usually breathe fresher air in rural areas and are less stressed than those in urban areas.

With the same survey data as these two papers, this study focuses on the impact of forced migration and a subsequent change of *Hukou* status due to land expropriation on people's physical and mental health outcomes. By controlling for exogenous characteristics only, the parameters identified in my model capture the full effect of displacement on physical and mental health. The results show that, on average, individuals who have changed *Hukou* status from rural to urban and were displaced due to land expropriation are more likely to be sick or injured and spend more on medical expenses than those who have always held rural *Hukou* and stayed in rural areas, while the impacts on mental health, self-rated health status, and body mass

¹For example, people in urban areas are usually employed and have unemployment insurance, while people in rural areas are usually not employed but are allocated a piece of land. Medical insurance in the urban area is also in a different system from the ones in the rural area. An analysis of medical insurance is explored as mechanisms or supplementary analysis due to the smaller sample size of medical insurance.

index (BMI) are not significant.

Besides the average effects, this study also explores the evolution of treatment effects over time. The results show that, in general, the impact on physical health is negative and significant for the recently displaced compared with the longer-time displaced. For mental health, a significantly negative impact is found for those who were displaced within recent years rather than the longer-time displaced. This study also investigates other possible effects that may provide a larger picture of what is happening to complement these results - including education, number of children and other aspects of health such as medical insurance and diseases - with a smaller sample. I find that the displaced had more years of education, fewer children, and were less likely to have medical insurance. Exploring the heterogeneity by where the displaced lived, I find a negative impact on physical health for the treated who moved to urban areas but a positive impact for the treated who stayed in rural areas. Investigating the various types of medical insurance, I find more of the treated population not having medical insurance, perhaps due to the fact that the majority had lost rural cooperative medical insurance, while only a few obtained other types of medical insurance that are available for urban residents, such as public and commercial medical insurance. The policy implication of these results is that there are both costs and benefits of development-related displacement.

For the rest of this chapter, Section 1.2 introduces background information on this study, and Section 1.3 describes the data and outcome variables. Section 1.4 demonstrates the empirical framework. Section 1.5 presents the main results and discusses possible mechanisms. Section 1.6 discusses robustness. Finally, Section 1.7 concludes.

1.2 Background Information

Due to China’s administrative hierarchy, the information is recorded based on different levels of aggregates. For clarification, in the rest of this chapter, the highest level is “state”, and the next highest is “province”. Each province administers several “prefectures”. Then each prefecture contains several “counties”.

1.2.1 Urbanisation and land expropriation

In over 60 years, the urban population in China has increased from 10.64% of the country’s total population to more than 55%². However, the growth of the urban population has not been stable, as illustrated in [Figure 1.1](#). [Figure 1.1](#) also suggests that China’s urbanisation progress can be split into several stages. [Chen \(2016\)](#) split it into five periods: “start (1949 - 1957)”, “unstable development (1958 - 1965)”, “plateau (1966 - 1978)”, “high-speed development (1979 - 1991)” and “stable development (1992 -)”.

Urbanisation is closely related to land expropriation in China. First, to expand urban areas, land in the rural areas needs to be converted for urban constructions such as skyscrapers, subway systems and viaducts. Second, building the infrastructures connecting different areas across long distances, like highways or railways, also requires land expropriation at the targeted construction sites. In general, the government has divided undeveloped areas into two types: development land and undeveloped land. Undeveloped land includes forests, grassland, mountainous areas, ancient sites, and arable land. As the government controls the amount of land that can be converted to construction sites, only the development land can be used for construction projects.

²National Bureau of Statistics, China.

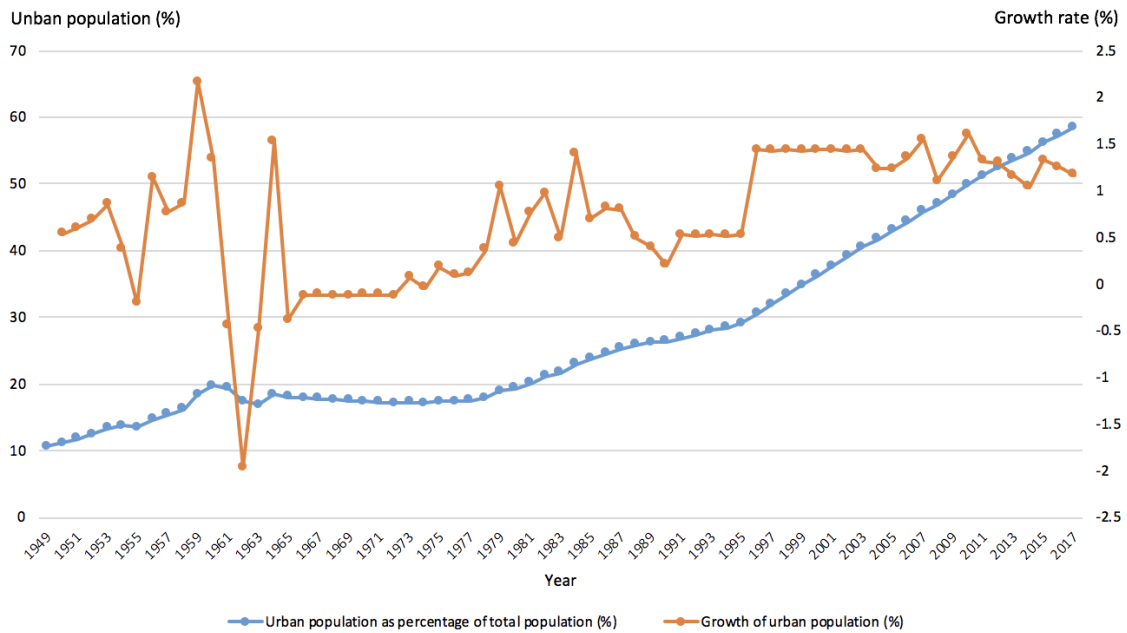


Figure 1.1: China Urbanization history
(Data source: National Bureau of Statistics, China)

1.2.2 How does the government make decisions on where to expropriate?

Displacement due to land expropriation is involuntary and can be taken as an external shock for the affected people. Still, the question remains as to whether or not the area got expropriated at random by the government. To answer this question, I researched official documents - “Urban and Rural Planning Law of the People’s Republic of China”. The process of urban planning can be summarised as follows: the State Council issues the Outline of the National Overall Planning on Land Use (Outline). Then the provincial governments make more specific planning on land use within each province according to the Outline and report their plans to get approved by the State Council. Similar procedures also apply to the lower levels of government planning. Prefecture-level plans need to be made according to provincial land use planning and get approval from their provincial government, and county-level plans

need to be made according to prefecture land use planning and get approval from their prefecture government. The only exception is the urban planning of the capital city of each province which also needs to get approval from the State Council. These urban plans usually span a time frame of 15-20 years.

Analysing provinces', prefectures' and counties' plans together, I find that the choice of areas depends on the upper level of government planning, which is subject to geographic and cultural characteristics. For example, natural resources like forests or lakes and historic sites (restricted areas) should be protected and not available for expropriation. Given that few people live in the restricted areas, no selection issue is expected for the government's decision between expropriated areas and restricted areas in my sample. However, the selection issue likely comes from the distance to cities. Areas that are close to cities are more likely to be expropriated than areas that are far away from cities. Due to a lack of information on how far away from cities the respondents were, I assume the decisions on which area to expropriate were random. The limitation of this assumption is discussed in section 1.6.

1.2.3 *Hukou*

Since China created *Hukou* in 1958, migrating from one area to another has been highly restricted. During the planned economy era in China, living necessities such as food and cloth were distributed by the central government based on *Hukou* registration. This phenomenon lasted until the 1980s, after the reform and opening-up policy was instituted, when the central government transferred its powers on *Hukou* management to local governments. Since the 1980s, the central government has introduced a series of policies regarding *Hukou* relaxation, including Blue Print *Hukou* (an intermediate or temporary *Hukou* status before getting local formal *Hukou* status, a policy started in the early 1990s and officially terminated in 2014), uniformed *Hukou* (as a replacement for the urban and rural dual household registration system, started in 2014), and a point-based system (which was initially proposed in 2011).

Each time the *Hukou* policy changed, the local governments adapted their policies accordingly. These changes mainly took place after the period covered by my analysis.³ Since *Hukou* restrictions were relaxed, growing populations have been moving from rural to urban areas and have been called “migrant workers”. Migrant workers are usually not eligible for urban *Hukou* at the beginning. Later on, depending on their circumstances, they may be able to convert their rural *Hukou* to urban *Hukou* if they choose to.

China has been focusing on industrialisation in urban areas, potentially leading to prosperity in urban areas and poverty in rural areas, since the 1950s. Although the registration system has been relaxed, especially since the 1990s, people who live in rural places are still not legally permitted to get urban *Hukou*, while people who live in urban areas are not permitted to get rural *Hukou*. Even though rural residents can work in urban areas and become rural migrants, they face a huge disadvantage due to their lack of formal urban registration (Liu, 2005); that is to say, people who do not hold urban *Hukou* are not eligible for the same rights and social security as their officially registered urban counterparts. With the aim of protecting the rights of workers, especially those who suffer from labour abuses, such as migrant workers, the Labour Contract Law became officially effective on 1 January 2008 (Jing, 2009). Although the promulgation of this law shows the progress of attention on disadvantaged workers in urban areas, the actual impact of the law on the displaced who moved to urban places is difficult to validate. While displaced individuals in the sample were interviewed in 2008 and 2009, their displacement happened earlier - over five years previously for many people. Even though the law may bring equal treatment to them, they have suffered for longer through being treated unequally. Whether they can recover one year after the implementation of the new law is still an unanswered question. Whether the law has brought equal treatment is also questionable. Zheng

³Although the Blue Print *Hukou* was terminated completely in 2014 in the city of Tianjin, after the period covered by my analysis, it has been terminated in other big cities since the early 2000s; for example, 2002 in Shanghai, 2007 in Guangzhou and Shenzhen.

(2009) argues that simply adding the workers' rights to the law does not guarantee their enforcement, which is the "real problem for the Chinese labour system". Due to the short period of promulgation and the unlikelihood of enforcing the law, this study assumes no critical effects of the Labour Contract Law on the well-being of the people.

1.2.4 Process of forced displacement and a change of *Hukou* status

From the official government documents, once the government has made the decision on where to expropriate, it sends an official notice to notify the residents of those areas about the expropriation decision, compensation and resettlement. If individuals have objections, they can apply for renegotiation. The compensation and resettlement policy varies across different expropriated areas. For example, some displaced residents were provided with resident housing elsewhere, while others received a single payoff and had the freedom to choose where to live. However, as the financial compensation may not match the purchasing power in urban areas, farmers still had difficulties surviving even with the highest level of compensation. In addition, employment opportunities and other social security were rarely provided to the displaced (He et al., 2009; Hui et al., 2013). For the rural individuals who got expropriated due to urbanisation (i.e., the areas where they lived became urban areas), their *Hukou* status would be changed from rural to urban. Urban *Hukou* allows them to receive equal rights for activities in the urban area in which they are registered; for example, certain types of jobs can only be applied for by local *Hukou* holders (Chan and Zhang, 1999). On the other hand, their loss of rural *Hukou* makes them ineligible for what they used to have in the rural areas, such as the use of land and medical insurance for rural residents.

1.3 Data

This study uses data from two waves of the [Longitudinal Survey on Rural Urban Migration in China \(RUMiC\)](#)⁴. Since RUMiC was established to study the patterns and effects of migration in China, surveys for seven waves have been conducted (2008 - 2014), with data from the first two waves publicly available. RUMiC covers interviewees from 8000 households in rural areas, 5000 households in urban areas and 5000 households who come from rural areas and temporarily reside in cities. The Rural Household Survey (RHS) was conducted in nine provinces: Anhui, Chongqing, Guangdong, Hebei, Henan, Hubei, Jiangsu, Sichuan, and Zhejiang. Migrant Household Survey (MHS) was conducted in 15 cities: Bengbu, Chengdu, Chongqing, Dongguan, Guangzhou, Hefei, Hangzhou, Luoyang, Nanjing, Ningbo, Shanghai, Shenzhen, Wuhan, Wuxi and Zhengzhou. The Urban Household Survey (UHS) was conducted in 19 cities, 15 cities are the same as the MHS, and four additional cities: Anyang, Jiande, Leshan and Mianyang.

In terms of health-related variables, RUMiC provides data on self-rated health status, weight, height, diseases and injuries, 12 questions on mental health and medical expenditures⁵. RUMiC also provides detailed information on *Hukou* status, including local urban *Hukou*, local rural *Hukou*, non-local urban *Hukou* and non-local rural *Hukou*.⁶ If the respondents hold urban *Hukou*, they will also indicate how they obtained urban *Hukou*. And those who obtained urban *Hukou* through “land expropriation” constitute my treatment group. The details of these variables are listed at

⁴The Longitudinal Survey on Rural Urban Migration in China (RUMiC) consists of three parts: the Urban Household Survey, the Rural Household Survey and the Migrant Household Survey. It was initiated by a group of researchers at the Australian National University, the University of Queensland and the Beijing Normal University and was supported by the Institute for the Study of Labor (IZA), which provides the Scientific Use Files. The financial support for RUMiC was obtained from the Australian Research Council, the Australian Agency for International Development (AusAID), the Ford Foundation, IZA and the Chinese Foundation of Social Sciences.

⁵Medical expenditures throughout this study mean all the expenses occurred due to utilisation of medical services, including the reimbursed or reduced amounts.

⁶Local *Hukou* means that the location of *Hukou* registration is the same as the location of survey respondents at the time of the interview. Non-local *Hukou* means that the location of *Hukou* registration is different from the location of survey respondents at the time of the interview.

the end of this section.

Table 1.1: Scoring coefficients from principal component factor analysis

	Mental Health Problem Index
Unable to concentrate	0.6466
Loss much sleep over worry	0.6602
Unable to play a useful part in things	0.6236
Incapable of making decisions	0.6255
Felt constantly under strain	0.6993
Unable to overcome difficulties	0.7393
Unable to enjoy day-to-day activities	0.6771
Unable to face problems	0.6585
Been feeling unhappy and depressed	0.7435
Been losing confidence in yourself	0.7529
Believe worthless	0.6993
Unhappy in general	0.6383

As the treatment is land expropriation with a *Hukou* status change from rural to urban regardless of whether they migrated to urban areas or not, the pre-treatment status for both treatment and control groups include living in rural areas and holding rural *Hukou*. The control group in my analysis is the rural residents who have always lived in rural areas and hold rural *Hukou*. The people in the treatment group are those who have changed their rural *Hukou* to urban *Hukou* through land expropriation. Therefore, all the observations in the control group are from RHS. The treatment group, on the other hand, comes from all three surveys as they may have the freedom to choose where to live after expropriation. If they move to urban areas, they will appear in UHS or MHS; while if they stay in rural areas, they will appear in RHS.

Using household heads and their spouses' data in the survey, my sample includes 27,709 observations (1,423 are displaced) in total, where 26,970 are from RHS (684 are displaced), 25 are from MHS (25 are displaced), and 714 are from UHS (714 are displaced).⁷ Therefore, these results are only supplementary. The balance between

⁷One exception for the sample size is for the mental health index, where the sample size was reduced to 19,876 in total, with 11,159 males and 8,717 females. This reduction mainly comes from the non-response due to absence as the mental health questions require respondents to be on the

the treatment and control group is shown in [Table 1.2](#). Column (3) of [Table 1.2](#) shows that the respondents in the treatment group in the full sample are, in general, comparable to the respondents in the control group in terms of gender, age, birth order, parents' education and parents' age, although the treatment group is slightly younger, slightly higher in their birth order and slightly lower in their parents' age compared with the control group.

Table 1.2: Sample balance

	Treatment: group mean	Control group mean	Difference: treatment vs. control
	(1)	(2)	(3)
Male (0 or 1)	0.476 (0.500)	0.501 (0.500)	-0.011 (0.018)
Age	47 (10.379)	49 (9.792)	-1.058*** (0.292)
Birth order	2.546 (1.519)	2.394 (1.440)	0.205*** (0.058)
Highest level of education among parents	3.026 (2.583)	2.140 (3.089)	0.146 (0.086)
Average age of parents	75.702 (11.279)	78.278 (10.598)	-0.848** (0.286)
Sample size	1,423	26,286	27,709

Notes to table: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. The results in columns "Differences: treatment vs. control" are obtained by regressing the variables of interest on treatment status and all other control variables. Besides the variables shown in the table, it also includes year and ethnicity.

It is worth mentioning that although people in the treatment group have obtained urban *Hukou* after displacement, close to half are still from RHS, meaning they are still living in rural areas. The reasons why these people still stay in rural areas are

scene. However, several individuals did not respond to mental health questions even though they were on the scene. To investigate whether these missing respondents are due to sample selection, I estimate the mental health model using the two-step Heckman correction method ([Heckman, 1979](#)). [Table 1A.2](#) in [Appendix 1A](#) displays the results. The insignificant coefficients on the inverse Mills ratio and the similar results for OLS and the selection model suggest no selection in response to mental health interview questions. However, without a valid exclusion restriction, the model can only be identified when the inverse Mills ratio is non-linear, which might not be the case for this dataset.

not available in the data. I interviewed some villagers in China, and they provided me with two possible reasons: 1. they face difficulty in finding a job in the urban areas, thus go to work with their relatives who are still in the rural areas; 2. senior people whose land got expropriated live with their adult children who live in rural areas.

Health-related Outcomes

Mental health: this is a mental health index combining 12 mental health questions by the principal component factor analysis method. The higher the index, the worse the mental health condition. Compared with the index built by simply adding up each component ([Tani, 2017](#)), this study uses principal component factor analysis to take the correlation between different components into account (see [Table 1.1](#) for the regression scoring coefficient of each component).

BMI scale: based on Asian-Pacific standard BMI classification ([WHO expert consultation, 2004](#)), each individual is classified as one of four categories: 0, underweight ($BMI < 18.5$); 1, normal ($18.5 \leq BMI < 23$); 2, overweight ($23 \leq BMI < 27.5$); or 3, obese ($BMI \geq 27.5$).

Self-rated health status: each individual self-reported their health status as one of the five categories: 1, excellent; 2, good; 3, average; 4, poor; 5, very poor. In [Table 1.3](#), “poor health” equals 1 if self-rated health status equals poor or very poor, and “poor health” equals 0 if self-rated health status equals average, good or excellent.

Medical expenditure: individuals reported their medical expenses in cash in 2007 (or 2008, depending on the survey year), including reimbursements or reduced amounts.

Whether sick or injured: individuals reported whether they had been sick or injured (including chronic or acute disease) in the last three months. Responses were recorded as 1 if the answer was yes, and 0 if their answer was no.

These outcomes are used to draw a complete picture of the well-being of people from different angles. [Table 1A.1](#) illustrates the correlation among these outcome variables (with BMI being numbers rather than categories): as expected, medical expenditure is positively correlated with whether being sick or injured (0.202), poor health (0.316) and mental health problem index (0.177); poor health is positively correlated with mental health problem index (0.277) and whether being sick or injured (0.195); and whether being sick or injured is also positively correlated with mental health problem index (0.114). However, as these correlations are weak, and BMI is basically not correlated with any other variables, not much information is missing when these outcomes are treated separately.

1.4 Empirical Framework

My aim is to estimate the difference in health outcomes for two individuals who were identical in all characteristics and lived in the same area before, then one got expropriated and one did not. I do not observe where individuals lived before expropriation, and need to assume that individuals do not change prefecture after the expropriation (as “prefecture” is the finest geographical unit observed in the data). I also assume that the decision on which areas to expropriate within the prefecture is as good as random. As discussed, areas closer to the cities are more likely to be expropriated, I need to assume that there are no systematic differences between individuals by the distance from the city as the distance to the city cannot be controlled due to data limitation. This assumption may not hold because of the reason mentioned above. This limitation is elaborated on in Section 1.6. [Table 1.2](#) shows that although differ-

Table 1.3: Health Outcomes

	Male	Female	All
Mental Health Problem Index	0.034 (0.059)	0.050 (0.059)	0.039 (0.041)
Poor health (0 or 1)	0.024 (0.016)	0.001 (0.015)	0.013 (0.011)
BMI	0.124 (0.158)	0.267 (0.234)	0.195 (0.142)
If sick or injured (in the last three months, 0 or 1)	0.022 (0.017)	0.020 (0.017)	0.021* (0.012)
Medical Expenditure	340.136 (188.962)	189.908 (99.428)	269.213* (105.652)
Sample size	13,857	13,852	27,709

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the individual level are shown in parentheses. BMI is a continuous variable calculated using height and weight. “Poor health” is a binary variable coded from self-reported health status. As poor health and if sick or injured are estimated using the random effects probit method, marginal effects are reported for these two outcomes. For the rest of the outcomes, equation (1.1) is estimated as linear random-effects regression and the reported figures are coefficients. Each coefficient/marginal effect is from a separate regression. The number of observations for mental health is different, with 11,159 male, 8,717 female and 19,876 in total.

ences in age, birth order and the average age of parents between the treatment group and control group are significant, these differences are small.

As displacement is likely to affect people's well-being through a change of income, ideally, the income should be controlled. However, as income is not available for some people in the dataset, we control for education instead. Education often acts as a shortcut to raising one's social status and has certainly played an essential role in transforming lifestyle from rural to urban. People who are initially from a rural area are more likely to be respected and secure a well-paid job if they have a higher level of education than those who do not. Another fundamental question is whether there are significant differences in the family background for those whose land got expropriated before they moved. However, this could not be answered directly from the dataset, as the survey was conducted ex-post in most households. Hence, parents' education and age acted as proxies for family background. Considering that the older the child in the family, the more responsibility and expectation the child will take, the analysis also controls for the birth order because an older child is usually more mature and better at handling issues.

The approach for modelling the effect of land expropriation with a change in *Hukou* status (EXP_{itp}) on health-related outcomes (H_{itp}) of individual i living in prefecture p in year t is regression with prefecture fixed-effects. The regression model is as below:

$$H_{itp} = \alpha_0 + X_{itp}\alpha_1 + \alpha_2 EXP_{itp} + \delta_t + \mu_p + \nu_{itp}, t = 2008, 2009 \quad (1.1)$$

where H_{itp} represents health-related outcomes of interest, including mental health, BMI, medical expenditure, self-rated health status and whether the individual is sick or injured in the last three months. X_{itp} is a vector of exogenous individual characteristics, including age, gender, birth order, ethnicity, parent's education and age. EXP_{itp} is a binary variable indicating whether the individual i , who was interviewed at prefecture p , has ever obtained urban *Hukou* through land expropriation by t or

not. μ_p represents prefecture effects for the prefecture where respondents were surveyed.⁸ And δ_t denotes year fixed-effects. The term $\nu_{itp} = c_i + u_{itp}$ is the composite error, which is the sum of the unobserved time-invariant individual-specific effect c_i and idiosyncratic error u_{itp} .

Given the different features of health-related outcomes, the estimation method for equation (1.1) varies. Mental health and medical expenditure are continuous variables, so **random-effects** is used for estimation. Weight categories are ordered categories where the boundaries for each category are known, so the estimation method is **random-effects interval regression**. For whether the individual is sick or injured or has poor health, equation (1.1) is estimated by **random-effects probit**. For self-rated health status, as the categories are ordered while the boundaries for each interval are unknown, equation (1.1) is estimated by **random-effects ordered probit**. For all these regressions, standard errors are clustered at the individual level.

In my sample, as only two individuals changed their treatment status from 0 in 2008 to 1 in 2009, there is not enough variation over time in *Hukou* status to use fixed-effects estimates. Therefore, all the regressions are estimated by random effects, assuming time-invariant individual-specific unobservable heterogeneity is uncorrelated with all other regressors. However, this assumption is quite strong and may not hold. This limitation is elaborated on in Section 1.6.

1.5 Results

1.5.1 Main results

Estimates of α_2 from equation (1.1) are reported for males, females and the two genders pooled together in [Table 1.3](#). As “poor health” and “if sick or injured” are estimated by random-effects probit, marginal effects are reported in [Table 1.3](#), instead of coefficients. For the rest of the outcomes in [Table 1.3](#), equation (1.1) is estimated

⁸All respondents were in the same prefecture in 2008 and 2009.

as linear random-effects regression and reported figures are coefficients.

Table 1.3 shows that the impact on whether the individuals had been sick or injured in the last three months is significant: displaced individuals were 2.1 percentage points more likely to report being injured or sick; no difference is found by respondents' gender. This is in line with the hypothesis. The possible explanation could be that people who obtained urban *Hukou* through land expropriation tended to enter labour-intensive industries, such as construction, in which they may get injured due to their unfamiliarity with the job.

Medical expenditure was significantly affected as expected; it increased by 269.21 yuan per year on average, equivalent to 55% of the mean medical expenditure for both genders pooled together. The increase is significant at the 10% significance level. This result makes intuitive sense because medical expenses are generally higher in urban areas than in rural areas. The estimated impact on medical expenditure for males is almost as twice that for females (340.14 yuan for males and 189.91 yuan for females); however, the impact for either gender is insignificant.

Increasing BMI does not directly indicate a better or worse health status, so I also estimate the effects of displacement on weight categories by random-effects interval regression. The results in Table 1.4, however, show no significant impact on weight categories for either gender or both genders pooled together. Table 1.5 displays the random-effects ordered probit estimates on self-rated health status. No significant impact is found for either gender or both genders pooled together. These results differ from the hypothesis. The explanation could be that the impact on health-related outcomes, if there is any, usually takes time to become evident and sometimes could take decades.

Given the main results, three questions are worth exploring. First, are the insignificant results due to the aggregation of treatment effects over time? Second, what are the mechanisms through which the treated population are more likely to be sick or injured? Third, are the insignificant results due to aggregation across heterogeneous

Table 1.4: Weight Categories

	Male	Female	All
Weight Categories			
Underweight (< 18.5)	-0.007 (0.005)	-0.003 (0.009)	-0.005 (0.005)
Normal ($\geq 18.5, < 23$)	-0.022 (0.019)	-0.005 (0.014)	-0.012 (0.012)
Overweight ($\geq 23, < 27.5$)	0.024 (0.020)	0.007 (0.020)	0.015 (0.014)
Obese (≥ 27.5)	0.005 (0.005)	0.001 (0.003)	0.003 (0.003)
Sample size	13,857	13,852	27,709

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the individual level are shown in parentheses. Each column (columns 2-4) is from one regression. Column 2 uses male data only, column 3 uses female data only and column 4 uses male and female data together. All regressions are estimated using random effects interval regression. Marginal effects are reported for each BMI category.

Table 1.5: Self-reported Health Status

	Male	Female	All
Health Status Categories			
Excellent	-0.016 (0.016)	-0.006 (0.014)	-0.011 (0.011)
Good	-0.003 (0.004)	-0.003 (0.006)	-0.003 (0.003)
Average	0.013 (0.014)	0.005 (0.012)	0.009 (0.009)
Poor	0.005 (0.005)	0.003 (0.006)	0.004 (0.004)
Very poor	0.001 (0.001)	0.001 (0.001)	0.001 (0.001)
Sample size	13,857	13,852	27,709

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the individual level are shown in parentheses. Each column (columns 2-4) is from one regression. Column 2 uses male data only, column 3 uses female data only, and column 4 uses male and female data together. All regressions are estimated using random effects ordered probit. Marginal effects are reported for each health status category.

groups? The following sections address these questions.

1.5.2 Supplementary results: time since displacement

To explore the first question, equation (1.2) is estimated with six dummies (0-3, 3-6, 6-9, 9-12, 12-20, over 20) which indicate the years since displacement, instead of the single treatment indicator EXP_{itp} in equation (1.1):

$$\begin{aligned}
H_{itp} = & \alpha_0 + X_{itp}\alpha_1 + \beta_1 EXP0_3_{itp} + \beta_2 EXP3_6_{itp} + \beta_3 EXP6_9_{itp} \\
& + \beta_4 EXP9_12_{itp} + \beta_5 EXP12_20_{itp} + \beta_6 EXP20_∞_{itp} + \alpha_3 t + \alpha_4 \mu_p + \nu_{itp}
\end{aligned}
\tag{1.2}$$

where $t = 2008, 2009$. where H_{itp} represents health-related outcomes of interest, including mental health, BMI, medical expenditure, self-rated health status and whether the individual has been sick or injured in the last three months. X_{itp} is a vector of exogenous individual characteristics, including age, gender, birth order, ethnicity, parent's education and parent's age. EXP_{y-z}_{itp} indicates whether or not individual i living in p has been displaced y to z years ago due to land expropriation for $y \in \{0, 3, 6, 9, 12, 20\}$ and $z \in \{3, 6, 9, 12, 20, \infty\}$. [Table 1.6](#) shows the results. For each health-related outcome, the effect for those who were displaced a long time ago is opposite to the effect for those who were displaced more recently. The impact on mental health is significantly positive for those displaced over 20 years ago, while the impact is significantly negative for those displaced in the past three to six years and nine to twelve years. Conversely, the impacts on the self-rated health status are significantly negative for those who were displaced over 20 years ago but positive for the recently displaced. The impacts on the probability of being physically sick or injured are similar. These findings suggest that the forced displacement with *Hukou* change has a relatively short-term negative impact on an individual's mental health, which may recover after adaptation to the new life, even when the physical health

status worsens. Following the literature on adaptation ([Baji and Biro, 2018](#); [Cubí-Mollá et al., 2017](#)), this could be explained by adaptation to the new circumstances over time (the migrant would only compare their life with a recent time reference, but not the time reference before they moved). Although land expropriation is not a health shock and does not share specific characteristics with health shocks, such as shortened life expectancy, it shares the characteristics of being adverse, unpredictable and a massive disruption of life. [Cubi-Molla et al. \(2017\)](#) find the adaptation effect only when the duration is over 20 years since the shock, which explains the impact of land expropriation on mental well-being in [Table 1.6](#). [Gilbert \(2006\)](#) also points out that the human brain is more sensitive to the “relative magnitude of stimulation” rather than the “absolute magnitude of stimulation”, which reflects adaption over time after the shock. It also points out that evaluating things could be subjective as it depends on the comparisons. This implies that even objective measures, such as self-reported health measures, have a solid subjective well-being component.

The effect on BMI is more consistent: being displaced increases BMI in general. [Table 1.7](#) shows a similar pattern as [Table 1.4](#): being displaced and changing to urban *Hukou* lowers the probability of being underweight or normal weight while increasing the probability of being overweight or obese. The effects are significant for those who were displaced 3-6 years ago and 9-12 years ago. This impact on weight categories may suggest the difference between urban and rural lifestyles. In terms of medical expenditure, people who were displaced six to nine years ago and over 20 years ago had significantly higher medical expenditures than the control group. However, the impact on medical expenditure is more complicated to explain. A higher expenditure may imply worse health conditions, better health consciousness, or better medical insurance⁹.

⁹As medical expenditure here includes the reimbursed or reduced amounts, medical expenditure reflects the utilisation of medical services. Better medical insurance can lead to higher utilisation of medical services.

1.5.3 Mechanisms

To explain the main effects, this section investigates some mechanisms that could potentially affect health. [Table 1.8](#) displays the effects of displacement on medical insurance the individuals hold, replacing health-related outcomes with medical insurance outcomes in equation (1.1). The binary outcomes are estimated by random-effects probit.¹⁰ The amount paid for medical insurance is estimated by linear random-effects regression. The results show that displacement with a change in *Hukou* status significantly lowers the probability for individuals to have medical insurance. This could be explained by the lower probability of getting rural cooperative medical insurance, as rural cooperative medical insurance mainly targets rural residents. Although the probability of obtaining public medical insurance and commercial medical insurance is higher, in absolute value, the increase is smaller than the decrease in the probability of having rural insurance. One reason for not getting urban medical insurance could be that the amount paid for urban medical insurance is significantly higher on average than rural cooperative medical insurance, which is in line with the fact that rural cooperative medical insurance is much cheaper than other medical insurance. This suggests that a smaller proportion of people have medical insurance because they cannot afford it after changing *Hukou*.

The main results in section 5.1 show that the displaced were more likely to be sick or injured in the last three months of the survey. To understand this result, this section investigates specific kinds of diseases which displaced people may have. The results show that the treated were more likely to have infectious diseases, diabetes, hypertension and cerebrovascular disease, which are closely related to stress level and living habits (see [Table 1.9](#) for details).

¹⁰The outcome of types of medical insurance cannot be coded as one categorical variable as the survey respondents may hold one or more types of medical insurance.

Table 1.6: Health outcomes with time interval since displacement

	Mental Health Index	Health Prob-	Poor Health (0 or 1)	BMI	If sick (in the last three months, 0 or 1)	Medical Expenditure	Sample size (EXP=1)	Sample size (EXP=0)
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
0-3 years		0.047 (0.113)	-0.019 (0.021)	0.089 (0.389)	-0.007 (0.029)	-208.449 (302.714)	115	27,594
3-6 years		0.165* (0.066)	-0.026** (0.013)	0.107 (0.230)	-0.019 (0.017)	-66.745 (190.191)	387	27,322
6-9 years		0.003 (0.061)	0.039** (0.019)	0.378 (0.208)	0.019 (0.019)	488.317** (172.744)	334	27,375
9-12 years		0.173* (0.085)	-0.015 (0.018)	0.322 (0.269)	0.018 (0.029)	-5.999 (227.991)	192	27,517
12-20 years		-0.092 (0.079)	0.001 (0.018)	-0.135 (0.254)	0.04 (0.025)	165.028 (202.699)	258	27,451
over 20 years		-0.275* (0.112)	0.057* (0.032)	0.621 (0.352)	0.079** (0.035)	686.022* (269.978)	178	27,531

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the individual level are shown in parentheses. Marginal effects are reported for each health outcome of interest. Each marginal effect is from one regression estimated. The estimation method in this for each health outcome is the same as that in Table 3.

Table 1.7: Weight categories with time interval since displacement

	Underweight (< 18.5) (1)	Normal ($\geq 18.5, < 23$) (2)	Overweight ($\geq 23, < 27.5$) (3)	Obese (≥ 27.5) (4)	Sample size (EXP=1) (5)	Sample size (EXP=0) (6)
0-3 years	-0.015 (0.012)	-0.038 (0.035)	0.044 (0.038)	0.009 (0.009)	115	27,594
3-6 years	-0.015** (0.007)	-0.038* (0.020)	0.044** (0.022)	0.009* (0.005)	387	27,322
6-9 years	0.01 (0.009)	0.019 (0.015)	-0.025 (0.021)	-0.004 (0.003)	334	27,375
9-12 years	-0.017** (0.008)	-0.043* (0.024)	0.051* (0.026)	0.010* (0.006)	192	27,517
12-20 years	0.006 (0.010)	0.011 (0.019)	-0.015 (0.025)	-0.002 (0.004)	258	27,451
over 20 years	-0.019* (0.010)	-0.048 (0.032)	0.055 (0.034)	0.011 (0.008)	178	27,531

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the individual level are shown in parentheses. Marginal effects are reported for each weight category. Each line is from one regression estimated. All regressions are estimated using random effects interval regression. Marginal effects are reported.

Table 1.8: Medical Insurance

	Male	Sample size	Female	Sample size	All	Sample size
Medical Insurance (0 or 1)	-0.078*** (0.019)	9,672	-0.081*** (0.018)	9,480	-0.075*** (0.012)	20,292
<i>Rural cooperative medical insurance (0 or 1)</i>	-0.256*** (0.024)	11,366	-0.279*** (0.023)	10,937	-0.258 *** (0.016)	23,212
<i>Public medical insurance (paid for by the government) (0 or 1)</i>	0.035*** (0.013)	7,041	0.025*** (0.010)	7,295	0.028*** (0.007)	16,183
<i>Commercial medical insurance (0 or 1)</i>	0.062*** (0.019)	3,551	- -	-	0.047*** (0.010)	9,280
Amount paid for medical insurance in 2007/2008 (including amounts deducted from payroll)	220.955*** (16.697)	13754	303.500*** (18.884)	13729	263.151*** (12.643)	27,483

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the individual level are shown in parentheses. Marginal effects are reported for each outcome and are shown with gender separation and together. Each marginal effect is from one regression. The estimation method for medical insurance, rural cooperative medical insurance, public medical insurance and commercial medical insurance are estimated with random effects probit, while the amount paid for medical insurance is estimated with random effects.

The literature suggests that education level and number of children may affect health in indirect ways ([Brunello et al., 2013](#); [Chow and Chen, 1994](#)), so I also explore these two factors. The results in [Table 1.10](#) show that the displaced people had more years of education but fewer children compared with the control group, and the effects are significant at the 5% level. The gender differences for these effects are small, but the effects on females are higher for education while lower for number of children.

1.5.4 Heterogeneous effects: living in urban areas versus rural areas

Although the displaced individuals in the sample have changed their *Hukou* status from rural to urban, only half of them actually moved to urban places (urban movers), and the rest stayed in rural places (rural stayers). These two groups may experience very different lives after displacement as urban movers may face more assimilation issues when they try to edge into urban society. For example, [Wong et al. \(2007\)](#) argue that migrant workers receive unequal treatment from urban natives and are likely to suffer from mental ill-health as a result of marginalisation. In order to address this potential heterogeneity, this section compares urban movers and rural stayers separately with the control group. The results are shown in [Appendix 1B.a](#) for urban movers and [Appendix 1B.b](#) for rural stayers.

The results show that urban movers are more likely to be sick or injured for both genders, while female rural stayers are less likely to be sick or injured. In terms of self-rated health status, urban movers are less likely to report good or excellent health and more likely to report average, poor or very poor health. On the contrary, rural stayers are more likely to report excellent health or good health for both genders, and less likely to report average, poor or very poor health. These results suggest that post-displacement life played an important part in the physical health of the displaced people and that urban life could be detrimental to their physical health.

Table 1.9: Most recent disease or injury in last 3 months

	Male	Female	All
infectious	0.010 (0.020)	0.034* (0.014)	0.024* (0.012)
parasitic	-0.002 (0.011)	0.000 (0.010)	-0.001 (0.007)
cancer	0.016 (0.014)	-0.005 (0.014)	0.007 (0.010)
benign	0.023 (0.018)	-0.060** (0.021)	-0.021 (0.014)
diabetes	0.051* (0.022)	0.032 (0.026)	0.040* (0.017)
blood	0.011 (0.010)	0.004 (0.016)	0.007 (0.010)
mental	0.000 (0.010)	-0.025 (0.019)	-0.014 (0.011)
eye	0.006 (0.017)	-0.008 (0.013)	-0.001 (0.011)
ear	-0.011 (0.013)	0.009 (0.013)	0.001 (0.009)
heart	0.038 (0.031)	-0.024 (0.028)	0.006 (0.021)
hypertension	0.091** (0.034)	0.067* (0.033)	0.076** (0.024)
cerebrovascular	0.014 (0.023)	0.066** (0.024)	0.043* (0.017)
bronchitis	-0.026 (0.035)	0.011 (0.030)	-0.007 (0.023)
gastritis	0.005 (0.044)	0.020 (0.040)	0.010 (0.029)
cirrhosis	-0.024 (0.015)	-0.003 (0.008)	-0.013 (0.008)
gallbladder	-0.006 (0.016)	-0.003 (0.019)	-0.007 (0.013)
genitourinary	0.000 (0.019)	0.029 (0.021)	0.018 (0.014)
skin	-0.014 (0.017)	0.025 (0.015)	0.008 (0.011)
arthritis	-0.007 (0.038)	-0.004 (0.042)	-0.003 (0.028)
perinatal	-0.002 (0.005)	-0.001 (0.006)	-0.001 (0.004)
injurpoison	0.027 (0.035)	-0.021 (0.023)	0.005 (0.020)
Sample size	1,428	1,721	3,149

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the individual level are shown in parentheses. Each result is from a separate regression. The estimation method is a linear probability model with random effects.

Table 1.10: Education and children

	Male	Sample size	Female	Sample size	All	Sample size
Years of Education	0.860*** (0.144)	13,770	0.901*** (0.146)	13,714	0.907*** (0.106)	27,484
Number of Children	-0.179*** (0.052)	13,634	-0.148** (0.046)	13,817	-0.164*** (0.035)	27,451

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the individual level are shown in parentheses. Each coefficient is from one regression. Years of education and number of children are both estimated using random effects.

In relation to medical expenditure, urban movers spent significantly more, while no significant effect is found for rural stayers. This result could be attributed to several factors, such as higher price levels and poorer air quality in urban areas. For mental health, no significant impact at the 5% level is found. As being either an urban mover or a rural stayer can be endogenous, this analysis is only descriptive and the results should be interpreted accordingly.

1.6 Robustness

This section demonstrates several robustness checks of the results and discusses the limitations of this study at the end of the section. For robustness checks, I first use the Mahalanobis-metric matching method to find a more comparable control group. Then, results with various dis-aggregations in years since treatment are examined to check the robustness of treatment effects over time.

1.6.1 Matching

As the size of the control group is much larger than the size of the treatment group in the sample (26,286 vs. 1,423), it is likely that observations in the control group are not all comparable to the observations in the treatment group. To improve the results from this case, I use Mahalanobis-metric matching (Rubin, 1979) to find a closely matched control group for the treatment group in terms of gender, age, birth order, parents' age and parents' education. The matching is conducted for each year, 2008

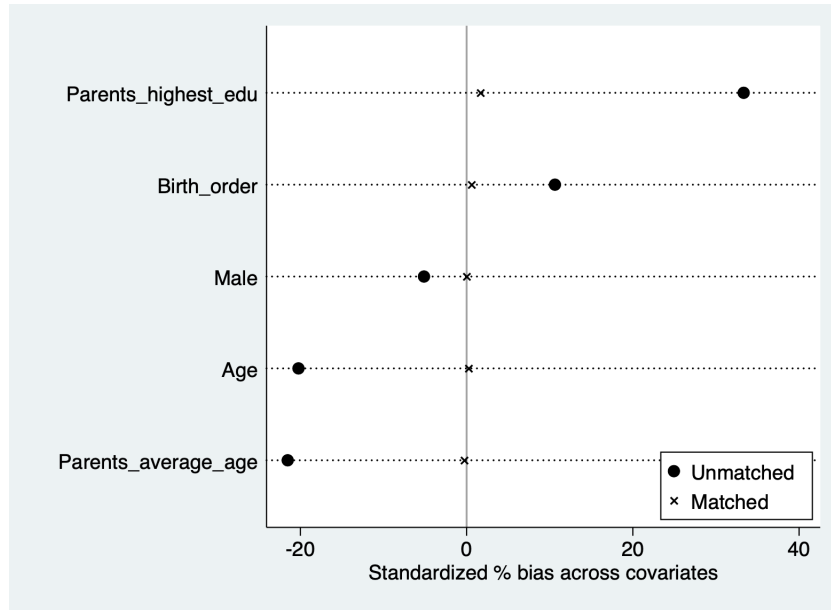


Figure 1.2: Matched sample versus unmatched sample, 2008

and 2009, separately. The matching quality before and after matching is displayed in [Figure 1.2](#) and [Figure 1.3](#). We can see that the matched sample is much more comparable to the unmatched sample in terms of the observables mentioned above. The bias reduction is over 90% for every control variable (see details in [Appendix 1C](#)).

Overall, the results based on the matched control sample are qualitatively similar to the baseline results, although there are differences in point estimates. I find a significant increase in medical expenditures with the matched control sample than the full sample. Also for the matched sample with both genders together, the impact on BMI becomes significant at the 10% level, with individuals in the treatment group less likely to be underweight or normal weight, but more likely to be overweight and obese, while the marginal effects are not significant for the full sample. In terms of medical insurance, the direction of the marginal effects in the matched sample is in line with those in the full sample, but the magnitude is greater.

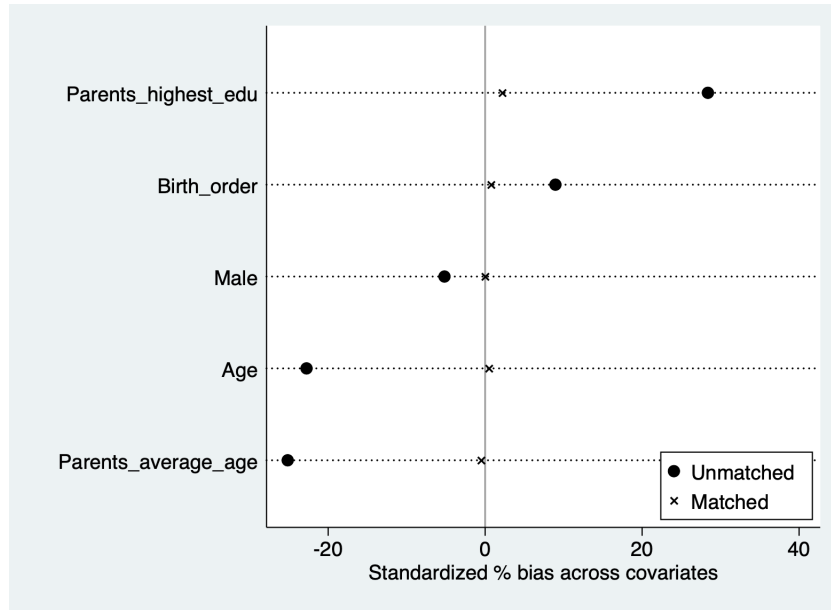


Figure 1.3: Matched sample versus unmatched sample, 2009

1.6.2 Treatment effects over time with different time intervals

To check the robustness of treatment effects over time, four different sets of time intervals are examined for the five main health-related outcomes¹¹ (see [Appendix 1D](#)). Similar to the results of time intervals in section 1.5.2, no changes over time are found for the medical expenditure and weight categories, and people were more likely to be sick or injured if they were long-term displaced. A negative impact on mental health is found for people who were displaced within six years. For the self-rated health status, the effect over time is not significant.

The rest of this section discusses some limitations of this study. As the identification strategy relies on spatial variation of land expropriation, it requires matching

¹¹First set: the individual has been displaced 0-4, 4-8, 8-12, 12-16, 16-20 and over 20 years ago. Second set: the individual has been displaced 0-6, 6-9, 9-20 and over 20 years ago. Third set: the individual has been displaced 0-5, 5-10, 10-15, 15-20 and over 20 years ago. Fourth set: the individual has been displaced 0-5, 5-10, 10-15 and over 15 years ago.

of project-induced expropriation to surveyed interviewees by location. Several issues are worth mentioning here.

i. Pre-treatment locations are unavailable

In terms of location, there are two issues. First, only the prefectures where households are located at the time of the survey are observed, which, for the treatment group, may differ from the prefectures where their land got expropriated. By controlling for prefecture effects, this study assumes that displaced people are from the same prefecture as the one indicated in the survey; in other words, they are displaced within the prefecture. However, this may not be the case. According to the China 2005 mini-census, almost one-third of people who were involuntarily displaced moved out of the prefecture, although (unlike people in my treatment group) these people did not change their *Hukou* status. The issue with some people changing prefectures is that the prefecture fixed-effects observed at the time of the survey do not fully control for the fact that the land in some prefectures may be more likely to be expropriated and individuals from these prefectures may have characteristics associated with the health outcomes.

Second, selection into location by households based on characteristics that are correlated with health outcomes potentially introduces selection bias. For example, depending on the compensation package they obtained, displaced people may have had the freedom to choose where to move to and may have selected a destination with better health care services or less pollution.

ii. Treatment may be non-random

The treatment is land expropriation with a change in *Hukou* status from rural to urban. This implies a displacement for the treated individuals. This displacement, depending on the local policy, is not always arranged by local authorities; that is, people may have the freedom to choose where they move to. As mentioned above, the

decision about which areas get expropriated is made by the government. Therefore, the selection is systematic and individuals have no power to manipulate the decision. Thus individuals cannot select themselves for the treatment. However, the decision by the government is based on specific features of the areas and may not be random. For example, the planner may choose to retain the land that is most productive for agricultural use or choose to preserve natural amenities or choose to redevelop places with poor prospects. Therefore the expropriated areas (those where the treatment group lives) may have some characteristics that do not match those of more remote rural areas where some in the control group live. For example, individuals who get displaced could be those who were living closer to the city centre - a more educated area with more job opportunities - and these areas may attract younger people rather than older ones. Controlling for age and parents' education attenuates these concerns, however, it cannot eliminate them. Moreover, the assumption that no systematic differences exist between treatment and control groups by the distance from the city may not hold. This assumption is made because the data on the pre-moving distance between people's houses and the city is unavailable in the dataset.

iii. Pre-treatment outcomes are not available

The pre-displacement health-related statuses are not controlled for due to data availability. For example, individuals' mental health status, BMI, medical expenditure and self-rated health status are unknown for the pre-displacement period. The inability to control for these pre-treatment outcomes could potentially confound the pre-treatment difference with the impact of treatment.

iv. The random effect assumption may not hold

By assuming time-invariant individual-specific unobservable heterogeneities are uncorrelated with all other regressors, all the models in this chapter are estimated by random effects. However, it is a strong assumption because certain unobserved het-

erogeneity might be correlated with some regressors. For example, personal history of addiction could be correlated with parents' education: a higher parents' education implies higher family income and social status, which makes it easier to get access to drugs but might make people less likely to consume them.

v. Heterogenous treatments

As the policy varies across prefectures, displaced people may receive different packages of compensation together with a change of *Hukou* status: some may get a compensation payoff, some may get monthly compensation, and some may get compensation with a job settlement. As information regarding the package received is unavailable, the effect of forced displacement with a change in *Hukou* status cannot be separated from the effect of getting specific compensation packages.

vi. Respondents from the control group may also have been displaced

Observations selected into the treatment group were based on one question in the questionnaires of the RUMiC survey ("How did you obtain the urban *Hukou*?"). If the answer is "land expropriation", then the observation was placed in the treatment group. For those in the control group who have always been rural *Hukou* holders, however, whether they have experienced forced displacement within rural areas or not cannot be observed. In other words, rural-to-rural forced displacement due to land expropriation is not observable from the data. However, the cases of displacement within rural places are much fewer than those from rural to urban.

vii. Multiple hypothesis testing

One concern of this study is the problem of multiple hypothesis testing. For example, in [Table 1.3](#), five different health measures are used. If we test each hypothesis separately, the probability of having false rejections is $1 - 0.95^5 = 19\%$. [Table 1E.1](#) compares the p-values adjusted for multiple hypothesis testing, the Westfall-Young

adjusted p-values, with unadjusted p-values.([Westfall et al., 1993](#)). From this table, certain differences exist between the Westfall-Young adjusted p-values and the unadjusted p-values. The impact on whether being sick or injured is statistically significant at a 5% level with Westfall-Young adjusted p-value while at a 10% level with the unadjusted p-value. The impact on medical expenditure is statistically significant at 1% level with Westfall-Young adjusted p-values while at 5% level with unadjusted p-values. For other outcomes, the impacts remain statistically insignificant with both p-values.

Besides the main analysis, further analysis of breaking down the periods since land expropriation and looking at different health conditions are more likely to have the problem of multiple hypothesis testing. For example, in [Table 1.6](#), five different health measures are used, interacting with six different periods; in [Table 1.9](#), 21 different medical conditions are reported, so the probability of having false rejections is so high that t-statistics are meaningless for statistical inference. Therefore, these analyses are supplementary and descriptive.

Unlike regressions with a single treatment period, regressions with multiple periods, as in [Table 1.6](#), have a much higher probability of false rejection. One solution would be to reduce the number of periods. For example, [Table 1E.2](#) uses only two periods: land expropriation happened within eight years or more than eight years. Results show that those whose land was expropriated were more likely to be sick or injured if they were expropriated eight years ago. It is worth mentioning that although the probability of false rejection on hypothesis testing with these two periods becomes lower than that with six periods, it is higher than that in [Table 1.3](#).

1.7 Conclusion

This chapter studies the health impacts on household heads and their spouses who have obtained urban *Hukou* through land expropriation and the possible explanations

for these effects. I find that, compared with those who have remained in rural areas with rural *Hukou*, people who have experienced forced displacement with a change in *Hukou* status are more likely to be sick or injured and have higher medical expenditures. These results may reflect that the treatment group are likely to obtain work in industries they are not familiar with, such as construction, where they suffer injuries. People from the treatment group incurred higher medical expenditure, most likely due to a combination of a higher price level of healthcare service and a higher medical service utilisation as a result of poor air quality and stressful lifestyles in urban areas. No significant impact is found on mental health, self-rated health status or BMI. By exploring the effects of land expropriation with a change in *Hukou* status on education, number of children and medical insurance, this study also finds the treated have more years of education and fewer children and are less likely to have medical insurance due to their ineligibility for rural medical insurance and a relatively low probability of obtaining other types of medical insurance.

As the insignificance of the impact on mental health and some measures of physical health may come from the aggregation of treatment effects over time and across heterogeneous groups, this study explores disaggregation of the treatment effects based on time intervals since displacement and whether the treated live in urban areas. I find a significantly negative impact on the mental health of the treated who were displaced within recent years, but no significant impact is found on people who were displaced a long time ago. The impact on BMI is consistent over time: the treated are more likely to be overweight or obese and less likely to be underweight or normal. No significant time effects on self-rated health status are found. By separating the treated into urban movers and rural stayers, I find a negative impact on physical health for the urban movers but a positive impact for the rural stayers.

The significantly lower probability of holding medical insurance for the treated suggests incomplete social security for the displaced people, even though they have urban *Hukou*. More years of education suggests a better environment of education,

which could explain the higher medical expenditure as individuals obtaining higher levels of education usually pay more attention to their health. The lower number of children suggests two possible points: 1. the one-child policy has been implemented more strictly in urban areas; 2. the displaced people did not live a settled family life, and thus they were more likely to undertake dangerous jobs and have an unstable lifestyle which made it easier to get sick or injured.

The main limitation of this study is that pre-treatment locations are not available, and thus the model cannot control for any pre-treatment location-related heterogeneity. For future research, if pre-treatment locations are available, it would be helpful to control for origin prefecture effects in the model and explore heterogeneity by whether individuals whose land got expropriated moved within or outside the prefecture. Also, I could not estimate the impact of treatment by fixed effects due to a lack of variation in the treatment status over the two waves (2008 and 2009). For future research, if more waves are publicly available and there is enough variation in the treatment status, it is worth estimating these impacts by fixed effects to control for unobserved characteristics. Other extensions could be considering additional mechanisms, such as occupation or industry.

These results suggest three points. First, mental health support in the early years of displacement and physical health support for the long-term displaced are expected to reduce the adverse effects of land expropriation on displaced individuals. Second, providing medical insurance support is essential for individuals who experience land expropriation and change their *Hukou* from rural to urban. Third, it would be vital to disseminate health-related knowledge to the displaced individuals so that they are less affected by the urban lifestyle that increases their risk of being overweight or obese over time.

Appendix 1A. Outcome variables and sample selection

Table 1A.1: Correlation among outcome variables

	Mental Health Problem Index	Poor Health	BMI	If sick or injured	Medical Expenditure
Mental Health Problem Index	1.000				
Poor Health	0.277	1.000			
BMI	-0.073	-0.045	1.000		
If sick or injured	0.114	0.195	-0.008	1.000	
Medical Expenditure	0.177	0.316	-0.022	0.202	1.000

Note: Correlations with Mental Health Problem Index are calculated in a sample size of 19,876. The remaining correlations are calculated in a sample size of 27,709.

Table 1A.2: Mental Health Selection effect Estimation

	2008		2009	
	OLS	Selection	OLS	Selection
Male				
Inverse Mills ratio	-	0.028	-	-0.051
Whether present	-	3.942***	-	3.849***
Land	-0.005	-0.006	0.147**	0.148**
Intercept	1.114***	1.112***	1.116***	1.125***
N (OLS respondents only)	5,641	7,051	5,518	6,800
Female				
Inverse Mills ratio	-	0.037	-	0.009
Whether present	-	3.551***	-	3.256***
Land	0.039	0.039	0.075	0.076
Intercept	0.839***	0.835***	0.350*	0.349*
N (OLS respondents only)	4,376	7,012	4,341	6,789
Both (with gender control)				
Inverse Mills ratio	-	0.033	-	-0.020
Whether present	-	3.641***	-	3.432***
Land	0.015	0.014	0.104**	0.104**
Intercept	0.785***	0.780***	0.636***	0.641***
N (OLS respondents only)	10,017	14,063	9,859	13,589
Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. OLS regressions are based on the sub-sample that excludes the missing values of the mental health index. Each regression also controls for age, birth order, ethnicity, prefecture, parents' education and parents' age.				

Appendix 1B.a Urban movers

Table 1B.a.1: Health Outcomes - Urban movers

	Male	Sample size	Female	Sample size	All	Sample size
Mental Health Problem (Factor Analysis)	-0.028 (0.105)	10,894	0.013 (0.096)	8,445	-0.012 (0.070)	19,339
Poor health (0 or 1)	0.021 (0.023)	13,340	0.032 (0.026)	13,347	0.028 (0.018)	26,930
BMI	0.289 (0.267)	13,515	0.326 (0.363)	13,510	0.300 (0.229)	27,025
If sick or injured (in the last three months, 0 or 1)	0.098*** (0.033)	13,481	0.105*** (0.031)	13,477	0.102*** (0.023)	26,993
Medical Expenditure	482.514 (313.083)	13,515	836.329*** (157.293)	13,510	689.067*** (169.478)	27,025

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the individual level are shown in parentheses. BMI is a continuous variable calculated using height and weight. “Poor health” is a binary variable coded from self-reported health status. As poor health and if sick or injured are estimated using the random effects probit method, marginal effects are reported for these two outcomes. Each coefficient/marginal effect is from a separate regression.

Table 1B.a.2: Weight Categories - Urban movers

	Male	Female	All
Weight Categories			
Underweight (< 18.5)	-0.014*	-0.009	-0.012
	(0.008)	(0.013)	(0.008)
Normal (≥ 18.5 & < 23)	-0.049	-0.014	-0.028
	(0.034)	(0.023)	(0.020)
Overweight (≥ 23 & < 27.5)	0.051	0.020	0.034
	(0.033)	(0.031)	(0.023)
Obese (≥ 27.5)	0.011	0.003	0.006
	(0.009)	(0.005)	(0.005)
Sample size	13,515	13,510	27,025

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the individual level are shown in parentheses. Each column (columns 2-4) is from one regression. Column 2 uses male data only, column 3 uses female data only, and column 4 uses male and female data together. All regressions are estimated using random effects interval regression. Marginal effects are reported for each BMI category.

Table 1B.a.3: Self-reported Health Status - Urban movers

	Male	Female	All
Health Status Categories			
Excellent	-0.083*** (0.021)	-0.059*** (0.017)	-0.070*** (0.013)
Good	-0.034** (0.015)	-0.038** (0.016)	-0.037*** (0.011)
Average	0.076*** (0.022)	0.056*** (0.017)	0.066*** (0.014)
Poor	0.032*** (0.011)	0.033*** (0.013)	0.033*** (0.008)
Very poor	0.009** (0.004)	0.008** (0.003)	0.008*** (0.002)
Sample size	13,515	13,510	27,025

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the individual level are shown in parentheses. Each column (columns 2-4) is from one regression. Column 2 uses male data only, column 3 uses female data only, and column 4 uses male and female data together. All regressions are estimated using random effects ordered probit. Marginal effects are reported for each health status.

Appendix 1B.b Rural stayers

Table 1B.b.1: Health Outcomes - Rural stayers

	Male	Sample size	Female	Sample size	All	Sample size
Mental Health Problem (Factor Analysis)	0.064 (0.071)	10,943	0.072 (0.074)	8,401	0.068 (0.051)	19,344
Poor health (0 or 1)	0.026 (0.021)	13,357	-0.020 (0.017)	13,376	0.003 (0.013)	26,952
BMI	0.038 (0.193)	13,522	0.231 (0.301)	13,448	0.134 (0.179)	26,970
If sick or injured (in the last three months, 0 or 1)	-0.022 (0.017)	13,515	-0.044** (0.018)	13,428	-0.032*** (0.012)	26,952
Medical Expenditure	255.237 (233.011)	13,522	-224.306 (125.316)	13,448	18.077 (132.704)	26,970

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the individual level are shown in parentheses. BMI is a continuous variable calculated using height and weight. “Poor health” is a binary variable coded from self-reported health status. As poor health and if sick or injured are estimated using the random effects probit method, marginal effects are reported for these two outcomes. Each coefficient/marginal effect is from a separate regression.

Table 1B.b.2: Weight Categories - Rural stayers

	Male	Female	All
Weight Categories			
Underweight (< 18.5)	-0.003 (0.007)	0.001 (0.011)	-0.001 (0.007)
Normal (≥ 18.5 & < 23)	-0.008 (0.023)	0.001 (0.017)	-0.003 (0.014)
Overweight (≥ 23 & < 27.5)	0.009 (0.025)	-0.001 (0.025)	0.004 (0.018)
Obese (≥ 27.5)	0.002 (0.005)	0.000 (0.003)	0.001 (0.003)
Sample size	13,522	13,448	26,970

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the individual level are shown in parentheses. Each column (columns 2-4) is from one regression. Column 2 uses male data only, column 3 uses female data only, and column 4 uses male and female data together. All regressions are estimated using random effects interval regression. Marginal effects are reported for each BMI category.

Table 1B.b.3: Self-reported Health Status - Rural stayers

	Male	Female	All
Health Status Categories			
Excellent	0.027 (0.022)	0.035* (0.020)	0.032** (0.015)
Good	0.002*** (0.001)	0.010** (0.004)	0.006*** (0.002)
Average	-0.021 (0.017)	-0.029* (0.016)	-0.025** (0.012)
Poor	-0.007 (0.005)	-0.013* (0.007)	-0.010** (0.004)
Very poor	-0.001 (0.001)	-0.003** (0.001)	-0.002** (0.001)
Sample size	13,522	13,448	26,970

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the individual level are shown in parentheses. Each column (columns 2-4) is from one regression. Column 2 uses male data only, column 3 uses female data only, and column 4 uses male and female data together. All regressions are estimated using random effects ordered probit. Marginal effects are reported for each health status.

Appendix 1C. Matching

Table 1C.1: Matching Quality

Variable	Unmatched Matched	Mean		% bias	% reduction of $ bias $	t-test		$V_e(T) /$ $v_e(C)$
		Treated on support	Treated off support			t	$p > t $	
Age	U	46.968	49.021	-20.2		-5.66	0.000	0.97
	M	46.968	46.944	0.2	98.8	0.05	0.963	1.03
Male	U	0.477	0.503	-5.1		-1.39	0.163	0.99
	M	0.477	0.477	0.0	100.0	0.00	1.000	1.00
Birth order	U	2.616	2.455	10.6		2.95	0.003	1.12
	M	2.616	2.607	0.6	94.4	0.12	0.908	1.02
Highest education of parents	U	3.087	2.141	33.3		8.38	0.000	0.82
	M	3.087	3.040	1.7	95.0	0.37	0.714	1.05
Average age of parents	U	75.446	77.802	-21.5		-6.04	0.000	0.98
	M	75.446	75.477	-0.3	98.7	-0.06	0.956	1.05

Note: Bias reduction of more than 90% for every control variables.

Sample	Ps R2	LR chi2	p>chi2	MeanBias	MedBias	B	R	%concern	%bad
Unmatched	0.016	98.08	0.000	18.2	20.2	38.5*	0.87	0	0
Matched	0.000	0.22	0.999	0.6	0.3	2.4	1.05	0	0

Table 1C.2: Health Outcomes - Matching

	Male	Sample size	Female	Sample size	All	Sample size
Mental Health Problem (Factor Analysis)	-0.092 (0.090)	992	0.090 (0.103)	1,007	-0.012 (0.067)	1,999
Poor health (0 or 1)	0.022 (0.025)	768	-0.028 (0.026)	951	0.002 (0.015)	2,162
BMI	0.253 (0.226)	1,306	0.307 (0.360)	1,426	0.323 (0.209)	2,732
If sick or injured (in the last three months, 0 or 1)	0.029 (0.031)	964	0.021 (0.030)	1,171	0.019 (0.020)	2,409
Medical Expenditure	548.933 (307.827)	1,306	320.166* (129.501)	1,426	425.404** (163.946)	2,732

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the individual level are shown in parentheses. BMI is a continuous variable calculated using height and weight. "Poor health" is a binary variable coded from self-reported health status. As poor health and if sick or injured are estimated using the random effects probit method, marginal effects are reported for these two outcomes. Each coefficient/marginal effect is from a separate regression.

Table 1C.3: Weight Categories - Matching

	Male	Female	All
Weight Categories			
Underweight (< 18.5)	-0.010 (0.007)	-0.012 (0.014)	-0.014* (0.008)
Normal (≥ 18.5 & < 23)	-0.043 (0.031)	-0.018 (0.022)	-0.033* (0.019)
Overweight (≥ 23 & < 27.5)	0.041 (0.029)	0.026 (0.031)	0.038* (0.022)
Obese (≥ 27.5)	0.013 (0.009)	0.004 (0.005)	0.009* (0.005)
Sample size	1,306	1,426	2,732

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the individual level are shown in parentheses. Each column (columns 2-4) is from one regression. Column 2 uses male data only, column 3 uses female data only, and column 4 uses male and female data together. All regressions are estimated using random effects interval regression. Marginal effects are reported for each BMI category.

Table 1C.4: Self-reported Health Status - Matching

	Male	Female	All
Health Status Categories			
Excellent	-0.021 (0.026)	0.018 (0.022)	-0.001 (0.017)
Good	-0.002 (0.002)	0.006 (0.008)	0.000 (0.004)
Average	0.017 (0.020)	-0.017 (0.021)	0.001 (0.015)
Poor	0.005 (0.007)	-0.006 (0.008)	0.000 (0.005)
Very poor	0.001 (0.001)	-0.001 (0.002)	0.000 (0.001)
Sample size	1,306	1,426	2,732

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the individual level are shown in parentheses. Each column (columns 2-4) is from one regression. Column 2 uses male data only, column 3 uses female data only, and column 4 uses male and female data together. All regressions are estimated using random effects ordered probit. Marginal effects are reported for each health status.

Table 1C.5: Medical Insurance - Matching

	Male	Sample size	Female	Sample size	All	Sample size
Medical Insurance (0 or 1)	-0.086** (0.034)	549	-0.059** (0.027)	651	-0.069*** (0.019)	1,348
<i>Rural cooperative medical insurance (0 or 1)</i>	-0.365*** (0.036)	800	-0.449*** (0.034)	867	-0.386*** (0.024)	1,837
<i>Public medical insurance (paid for by the government) (0 or 1)</i>	0.058** (0.026)	646	0.065*** (0.020)	627	0.058*** (0.015)	1,474
<i>Commercial medical insurance (0 or 1)</i>	- (0.026)	-	0.051* (0.026)	525	0.064*** (0.013)	1,110
Amount paid for medical insurance in 2007/2008 (including amounts deducted from payroll)	217.569*** (54.826)	1,252	288.723*** (76.394)	1,364	252.086*** (46.747)	2,616

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the individual level are shown in parentheses. Marginal effects are reported for each outcome and are shown with gender separation and together. Each marginal effect is from one regression. The estimation method for medical insurance, rural cooperative medical insurance, public medical insurance and commercial medical insurance are estimated with random effects probit, while the amount paid for medical insurance is estimated with random effects.

Table 1C.6: Education and children - Matching

	Male	Sample size	Female	Sample size	All	Sample size
Years of Education	0.732** (0.227)	1,293	0.869*** (0.245)	1,418	0.821*** (0.168)	2,711
Number of Children	-0.184** (0.065)	1,288	-0.155* (0.067)	1,421	-0.166*** (0.046)	2,709

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the individual level are shown in parentheses. Each coefficient is from one regression. Years of education and number of children are both estimated using random effects.

Appendix 1D. Treatment effects over time with different time intervals

Table 1D.1: Health Outcomes - Time Interval version 1

	Mental Health Problem (Factor Anal- ysis)	S.E.	Poor health (0 or 1)	S.E.	BMI	S.E.	If sick (in the last three months, 0 or 1)	S.E.	Medical Expen- diture	S.E.
0-4 years										
Male	0.091	(0.149)	-	-	-0.148	(0.405)	-0.032	(0.033)	-225.871	(506.703)
Female	-0.152	(0.153)	0.045	(0.042)	0.056	(0.588)	0.021	(0.044)	-37.292	(266.170)
All	-0.032	(0.106)	0.008	(0.025)	-0.071	(0.359)	-0.002	(0.028)	-67.474	(283.440)
4-8 years										
Male	0.065	(0.077)	-0.005	(0.017)	0.152	(0.213)	0.010	(0.021)	482.689	(264.476)
Female	0.047	(0.079)	0.001	(0.021)	0.47	(0.316)	0.000	(0.022)	29.881	(138.912)
All	0.055	(0.055)	-0.002	(0.013)	0.329	(0.191)	0.005	(0.015)	249.313	(148.139)
8-12 years										
Male	0.211	(0.110)	0.027	(0.032)	0.17	(0.277)	-0.007	(0.031)	-132.994	(354.709)
Female	0.225*	(0.107)	-0.016	(0.026)	0.351	(0.402)	0.002	(0.033)	149.184	(185.058)
All	0.218**	(0.076)	0.004	(0.020)	0.274	(0.245)	0.000	(0.023)	5.704	(198.784)
12-16 years										
Male	-0.231	(0.141)	0.050	(0.040)	0.496	(0.346)	0.066	(0.048)	134.701	(461.239)
Female	-0.253	(0.139)	-0.025	(0.030)	-0.284	(0.517)	0.060	(0.049)	257.068	(245.256)
All	-0.248*	(0.098)	0.014	(0.025)	0.057	(0.310)	0.059*	(0.033)	186.559	(262.143)
16-20 years										
Male	0.077	(0.211)	0.011	(0.046)	-0.771	(0.437)	0.090	(0.069)	81.088	(581.264)
Female	0.226	(0.145)	-0.037	(0.027)	-0.05	(0.544)	-0.025	(0.038)	128.45	(249.695)
All	0.162	(0.116)	-0.014	(0.024)	-0.433	(0.352)	0.012	(0.033)	127.648	(289.947)
over 20 years										
Male	-0.381*	(0.179)	0.075	(0.046)	0.807*	(0.387)	0.063	(0.046)	848.335	(478.279)
Female	-0.176	(0.150)	0.030	(0.042)	0.439	(0.592)	0.091*	(0.051)	535.715*	(258.389)
All	-0.263*	(0.113)	0.059*	(0.032)	0.601	(0.352)	0.077**	(0.034)	684.186*	(270.029)
Sample size										
Male	11,159		13,594		13,857		13,823		13,857	
Female	8,717		13,689		13,852		13,819		13,852	
All	19,876		27,614		27,709		27,677		27,709	

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the individual level are shown in parentheses. BMI is a continuous variable calculated using height and weight. "Poor health" is a binary variable coded from self-reported health status. As poor health and if sick or injured are estimated using the random effects probit method, marginal effects are reported for these two outcomes. Each coefficient/marginal effect is from a separate regression.

Table 1D.2: Weight Categories - Time Interval version 1

	Underweight (< 18.5)	S.E.	Normal (≥ 18.5 & < 23)	S.E.	Overweight (≥ 23 & < 27.5)	S.E.	Obese (≥ 27.5)	S.E.
0-4 years								
Male	0.003	(0.016)	0.008	(0.045)	-0.009	(0.052)	-0.002	(0.010)
Female	-0.015	(0.019)	-0.026	(0.039)	0.035	(0.050)	0.005	(0.009)
All	-0.006	(0.012)	-0.014	(0.030)	0.017	(0.036)	0.003	(0.007)
4-8 years								
Male	-0.003	(0.008)	-0.008	(0.025)	0.009	(0.027)	0.002	(0.006)
Female	-0.003	(0.012)	-0.004	(0.019)	0.006	(0.027)	0.001	(0.004)
All	-0.003	(0.007)	-0.007	(0.016)	0.009	(0.019)	0.002	(0.003)
8-12 years								
Male	-0.014*	(0.008)	-0.051	(0.034)	0.052	(0.033)	0.012	(0.009)
Female	-0.018	(0.013)	-0.032	(0.027)	0.043	(0.034)	0.007	(0.006)
All	-0.016**	(0.007)	-0.041*	(0.022)	0.047**	(0.024)	0.009*	(0.005)
12-16 years								
Male	-0.022***	(0.008)	-0.090**	(0.043)	0.088**	(0.037)	0.024*	(0.014)
Female	0.011	(0.022)	0.015	(0.027)	-0.023	(0.044)	-0.003	(0.005)
All	-0.010	(0.010)	-0.024	(0.027)	0.029	(0.031)	0.005	(0.006)
16-20 years								
Male	0.023	(0.022)	0.054	(0.040)	-0.066	(0.054)	-0.011	(0.007)
Female	0.030	(0.027)	0.035	(0.023)	-0.058	(0.045)	-0.007	(0.004)
All	0.033*	(0.018)	0.051**	(0.020)	-0.074**	(0.034)	-0.010***	(0.004)
over 20 years								
Male	-0.022**	(0.009)	-0.092*	(0.050)	0.090**	(0.043)	0.024	(0.016)
Female	-0.005	(0.022)	-0.008	(0.036)	0.011	(0.051)	0.002	(0.008)
All	-0.018*	(0.010)	-0.045	(0.032)	0.052	(0.034)	0.010	(0.008)

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Sample size: 13,857 male, 13,852 female. Standard errors clustered at the individual level are shown in parentheses. Marginal effects are reported for each BMI category and are shown with gender separation and together. Each line is from one regression estimated. All regressions are estimated using random effects interval regression. Marginal effects are reported.

Table 1D.3: Self-reported Health Status - Time Interval version 1

	Excellent	S.E.	Good	S.E.	Average	S.E.	Poor	S.E.	Very poor	S.E.
0-4 years										
Male	0.031	(0.047)	0.002***	(0.001)	-0.024	(0.035)	-0.007	(0.010)	-0.002	(0.002)
Female	-0.053*	(0.029)	-0.033	(0.025)	0.050*	(0.029)	0.029	(0.020)	0.007	(0.005)
All	-0.021	(0.026)	-0.007	(0.010)	0.018	(0.023)	0.008	(0.011)	0.002	(0.002)
4-8 years										
Male	-0.009	(0.022)	-0.001	(0.004)	0.007	(0.019)	0.002	(0.006)	0.001	(0.001)
Female	-0.013	(0.019)	-0.006	(0.009)	0.012	(0.017)	0.006	(0.009)	0.001	(0.002)
All	-0.011	(0.014)	-0.003	(0.005)	0.009	(0.013)	0.004	(0.005)	0.001	(0.001)
8-12 years										
Male	0.020	(0.032)	0.002	(0.002)	-0.016	(0.025)	-0.005	(0.007)	-0.001	(0.002)
Female	0.054*	(0.030)	0.013***	(0.003)	-0.045*	(0.023)	-0.019**	(0.009)	-0.004**	(0.001)
All	0.038*	(0.022)	0.006***	(0.002)	-0.030*	(0.017)	-0.011**	(0.006)	-0.002**	(0.001)
12-16 years										
Male	-0.028	(0.034)	-0.006	(0.010)	0.024	(0.030)	0.009	(0.012)	0.002	(0.003)
Female	-0.012	(0.032)	-0.006	(0.015)	0.011	(0.029)	0.006	(0.015)	0.001	(0.003)
All	-0.019	(0.023)	-0.006	(0.009)	0.017	(0.021)	0.007	(0.009)	0.002	(0.002)
16-20 years										
Male	-0.038	(0.042)	-0.010	(0.015)	0.033	(0.038)	0.012	(0.015)	0.003	(0.004)
Female	-0.027	(0.031)	-0.014	(0.019)	0.025	(0.029)	0.013	(0.017)	0.003	(0.004)
All	-0.032	(0.025)	-0.012	(0.012)	0.029	(0.023)	0.013	(0.011)	0.003	(0.003)
over 20 years										
Male	-0.061*	(0.034)	-0.020	(0.018)	0.054*	(0.033)	0.021	(0.015)	0.005	(0.004)
Female	-0.001	(0.036)	-0.001	(0.015)	0.001	(0.032)	0.001	(0.016)	0.000	(0.003)
All	-0.033	(0.024)	-0.012	(0.012)	0.029	(0.023)	0.013	(0.011)	0.003	(0.003)

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Sample size: 13,857 male, 13,852 female. Standard errors clustered at the individual level are shown in parentheses. Marginal effects are reported for each health status and are shown with gender separation and together. Results for each gender are from one regression estimated. All regressions are estimated using random effects ordered probit. Marginal effects are reported.

Table 1D.4: Health Outcomes - Time Interval version 2

	Mental Health Problem (Factor Anal- ysis)	S.E.	Poor health (0 or 1)	S.E.	BMI	S.E.	If sick (in the last three months, 0 or 1)	S.E.	Medical Expen- diture	S.E.
0-6 years										
Male	0.198*	(0.085)	-0.041***	(0.009)	0.352	(0.234)	-0.015	(0.021)	-189.759	(300.549)
Female	0.081	(0.086)	-0.010	(0.021)	-0.149	(0.338)	-0.021	(0.022)	-91.67	(154.103)
All	0.138*	(0.060)	-0.024**	(0.012)	0.100	(0.207)	-0.016	(0.015)	-105.541	(166.835)
6-9 years										
Male	-0.045	(0.086)	0.042	(0.028)	-0.249	(0.234)	0.016	(0.026)	834.094**	(310.101)
Female	0.032	(0.089)	0.036	(0.028)	0.916**	(0.338)	0.024	(0.029)	179.96	(158.313)
All	-0.009	(0.061)	0.040**	(0.019)	0.359	(0.206)	0.020	(0.019)	486.835**	(172.532)
9-20 years										
Male	0.019	(0.093)	0.017	(0.023)	0.12	(0.229)	0.046	(0.030)	10.88	(286.550)
Female	0.044	(0.090)	-0.028	(0.018)	0.123	(0.336)	0.022	(0.027)	171.445	(149.220)
All	0.027	(0.064)	-0.005	(0.014)	0.076	(0.204)	0.032	(0.020)	90.769	(159.516)
over 20 years										
Male	-0.389*	(0.179)	0.074	(0.046)	0.853*	(0.387)	0.064	(0.047)	849.839	(477.794)
Female	-0.187	(0.150)	0.030	(0.042)	0.46	(0.592)	0.094*	(0.052)	538.183*	(258.435)
All	-0.271*	(0.113)	0.057*	(0.032)	0.632	(0.352)	0.078**	(0.035)	685.333*	(269.955)
Sample size										
Male	11,159		13,653		13,857		13,823		13,857	
Female	8,717		13,689		13,852		13,819		13,852	
All	19,876		27,614		27,709		27,677		27,709	

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the individual level are shown in parentheses. BMI is a continuous variable calculated using height and weight. Poor health is a binary variable, coded from self-reported health status. As poor health and if sick or injured are estimated using the random effects probit method, marginal effects are reported for these two outcomes. Each coefficient/marginal effect is from a separate regression.

Table 1D.5: Weight Categories - Time Interval version 2

	Underweight (< 18.5)	S.E.	Normal (≥ 18.5 & < 23)	S.E.	Overweight (≥ 23 & < 27.5)	S.E.	Obese (≥ 27.5)	S.E.
0-6 years								
Male	-0.015**	(0.006)	-0.057*	(0.029)	0.058**	(0.028)	0.014*	(0.008)
Female	-0.012	(0.012)	-0.020	(0.022)	0.028	(0.029)	0.004	(0.005)
All	-0.015**	(0.006)	-0.037**	(0.018)	0.044**	(0.020)	0.008*	(0.004)
6-9 years								
Male	0.016	(0.011)	0.040*	(0.023)	-0.048	(0.030)	-0.008*	(0.004)
Female	0.006	(0.014)	0.008	(0.019)	-0.012	(0.029)	-0.002	(0.004)
All	0.011	(0.009)	0.021	(0.015)	-0.028	(0.021)	-0.004	(0.003)
9-20 years								
Male	-0.011	(0.007)	-0.039	(0.028)	0.041	(0.028)	0.009	(0.007)
Female	-0.001	(0.013)	-0.002	(0.020)	0.002	(0.028)	0.000	(0.004)
All	-0.006	(0.007)	-0.013	(0.017)	0.016	(0.020)	0.003	(0.004)
over 20 years								
Male	-0.023***	(0.008)	-0.100**	(0.051)	0.096**	(0.042)	0.027	(0.017)
Female	-0.006	(0.022)	-0.010	(0.037)	0.014	(0.051)	0.002	(0.008)
All	-0.019*	(0.010)	-0.050	(0.032)	0.057*	(0.034)	0.012	(0.008)

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Sample size: 13,857 male, 13,852 female. Standard errors clustered at the individual level are shown in parentheses. Marginal effects are reported for each BMI category and are shown with gender separation and together. Each line is from one regression estimated. All regressions are estimated using random effects interval regression. Marginal effects are reported.

Table 1D.6: Self-reported Health Status - Time Interval version 2

	Excellent	S.E.	Good	S.E.	Average	S.E.	Poor	S.E.	Very poor	S.E.
0-6 years										
Male	0.058**	(0.029)	0.001	(0.003)	-0.044**	(0.020)	-0.013**	(0.005)	-0.003***	(0.001)
Female	0.009	(0.022)	0.003	(0.007)	-0.008	(0.019)	-0.004	(0.009)	-0.001	(0.002)
All	0.030*	(0.018)	0.006***	(0.002)	-0.024*	(0.014)	-0.009*	(0.005)	-0.002**	(0.001)
6-9 years		-		-		-		-		-
Male	-0.058***	(0.021)	-0.018*	(0.011)	0.051**	(0.020)	0.020**	(0.009)	0.005*	(0.003)
Female	-0.035*	(0.019)	-0.019	(0.013)	0.032*	(0.018)	0.018	(0.011)	0.004	(0.003)
All	-0.045***	(0.014)	-0.019**	(0.008)	0.041***	(0.013)	0.019***	(0.007)	0.004**	(0.002)
9-20 years		-		-		-		-		-
Male	-0.008	(0.024)	-0.001	(0.005)	0.007	(0.020)	0.002	(0.007)	0.001	(0.002)
Female	0.003	(0.021)	0.001	(0.008)	-0.003	(0.019)	-0.001	(0.009)	0.000	(0.002)
All	-0.002	(0.016)	0.000	(0.004)	0.002	(0.014)	0.001	(0.006)	0.000	(0.001)
over 20 years		-		-		-		-		-
Male	-0.059*	(0.034)	-0.019	(0.018)	0.053	(0.033)	0.020	(0.015)	0.005	(0.004)
Female	-0.002	(0.036)	-0.001	(0.015)	0.002	(0.032)	0.001	(0.016)	0.000	(0.003)
All	-0.032	(0.024)	-0.012	(0.012)	0.028	(0.023)	0.013	(0.011)	0.003	(0.003)

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Sample size: 13,857 male, 13,852 female. Standard errors clustered at the individual level are shown in parentheses. Marginal effects are reported for each health status and are shown with gender separation and together. Results for each gender are from one regression estimated. All regressions are estimated using random effects ordered probit. Marginal effects are reported.

Table 1D.7: Health Outcomes - Time Interval version 3

	Mental Health Problem (Factor Anal- ysis)	S.E.	Poor health (0 or 1)	S.E.	BMI	S.E.	If sick (in the last three months, 0 or 1)	S.E.	Medical Expen- diture	S.E.
0-5 years										
Male	0.237*	(0.104)	-	-	-0.008	(0.288)	-0.016	(0.026)	-191.238	(377.413)
Female	0.119	(0.111)	0.025	(0.032)	0.346	(0.414)	0.012	(0.032)	-20.274	(193.160)
All	0.175*	(0.076)	-0.009	(0.017)	0.177	(0.253)	0.000	(0.021)	-69.909	(209.586)
5-10 years										
Male	0.051	(0.074)	0.013	(0.020)	-0.015	(0.199)	0.003	(0.020)	431.894	(247.855)
Female	0.043	(0.075)	0.009	(0.021)	0.438	(0.295)	0.004	(0.021)	133.679	(130.128)
All	0.047	(0.052)	0.012	(0.014)	0.235	(0.179)	0.004	(0.014)	281.488*	(138.911)
10-15 years										
Male	-0.018	(0.115)	0.017	(0.029)	0.759*	(0.298)	0.042	(0.038)	2.131	(376.564)
Female	-0.014	(0.122)	-0.033	(0.021)	-0.021	(0.449)	-0.006	(0.033)	104.764	(202.645)
All	-0.016	(0.083)	-0.006	(0.018)	0.326	(0.270)	0.016	(0.025)	33.667	(213.834)
15-20 years										
Male	-0.054	(0.174)	0.031	(0.043)	-0.488	(0.379)	0.085	(0.056)	31.333	(487.113)
Female	0.096	(0.137)	-0.043**	(0.021)	-0.117	(0.509)	0.025	(0.042)	27.043	(229.243)
All	0.024	(0.105)	-0.010	(0.021)	-0.355	(0.320)	0.044	(0.032)	54.531	(255.841)
over 20 years										
Male	-0.381*	(0.179)	0.074	(0.046)	0.826*	(0.387)	0.063	(0.046)	842.462	(478.145)
Female	-0.184	(0.150)	0.030	(0.042)	0.43	(0.592)	0.092*	(0.052)	524.558*	(258.351)
All	-0.269*	(0.113)	0.058*	(0.032)	0.614	(0.352)	0.078**	(0.035)	680.009*	(269.973)
Sample size										
Male	11,159		13,542		13,857		13,823		13,857	
Female	8,717		13,689		13,852		13,819		13,852	
All	19,876		27,614		27,709		27,677		27,709	

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the individual level are shown in parentheses. BMI is a continuous variable calculated using height and weight. Poor health is a binary variable, coded from self-reported health status. As poor health and if sick or injured are estimated using the random effects probit method, marginal effects are reported for these two outcomes. Each coefficient/marginal effect is from a separate regression.

Table 1D.8: Weight Categories - Time Interval version 3

	Underweight (< 18.5)	S.E.	Normal (≥ 18.5 & < 23)	S.E.	Overweight (≥ 23 & < 27.5)	S.E.	Obese (≥ 27.5)	S.E.
0-5 years								
Male	0.004	(0.011)	0.011	(0.031)	-0.013	(0.036)	-0.002	(0.007)
Female	-0.013	(0.014)	-0.022	(0.027)	0.030	(0.035)	0.005	(0.006)
All	-0.005	(0.009)	-0.010	(0.021)	0.013	(0.025)	0.002	(0.005)
5-10 years								
Male	-0.001	(0.007)	-0.003	(0.023)	0.004	(0.025)	0.001	(0.005)
Female	-0.006	(0.011)	-0.009	(0.018)	0.013	(0.025)	0.002	(0.004)
All	-0.004	(0.006)	-0.009	(0.015)	0.011	(0.018)	0.002	(0.003)
10-15 years								
Male	-0.026***	(0.006)	-0.117***	(0.039)	0.110***	(0.031)	0.032**	(0.014)
Female	-0.001	(0.017)	-0.001	(0.026)	0.002	(0.038)	0.000	(0.005)
All	-0.018**	(0.008)	-0.047*	(0.024)	0.054**	(0.026)	0.011*	(0.006)
15-20 years								
Male	0.007	(0.016)	0.018	(0.041)	-0.021	(0.048)	-0.004	(0.008)
Female	0.018	(0.023)	0.023	(0.025)	-0.037	(0.043)	-0.005	(0.005)
All	0.017	(0.014)	0.030	(0.022)	-0.041	(0.032)	-0.006	(0.004)
over 20 years								
Male	-0.023***	(0.009)	-0.095*	(0.051)	0.092**	(0.043)	0.025	(0.016)
Female	-0.005	(0.022)	-0.008	(0.036)	0.011	(0.051)	0.002	(0.008)
All	-0.018*	(0.010)	-0.047	(0.032)	0.055	(0.034)	0.011	(0.008)

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Sample size: 13,857 male, 13,852 female. Standard errors clustered at the individual level are shown in parentheses. Marginal effects are reported for each BMI category and are shown with gender separation and together. Each line is from one regression estimated. All regressions are estimated using random effects interval regression. Marginal effects are reported.

Table 1D.9: Self-reported Health Status - Time Interval version 3

	Excellent	S.E.	Good	S.E.	Average	S.E.	Poor	S.E.	Very poor	S.E.
0-5 years										
Male	0.031	(0.033)	0.002***	(0.001)	-0.024	(0.025)	-0.008	(0.007)	-0.002	(0.001)
Female	-0.007	(0.026)	-0.003	(0.011)	0.006	(0.023)	0.003	(0.012)	0.001	(0.002)
All	0.009	(0.021)	0.002	(0.004)	-0.008	(0.017)	-0.003	(0.007)	-0.001	(0.001)
5-10 years										
Male	0.003	(0.022)	0.000	(0.003)	-0.002	(0.017)	-0.001	(0.006)	0.000	(0.001)
Female	0.000	(0.018)	0.000	(0.007)	0.000	(0.016)	0.000	(0.008)	0.000	(0.002)
All	0.001	(0.014)	0.000	(0.004)	-0.001	(0.012)	0.000	(0.005)	0.000	(0.001)
10-15 years										
Male	-0.029	(0.029)	-0.007	(0.009)	0.025	(0.026)	0.009	(0.010)	0.002	(0.002)
Female	-0.009	(0.027)	-0.004	(0.012)	0.008	(0.025)	0.004	(0.013)	0.001	(0.003)
All	-0.019	(0.020)	-0.006	(0.008)	0.016	(0.018)	0.007	(0.008)	0.002	(0.002)
15-20 years										
Male	-0.029	(0.037)	-0.007	(0.012)	0.025	(0.033)	0.009	(0.013)	0.002	(0.003)
Female	-0.022	(0.029)	-0.011	(0.017)	0.020	(0.027)	0.011	(0.015)	0.002	(0.003)
All	-0.026	(0.023)	-0.009	(0.010)	0.022	(0.021)	0.010	(0.010)	0.002	(0.002)
over 20 years										
Male	-0.060*	(0.034)	-0.019	(0.018)	0.054	(0.033)	0.021	(0.015)	0.005	(0.004)
Female	-0.004	(0.036)	-0.002	(0.015)	0.004	(0.032)	0.002	(0.016)	0.000	(0.003)
All	-0.034	(0.024)	-0.013	(0.012)	0.030	(0.022)	0.013	(0.011)	0.003	(0.003)

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Sample size: 13,857 male, 13,852 female. Standard errors clustered at the individual level are shown in parentheses. Marginal effects are reported for each health status and are shown with gender separation and together. Results for each gender are from one regression estimated. All regressions are estimated using random effects ordered probit. Marginal effects are reported.

Table 1D.10: Health Outcomes - Time Interval version 4

	Mental Health Problem (Factor Anal- ysis)	S.E.	Poor health (0 or 1)	S.E.	BMI	S.E.	If sick (in the last three months, 0 or 1)	S.E.	Medical Expen- diture	S.E.
0-5 years										
Male	0.241*	(0.104)	-	-	-0.034	(0.288)	-0.016	(0.027)	-207.806	(377.183)
Female	0.121	(0.111)	0.025	(0.032)	0.345	(0.414)	0.011	(0.032)	-22.154	(193.168)
All	0.178*	(0.076)	-0.009	(0.017)	0.167	(0.253)	0.000	(0.021)	-76.655	(209.562)
5-10 years										
Male	0.054	(0.074)	0.013	(0.020)	-0.026	(0.199)	0.003	(0.020)	422.524	(247.749)
Female	0.044	(0.075)	0.008	(0.021)	0.438	(0.295)	0.004	(0.021)	132.554	(130.136)
All	0.048	(0.052)	0.012	(0.014)	0.233	(0.179)	0.004	(0.014)	277.818*	(138.904)
10-15 years										
Male	-0.022	(0.115)	0.019	(0.029)	0.789**	(0.298)	0.041	(0.038)	3.182	(376.581)
Female	-0.018	(0.122)	-0.028	(0.022)	-0.012	(0.449)	-0.006	(0.033)	106.08	(202.658)
All	-0.02	(0.083)	-0.003	(0.018)	0.347	(0.269)	0.017	(0.025)	33.573	(213.847)
over 15 years										
Male	-0.213	(0.128)	0.056*	(0.033)	0.154	(0.283)	0.072*	(0.037)	444.912	(351.551)
Female	-0.029	(0.107)	-0.009	(0.023)	0.109	(0.410)	0.056	(0.034)	243.003	(179.627)
All	-0.111	(0.080)	0.024	(0.020)	0.077	(0.250)	0.061**	(0.025)	348.955	(193.205)
Sample size										
Male	11,159		13,542		13,857		13,823		13,857	
Female	8,717		13,689		13,852		13,819		13,852	
All	19,876		27,614		27,709		27,677		27,709	

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the individual level are shown in parentheses. BMI is a continuous variable calculated using height and weight. "Poor health" is a binary variable coded from self-reported health status. As poor health and if sick or injured are estimated using the random effects probit method, marginal effects are reported for these two outcomes. Each coefficient/marginal effect is from a separate regression.

Table 1D.11: Weight Categories - Time Interval version 4

	Underweight (< 18.5)	S.E.	Normal (≥ 18.5 & < 23)	S.E.	Overweight (≥ 23 & < 27.5)	S.E.	Obese (≥ 27.5)	S.E.
0-5 years								
Male	0.005	(0.011)	0.013	(0.031)	-0.015	(0.036)	-0.003	(0.006)
Female	-0.013	(0.014)	-0.022	(0.027)	0.030	(0.035)	0.005	(0.006)
All	-0.004	(0.009)	-0.010	(0.021)	0.012	(0.025)	0.002	(0.005)
5-10 years								
Male	-0.001	(0.007)	-0.003	(0.023)	0.003	(0.025)	0.001	(0.005)
Female	-0.006	(0.011)	-0.009	(0.018)	0.013	(0.025)	0.002	(0.004)
All	-0.004	(0.006)	-0.009	(0.015)	0.011	(0.018)	0.002	(0.003)
10-15 years								
Male	-0.027***	(0.006)	-0.120***	(0.039)	0.113***	(0.031)	0.034**	(0.014)
Female	-0.001	(0.017)	-0.002	(0.027)	0.003	(0.038)	0.000	(0.005)
All	-0.019**	(0.008)	-0.049**	(0.025)	0.056**	(0.026)	0.011*	(0.006)
over 15 years								
Male	-0.010	(0.009)	-0.033	(0.035)	0.035	(0.035)	0.008	(0.009)
Female	0.008	(0.017)	0.011	(0.022)	-0.017	(0.035)	-0.002	(0.004)
All	-0.001	(0.009)	-0.001	(0.020)	0.001	(0.025)	0.000	(0.004)

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Sample size: 13,857 male, 13,852 female. Standard errors clustered at the individual level are shown in parentheses. Marginal effects are reported for each BMI category and are shown with gender separation and together. Each line is from one regression estimated. All regressions are estimated using random effects interval regression. Marginal effects are reported.

Table 1D.12: Self-reported Health Status - Time Interval version 4

	Excellent	S.E.	Good	S.E.	Average	S.E.	Poor	S.E.	Very poor	S.E.
0-5 years										
Male	0.032	(0.033)	0.002***	(0.000)	-0.025	(0.025)	-0.008	(0.007)	-0.002	(0.001)
Female	-0.007	(0.026)	-0.003	(0.011)	0.006	(0.023)	0.003	(0.012)	0.001	(0.002)
All	0.009	(0.021)	0.002	(0.004)	-0.008	(0.017)	-0.003	(0.007)	-0.001	(0.001)
5-10 years										
Male	0.003	(0.022)	0.000	(0.003)	-0.003	(0.017)	-0.001	(0.006)	0.000	(0.001)
Female	0.000	(0.018)	0.000	(0.007)	0.000	(0.016)	0.000	(0.008)	0.000	(0.002)
All	0.001	(0.014)	0.000	(0.004)	-0.001	(0.012)	0.000	(0.005)	0.000	(0.001)
10-15 years										
Male	-0.030	(0.029)	-0.007	(0.009)	0.025	(0.026)	0.009	(0.010)	0.002	(0.002)
Female	-0.008	(0.027)	-0.004	(0.012)	0.007	(0.025)	0.004	(0.013)	0.001	(0.003)
All	-0.019	(0.020)	-0.006	(0.008)	0.016	(0.018)	0.007	(0.008)	0.002	(0.002)
over 15 years										
Male	-0.045*	(0.026)	-0.012	(0.011)	0.039	(0.024)	0.014	(0.010)	0.004	(0.003)
Female	-0.015	(0.024)	-0.007	(0.012)	0.013	(0.022)	0.007	(0.012)	0.001	(0.003)
All	-0.029*	(0.018)	-0.011	(0.008)	0.026	(0.016)	0.011	(0.008)	0.003	(0.002)

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Sample size: 13,857 male, 13,852 female. Standard errors clustered at the individual level are shown in parentheses. Marginal effects are reported for each health status and are shown with gender separation and together. Results for each gender are from one regression estimated. All regressions are estimated using random effects ordered probit. Marginal effects are reported.

Appendix 1E

Table 1E.1: Unadjusted p-value versus Westfall-Young adjusted p-value

	Mental Health Problem Index			Poor Health (0 or 1)			BMI			If sick (in the last three months, 0 or 1)			Medical Expenditure		
	Unadjusted p-value	Westfall-Young adjusted p-value	Westfall-Young p-value	Unadjusted p-value	Westfall-Young adjusted p-value	Westfall-Young p-value	Unadjusted p-value	Westfall-Young adjusted p-value	Westfall-Young p-value	Unadjusted p-value	Westfall-Young adjusted p-value	Westfall-Young p-value	Unadjusted p-value	Westfall-Young adjusted p-value	Westfall-Young p-value
Land expropriation	0.340	0.167	0.202	0.167	0.172	0.167	0.061	0.033	0.011	0.000					

Table 1E.2: Health outcomes with time interval since displacement - two-interval version

	Mental Health Problem Index	Poor Health (0 or 1)	BMI	If sick (in the last three months, 0 or 1)	Medical Expenditure	Sample size (EXP=1)	Sample size (EXP=0)
0-8 years	0.038 (0.051)	-0.003 (0.232)	0.252 (0.178)	0.003 (0.014)	180.398 (136.140)	838	26,871
over 8 yeears	0.020 (0.055)	0.296 (0.223)	0.167 (0.180)	0.034** (0.017)	203.173 (137.090)	887	26,822

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Standard errors clustered at the individual level are shown in parentheses. Marginal effects are reported for each health outcome of interest. The estimation method for each health outcome is the same as in Table 1.6.

References

- Baji, P. and Bíró, A. (2018) Adaptation or recovery after health shocks? Evidence using subjective and objective health measures. *Health Economics*, 27(5), 850-864.
- Brunello, G., Fabbri, D., and Fort, M. (2013). The causal effect of education on bodymass: Evidence from Europe. *Journal of Labor Economics*, 31(1), 195-223.
- Cao, Y., Hwang, S. S. and Xi, J. (2012). Project-induced Displacement, Secondary Stressors, and Health. *Social Science & Medicine*, 74(7), 1130-1138.
- Chan, K. W. and Zhang, L. (1999). The Hukou System and Rural-Urban Migration in China: Processes and Changes. *The China quarterly*, 160, 818-855.
- Chen, B. (2016). The History and Future of China Urbanization. State Information Center, China. <http://www.sic.gov.cn/News/455/6167.htm> (in Chinese)
- Chow, E.N. and Chen, K. (1994). The impact of the one-child policy on women and the patriarchal family in the People's Republic of China. In E. N. Chow and C. W. Berheide (Eds.), *Women, the Family, and Policy: A Global Perspective*. SUNY Press.
- Cubí-Mollá, P., Jofre-Bonet, M., Serra-Sastre, V., 2017. Adaptation to health states: sick yet better off? *Health Economics*, 26 (12), 1826–1843.
- Eisenman, D.P., Gelberg, L., Liu, H. and Shapiro, M. F. (2003). Mental health and health-related quality of life among adult Latino primary care patients living in the United States with previous exposure to political violence. *Journal of the American Medical Association*, 290 (5), 627-634.
- Fazel, M., Wheeler, J. and Danesh, J. (2005). Prevalence of serious mental disorder in 7000 refugees resettled in western countries: a systematic review. *Lancet*, 365(9467), 1309-1314.
- Fussell, E. and Lowe, S. R. (2014). The impact of housing displacement on the mental health of low-income parents after Hurricane Katrina. *Social Science & Medicine*, 113(2014), 137-144.
- Gilbert, D. (2006). *Stumbling on happiness*. Alfred A. Knopf.
- He, S., Liu, Y., Webster, C. and Wu, F. (2009). Property Rights Redistribution, Entitlement Failure and the Impoverishment of Landless Farmers in China. *Urban*

Studies, 46 (9), 1925-1949.

Heckman, James J. (1979). Sample Selection Bias as Specification Error. *Econometrica*, 47 (1), 153-161.

Hui, E. C. M., Bao, H. J. and Zhang, X. L. (2013). The policy and praxis of compensation for land expropriations in China: An appraisal from the perspective of social exclusion. *Land Use Policy*, 32 (2013), 309-316.

Hwang, SS., Xi, J., Cao, Y., Feng, X. and Qiao, X. (2007). Anticipation of migration and psychological stress and the Three Gorges Dam project, China. *Social science & medicine*, 65 (5), 1012-1024.

Hwang, SS., Cao, Y. and Xi, J. (2010). Project-induced migration and depression: A panel analysis. *Social science & medicine*, 70 (11), 1765-1772.

Institute of Labor Economics (IZA). Australian National University, the University of Queensland and the Beijing Normal University (2014), Longitudinal Survey on Rural Urban Migration in China (RUMiC) (2008 – 2009). International Data Service Center of IZA (IDSC). Version 1.0. doi:10.15185/izadp.7680.1

Jing, L. (2009). China's New Labor Contract Law and Protection of Workers. *Fordham International Law Journal*, 32(3), 1083-1132.

Kasteler, J. M., Gray, R. M. and Carruth, M. L. (1968). Involuntary relocation of the elderly. *The Gerontologist*, 8(4), 276-279.

Liu, Z.Q. (2005). Institution and inequality: the hukou system in China. *Journal of Comparative Economics*, 33 (1), 133-157.

Liu, S., Chen, Z. and Chen, G. (2004). Impact of Urbanization and Resettlement on Employment of Rural Inhabitants in China: A Case Study in Changling Town, Chongqing Municipality, Three Gorges Reservoir (TGR) Area. *Mountain Research and Development*, 24(3), 228-233.

Porter, M., Haslam, N., (2005). Predisplacement and postdisplacement factors associated with mental health of refugees and internally displaced persons: a meta-analysis. *Journal of the American Medical Association*, 294 (5), 602-612.

Randell, H. (2016). The short-term impacts of development-induced displacement on wealth and subjective well-being in the Brazilian Amazon. *World Development*, 87, 385-400.

- Rubin, D.B. (1979). Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies. *Journal of the American Statistical Association*, 74(366), 318-328.
- Tang, W. S. (1997). Urbanization in China: a review of its causal mechanisms and spatial relations. *Progress in Planning*, 48(1), 1-65.
- Tani, M. (2017). Hukou Changes and Subjective Well-Being in China. *Social Indicators Research*, 132(1), 47-61.
- Tolley, G. S. (1987). Urbanization and economic development. In G. S. Tolley and V. Thomas (Eds.), *The Economics of Urbanization and Urban Policies in Developing Countries (A World Bank Symposium)* (pp. 15-31). The World Bank.
- Wang, C., Akguc, M., Liu, X. and Tani, M. (2020). Expropriation with hukou change and labour market outcomes in China. *China Economic Review*, 60, 101391
- Wang, Q. (2014). Effects of urbanisation on energy consumption in China. *Energy Policy*, 65(2014), 332-339.
- Westfall, P.H., Young, S.S. and Wright, S.P. (1993). On Adjusting P-Values for Multiplicity. *Biometrics*, 49(3), 941-945.
- WHO expert consultation (2004). Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. *Lancet*, 363(9403), 157-163.
- Wong, D.F.K., Chang, Y.L. and He, X.S. (2007). Rural migrant workers in urban China: living a marginalised life. *International journal of social welfare*, 16(1), 32-40.
- Xi, J., Hwang, SS. and Drentea, P. (2013). Experiencing a Forced Relocation at Different Life Stages: The Effects of China's Three Gorges Project-induced Relocation on Depression. *Society and Mental Health*, 3(1), 59-76.
- Zheng, Y. (2009). It's Not What is on Paper, but What is in Practice: China's New Labor Contract Law and the Enforcement Problem. *Washington University Global Studies Law Review*, 8(3), 595-618.

Chapter 2

Impact of forced displacement on mental well-being

2.1 Introduction

Being one of the largest hydraulic engineering projects in China, the Three Gorges Dam (TGD) on the Yangtze River, located in the Hubei province, was built with the objectives of controlling floods, generating hydroelectric power and improving navigation on the river ([Stone, 2008](#)). However, this project required submerging a populated area in China, thus leading to the relocation of 1.3 million people, the majority of whom were from Hubei Province and Chongqing Municipality. ([Jackson and Sleight, 2000](#); [Xi et al., 2007](#))

As 1.3 million people were forced to migrate due to the Three Gorges Dam project (TGDP), how to displace people with minimum impact on their lives became a major challenge to policymakers. This project was conceived as a way to improve the livelihood of the general population in China. However, if the people affected by the construction cannot be resettled appropriately and suffer long-term distress due to the displacement, we cannot say this project is a success for all people. Mitigating the negative consequence of this sort of project is important for maximising the benefits. As China has been developing such water conservancy projects, and more people will be affected by them, it is crucial to understand how such projects will affect people's

well-being and what we can learn from them. Although the government intended a “developmental resettlement” for TGDP, which means that displacement should not be detrimental for displaced people ([Padovani, 2016](#)), anecdotal evidence suggests that a certain number of TGDP migrants returned to the region where they came from due to their inability to adapt to their new environment ([Li, 2009](#); [Peng, 2011](#)). Although the relocatees were provided with housing, land or even jobs as compensation for their loss, many of them found life more difficult ([Li et al., 2001](#)). The duration of such feelings could result in mental ill-health. The majority of the literature focuses on the impact of displacement on mental health but regards the displaced group as homogeneous. However, individual circumstances may play a significant role in the process of displacement and adaptation, a factor that is not often considered.

First, due to the mountainous location, limited land in nearby areas of TGDP was available for the relocatees. As a result, some relocatees had to migrate far away from their origin, facing more challenges (economically and mentally) than relocatees who migrated closer to their original homes ([Li et al., 2001](#); [Feng, 2004](#)). Second, whether moving together with acquaintances or moving with their household alone may also affect the well-being of relocatees in different ways ([Li, 1994](#)).

Third, in addition to the factors related to displacement itself, relocatees may have different displacement experiences due to various personal backgrounds, which could also contribute to different mental well-being statuses. For example, the majority of the literature shows that women are likely to suffer more than men after traumatic experiences or events such as forced displacement ([Araya et al., 2007](#); [Taha and Sijbrandij, 2021](#); [Shevlin et al., 2017](#); [Sastry and VanLandingham, 2009](#); [Mujeeb, 2015](#)), while some found no differences between women and men, such as [Salah et al.\(2013\)](#). In addition, evidence suggests that education might be correlated with psychological well-being and migrants’ adaptation to their new life at the location of the destination ([Glenn and Weaver, 1981](#); [Kessler, 1982](#); [Padovani, 2016](#)). This correlation could come from various sources, including the level of income (people with

a higher level of education usually have higher incomes which could help to secure a less stressful life) and awareness of mental well-being (people who have higher levels of education usually are more likely to be aware of, and thus take more care of, their mental well-being).

Finally, evidence also suggests a correlation between mental health and marital status, although the direction of the impact is not straightforward ([Booth and Amato, 1991](#); [Glenn, 1975](#); [Wade and Pevalin, 2004](#)). Therefore, displacement may have differential impacts on mental well-being among those with different education levels and marital statuses.

This chapter hypothesizes that displacement due to the Three Gorges Dam project will negatively affect people's mental well-being on average. The outcome measures for mental well-being include depression, anxiety, low mood and low self-esteem, which are four factors derived from 20 questions in [Table 2A.1](#) by principle component factor analysis. For overall impact, depression, anxiety and low mood are expected to be adversely affected by displacement as these reflect people's distress for displacement and the negative feelings this generates. On the other hand, we may not find an adverse impact on low self-esteem because this policy is an exogenous shock to where people live and potentially who they live with.

As people are from various backgrounds, many factors could affect their ability and mentality to handle displacement. Women's mental well-being is more affected than men's due to gender differences in dealing with such life-changing events. Negative impacts are expected for those with a lower education level because less education means more difficulty securing a job, especially in a less familiar environment. People who were widowed and married are also more likely to suffer more from displacement as those who were widowed would not want to leave the memories with their previous partner or their familiar setting and established networks. Married people would not be happy to change after establishing their family, given the distress of getting new jobs and finding schools for children. In addition, those who were displaced farther

away from their original home would be more likely to feel homesick. Finally, those who moved with individual households would find it more challenging to adapt to the new life than those who moved together with village neighbours.

The results of heterogeneous impacts of this study show that variations in moving distance, moving models, education and marital status of relocatees could result in differential impacts of displacement.

This study fits in several lines of research. First, it contributes to the literature on displacement and resettlement ([Eisenman et al., 2003](#); [Fussell and Lowe, 2014](#); [Fazel et al., 2005](#); [Porter and Haslam, 2005](#); [Kasteler et al., 1968](#); [Randell, 2016](#)). Second, it adds to the literature on displacement due to TGDP, which is mainly qualitative ([McDonald-Wilmsen, 2009](#); [Wilmsen et al., 2011](#); [Duan and Wilmsen, 2012](#); [Wilmsen, 2016](#); [Wilmsen and Van Hulten, 2017](#); [Wilmsen, 2018](#); [Tan et al., 2005](#)). These studies are based on simple comparisons between pre- and post-displacement and therefore do not provide convincing results for the causal impacts of displacement.

A series of studies on TGD displacement using the same survey data as this study have found a negative impact on the mental well-being of relocatees. Comparing the non-migrants and the designated migrants who have already been displaced, [Hwang et al. \(2011\)](#) found negative and statistically significant short-term impacts of the forced displacement due to TGDP on mental and physical health. [Hwang et al. \(2010\)](#) explored the causal relationship between forced displacement and depression by comparing non-migrants and designated migrants regardless of whether or not they had moved. Their results showed that forced migration could aggravate depression levels. And this effect may come directly from forced migration; it may also come indirectly through other channels. Following the “stress process model” and Cernea’s “Impoverishment Risks and Reconstruction model”, [Cao et al. \(2012\)](#) found a positive correlation between displacement and migrants’ depression levels and that the mechanisms could be changing “social integration”, “socioeconomic status” and “community resources”. However, these studies consider relocatees as a homoge-

neous group and fail to consider the possible differential impacts on subgroups with various characteristics. [Xi et al. \(2013\)](#) explored differential impacts for different age groups - 18 to 30 years old, 31 to 50 years old, and over 50 years old - and found no significant heterogeneous impacts. To the best of my knowledge, this study is the only quantitative study addressing several dimensions of heterogeneous impacts of TGDP.

This paper contributes to the literature in two ways. First, by looking at both the average intention-to-treat effect and the average treatment effect on the treated, it finds that the experience of displacement is more detrimental to depression levels than the anticipation of displacement. Second, it quantifies the differential impacts of displacement due to the TGDP on the mental well-being of various subgroups of people in terms of their gender, levels of education, marital statuses, displacement within the town versus out of town, and displacement with the village versus individually.

For the rest of this chapter, Section 2.2 introduces background information on this study, and Section 2.3 describes the data. Section 2.4 demonstrates the empirical framework, and Section 2.5 presents estimates of both the average and heterogeneous impacts of TGDP. Section 2.6 shows robustness checks with propensity score matching and discusses the limitations of this study. Finally, Section 2.8 concludes.

2.2 Background Information

The administration of displacement was managed from the top down: the local authority followed the instruction from the provincial government, which was, in turn, guided by the central government. The TGDP progressed in four stages ([Padovani, 2016](#)). During the first stage (1992-1997), people who lived on Sandouping Island and close to the foundation of TGD were displaced within the same district but to a higher altitude. During the second stage (1998-2003), the water was raised up to 135

metres¹. Then the water was raised up to 156 metres in the third stage (2004-2006), followed by a final rise to 175 metres in the fourth stage (2007-2009). In this way, the relocatees could be displaced at different times, which helped release some pressure from this mass displacement programme and allowed the government to have enough time to make plans for their housing, land and employment.² As a result, the designated relocatees in the survey did not have to move simultaneously. Some of them had not yet moved by the second wave of the survey in 2006. In any case, residents living below 175 metres in the affected areas had to relocate eventually.

The initial idea was to relocate people within the same area or at least within the same provincial district. Hubei and Chongqing are the two provincial districts affected by TGDP, with Chongqing being affected more. As a consequence, Hubei was able to relocate all its relocatees within the province, while Chongqing could not. Relocatees from Chongqing had three options regarding where to relocate: move uphill (close to their original homes), move to live with relatives or relocate outside the province (Padovani, 2016). As our sample is from Chongqing, the individuals in the sample faced all these choices for resettlement. The relocatees had to deal with numerous situations in their destination, and their well-being largely depended on many factors, such as the culture at the destination and the skills of the relocatees.

2.3 Data

The data used for this study include two waves of interviews: the first wave is from 2003, and the second is from 2006.³ From Hwang et al., 2007, the first wave of interviews started in 2002 and was completed in 2003; 975 designated migrants and 555 non-migrants responded to the survey. They came from five randomly selected

¹Although Padovani (2016) indicates 139 metres, the Chinese government official website indicates 135 metres. https://china.gov.cn.admin.kyber.vip/jrzg/2006-10/27/content_425537.htm

²Chinese government website. 2006 Introduction to the Three Gorges Dam Project. http://www.gov.cn/test/2006-05/12/content_278989.htm

³The Three Gorges Dam migrants survey data is kindly shared by Dr Juan Xi.

communities in the “Wanxian Relocation and Development Region” (80% of the designated migrants lived in this region). The interview response rate for the first wave was 99%. Although this rate is very high, it is regarded as normal for face-to-face interviews conducted in China (Hwang et al., 2007; Hwang et al., 2010; Hwang et al., 2011). Due to TGDP, residents who lived at altitudes lower than 175 metres above sea level had to relocate, while those who lived above did not. 1,070 respondents were successfully followed and interviewed again for the second wave of the survey in 2006 (success rate 70%).

The respondent from each household for the survey was usually the household head or anyone in the household who could answer survey questions well. The sample size for my analysis is 1,510 for the first wave and 1,023 for the second wave, after eliminating 14 ineligible respondents who had already migrated before the first survey but were included in the survey by mistake and 53 observations who did not answer the mental well-being survey questions. As mental health is the focus of this study, dropping these observations could be problematic; people with worse mental health conditions may be less likely to respond to related questions as they may be unwilling to reveal their symptoms. To check whether this was the case, I regressed the indicator of missing mental health response on the treatment status and the covariates (age, gender, whether they own their house, number of household members, household income, whether they were urban Hukou holder, and a dummy for missing income data). The results show that the interviewees who did not answer the mental health questions were not significantly different from those who responded regarding treatment status and the controlled characteristics (Table 2.1), which alleviates my worries concerning sample selection after dropping these observations.

Summary statistics of the sample can be found in Tables 2.2-2.4. Table 2.2 displays the summary statistics of the outcome variables of interest, four standardised mental health factors (anxiety, depression, low mood and low self-esteem), in Wave 1 and Wave 2 separately. The mean and standard deviation of each mental health factor

Table 2.1: Correlations between missing mental health information and individual characteristics

	(1)
Treat	0.002 (0.006)
Age	0.000 (0.000)
Male	-0.004 (0.006)
House owner	0.014 (0.009)
Number of household members	0.001 (0.002)
Household income	-0.006 (0.010)
Urban Hukou holder	-0.001 (0.006)
Dummy for missing income data	-0.014 (0.017)
No. of Obs.	1,530
Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Column 1 presents the coefficients of the regression of missing mental health indicators and personal and household characteristics.	

Table 2.2: Summary Statistics - Standardized Mental Health Factors

		No. of Obs.	Mean	Std. Dev.	Min	Max
1st wave	Anxiety	1,510	0.020	1.021	-3.340	3.378
	Depression	1,510	0.051	1.015	-4.106	2.043
	Low mood	1,510	0.062	1.051	-3.687	2.104
	Low self-esteem	1,510	0.035	1.026	-3.880	2.049
2nd wave	Anxiety	1,023	0.005	0.996	-3.042	2.752
	Depression	1,023	-0.103	0.995	-3.268	2.221
	Low mood	1,023	-0.095	0.968	-3.913	2.213
	Low self-esteem	1,023	-0.005	0.971	-3.416	2.091

are around 0 and 1, respectively. The mean of each factor in Wave 1 is higher than that in Wave 2, which shows that, on average, the individuals had a better mental health status in 2003 before anyone moved. The range of each mental health factor is between -4.5 and 3.5. [Table 2.3](#) shows summary statistics of individual and household characteristics. The individual characteristics pertain to the individual in the household who responded to the survey. In the baseline survey, the average age of respondents was 45. 45% were male, and 86% of respondents owned their house. The average number of household members was four. 51% were urban residents, and 49% were rural. As rural residents did not reside as concentrated as urban residents, it was necessary to oversample urban residents to lower expenses ([Hwang et al., 2007](#)). In the second wave, the sample characteristics look similar to the baseline survey. The only exception is that the proportion of urban residents (as defined by their location in Wave 1) dropped to 47% of the sample, which shows the rural residents were more likely to respond again in the second wave, while urban residents were missing, possibly due to death, migration or refusal to respond.

Table 2.3: Summary Statistics - Covariates

		No. of Obs.	Mean	Std. Dev.	Min	Max
1st wave	Age	1,510	45.060	13.821	15	97
	Male	1,510	0.447	0.497	0	1
	House owner	1,510	0.862	0.345	0	1
	Number of household members	1,510	3.869	1.431	1	9
	Household income	1,510	0.120	0.301	0.001	10
	Urban Hukou holder	1,510	0.511	0.500	0	1
2nd wave	Age	1,023	44.390	13.111	17	80
	Male	1,023	0.451	0.498	0	1
	House owner	1,023	0.879	0.327	0	1
	Number of household members	1,023	3.950	0.327	1	9
	Household income	1,023	0.117	0.334	0.001	10
	Urban Hukou holder	1,023	0.470	0.499	0	1

*Household income in 100,000 RMB.

[Table 2.4](#) summarises the characteristics of the first and second waves, on which

the heterogeneity analysis is based; these include gender, education, marital status, displacement distance and whether displaced alone. In both waves, around 45% are male, and 55% are female. More than 80% of the respondents have general education⁴, while more than 11% of the sample are illiterate, around 90% of the sample are married, and more than 5% are widowed. The proportion of designated migrants who would be displaced within the town increased from 31.2% in the first wave to 37.4% in the second wave, and the proportion of the designated migrants who would move with their individual household increased from 15.5% in the first wave to 19% in the second wave.

Table 2.5 shows the mean characteristics of the treatment group (Column 1) and the control group (Column 2) in Wave 1. The characteristics include gender, age, whether the individual owned a house, number of household members, household income and whether the individual held urban Hukou. To understand the difference between the treatment and control groups for each characteristic and whether these differences are statistically significant, I regress each characteristic on the indicator of the treatment group. Column 3 presents these regression coefficients, showing that in 2003 (pre-displacement period), the treatment group has more males who are less likely to own a house, have more people in the household, and are less likely to have an urban Hukou, which demonstrates the disadvantages of the designated migrants.

2.4 Empirical Framework

This section first explores the average effect of displacement on mental well-being, including the average intention-to-treat effect (ITT) and average treatment effect on the treated (ATT). Then the heterogeneous effect of the displacement is investigated.

All the regression analyses adjust for individual characteristics, including age and gender, and socioeconomic characteristics, including whether they own a house, num-

⁴Any education level from lower elementary to senior high school.

Table 2.4: Summary Statistics - Characteristics for heterogeneity analysis

	No. of Obs.		Percentage	
	First wave	Second wave	First wave	Second wave
Gender	1,510	1,023		
Male	675	460	44.7%	45.0%
Female	835	563	55.3%	55.0%
Education	1,510	1,023		
Illiterate	167	115	11.1%	11.2%
General Education	1,219	831	80.7%	81.2%
Technical Education	101	57	6.7%	5.6%
College	23	20	1.5%	2.0%
Marital Status	1,510	1,023		
Single	66	16	4.4%	1.6%
Married	1,341	932	88.8%	91.1%
Divorced	25	18	1.7%	1.8%
Cohabit	1	2	0.1%	0.2%
Widowed	77	55	5.1%	5.4%
Displacement Distance	898	741		
Out of town	131	127	14.6%	17.1%
Within town	280	277	31.2%	37.4%
No displacement	487	337	54.2%	45.5%
Displaced alone vs. not	1,356	752		
Alone	210	143	15.5%	19.0%
With village	531	189	39.2%	25.1%
Others	128	83	9.4%	11.0%
No displacement	487	337	35.9%	44.8%

Note: For education, the category “general education” includes any education level from lower elementary school to senior high school except for the technical high school, and the category “technical education” includes technical high school and technical college.

Table 2.5: Sample Balance

	(1)	(2)	(3)
	Treatment group mean	Control group mean	Difference treatment vs. control
Male	0.468 (0.499)	0.402 (0.491)	0.066** (0.027)
Age	45.451 (13.625)	44.240 (14.203)	1.210 (0.761)
House Owner	0.834 (0.372)	0.920 (0.272)	-0.097*** (0.022)
Number of household members	3.917 (1.522)	3.768 (1.213)	0.149* (0.079)
Household income	0.120 (0.348)	0.119 (0.163)	0.001 (0.017)
Urban Hukou holder	0.495 (0.500)	0.544 (0.499)	-0.050* (0.027)
Sample size	1,023	487	1,510

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. The results in column 3 are obtained by regressing the variables of interest on treatment status.

ber of household members, household income and whether they hold urban Hukou. People in different age ranges or gender are likely to react differently to change, especially events like displacement. Whether owning a house in China is usually treated as a symbol of stability in family life. It would be much easier for someone without their own house to move than those with one. Household income and the number of household members are related to how easily the household handles displacement. It would be easier to displace with higher income and fewer household members. Urban Hukou demonstrates which welfare system people were involved in, as there is a clear distinction between urban and rural welfare systems.

2.4.1 Average effect

I expect an anticipation effect exists: people were affected as soon as they knew they would be displaced before anyone was actually displaced. It is different from ITT, as

ITT demonstrates the impact of displacement on the designated relocatees regardless of whether or not they have moved. Differences are also expected between the anticipation effect and ATT: people who moved earlier were more negatively affected than the designated relocatees who had not yet moved, as these designated relocatees had more time to plan for displacement, and they also had not experienced the displacement yet, which was stressful itself. That is, the effect on the designated relocatees who had not yet moved originated from their anticipation of the displacement rather than from displacement itself. This study first estimates both ITT and ATT.

i. Difference-in-differences analysis to estimate average intention-to-treat effect

For ITT, the treatment group includes those who lived below 175m altitude, regardless of whether or not they had been displaced.⁵ The econometric model is as below:

$$MH_{it} = \alpha_i + \beta_1 Treat_i + \beta_2 Post_t + \beta_3 Treat_i * Post_t + \delta X_i + u_{it} \quad (2.1)$$

where MH represents a specific standardised mental health factor from principal component factor analysis, $Treat$ is a binary variable indicating the intention to treat group, that is, those living under 175 metres above sea level, $Post$ indicates the year 2006, and X includes age, gender, household income, whether individuals were registered as urban residents or rural residents (1 indicates they are urban residents), whether they owned a house, and number of household members in Wave 1.

Standardised Mental health factors: the outcomes of interest in this study relate to mental well-being, the analysis of which is constructed from 20 survey questions on mental health symptoms. These questions ask respondents whether they have experienced certain mental health symptoms in the past week, for example, whether

⁵From the data, 606 designated migrants were aware in Wave 1 that they would be displaced, and 80 designated migrants were not aware. Among them, 470 designated migrants knew which year exactly they would be displaced.

they felt scared, whether they did not feel like talking, whether they felt that they could not get going, and so on. Each question uses a 1-4 response format to show the severity of the symptom: 1, always; 2, sometimes; 3, rarely; 4, never. Unlike previous studies which used the CES-D scale⁶ to represent the mental health status (Hwang et al., 2007; Xi et al., 2012; Hwang et al., 2010; Cao et al., 2012; Hwang et al., 2011; Xi and Hwang, 2011a; Xi and Hwang, 2011b; Xi, 2016), this study uses principal component factor analysis to reduce the dimensionality of these 20 mental well-being questions to four uncorrelated standardised mental health factors: anxiety, depression, low mood and low self-esteem. The summary statistics for these four components are shown in Table 2.2. A higher value of factor implies better mental health condition in that factor. Detailed information about which factors are more likely to correspond to which mental health questions is available in Appendix 2A.

ii. Instrumented difference-in-differences analysis for the treatment effect on the treated

By 2006 (the second wave of the survey), around 40% of the designated relocatees in the sample had not yet moved. As I suppose the impact on the people who were expecting to be displaced could be different from that on the people who had already been displaced, I want to look at the treatment effect on the treated. The reason why some designated relocatees had not yet moved by 2006 could be that the phase of the TGD project allowed them to stay longer at their origin; Hwang et al. (2007) points out that when to relocate migrants depended on things such as the process of building the dam. However, I cannot completely rule out personal reasons for differences in the timing of displacement. For example, some people might have been reluctant to move and waited until the last moment because they missed their old house and the area, while some moved without any delay because they looked

⁶CES-D indicates “Center for Epidemiologic Studies Depression Scale”, a 20-item scale measuring depressive symptom severity. (Julian et al., 2011)

forward to more opportunities in the future. These unknown reasons could affect the time of moving and potentially be endogenous. If I simply use the group who have moved as the treatment group, it could be problematic due to self-selection into the treatment. To solve this problem, I employ instrumented difference-in-differences⁷ (DID-IV) approach (Hudson et al., 2017; de Chaisemartin, 2010; de Chaisemartin and D’Haultfoeuille, 2018) and estimate the model by two-stage least squares using $Post_t * Treat_i$ as an instrumental variable for ‘propensity to relocate’ (1 indicates that designated relocatees who have actually moved and 0 otherwise).

First stage equation:

$$PropensityToRelocate_{it} = \gamma_i + \pi_1 Post_t + \pi_2 Treat_i + \pi_3 Treat_i * Post_t + \eta X_i + \epsilon_{it} \quad (2.2)$$

where *PropensityToRelocate* is a binary variable indicating whether a designated relocatee has been displaced or not, *Post* indicates the year 2006, *Treat* represents residents who lived below 175 metres’ altitude, and *X* includes age, gender, household income, whether they were registered as urban residents or rural residents (1 indicates they were urban residents), whether they owned a house, and the number of household members.

Second stage equation:

$$MH_{it} = \alpha_i + \beta_1 Post_t + \beta_2 Treat_i + \beta_3 \widehat{PropensityToRelocate}_{it} + \delta X_i + u_{it} \quad (2.3)$$

where *MH* represents a specific standardised mental health factor derived from principal component factor analysis; *Post* indicates the year 2006; *Treat* represents residents who lived below 175 metres’ altitude; $\widehat{PropensityToRelocate}_{it}$ is the predicted propensity to relocate rate from the first-stage estimation; and *X* includes age, gender, household income, whether registered as urban residents, whether they own a

⁷Some places may call it “difference-in-differences instrumental variable”.

house, and the number of household members. For the coefficient β_3 to be a consistent estimate of ATT, $Treat * Post$ needs to satisfy the exclusion restriction, that is, it only needs to affect MH via propensity to relocate, not directly. A caveat worth noting is that this method is invalid if $Treat * Post$ is correlated with mental health status via anticipation effects, that is, if the mental health of those who had not yet moved but would have to move was affected.

2.4.2 Heterogeneous impact

The main focus of this study is the heterogeneous impact of displacement, as the expectation is that people with different characteristics could be affected differently. The econometric model to identify heterogeneous impact is:

$$\begin{aligned}
MH_{it} = & \alpha_0 + \sigma X_i + \beta_1 Treat_i + \beta_2 Post_t + \beta_3 W_i + \gamma_1 Treat_i * Post_t \\
& + \gamma_2 Post_t * W_i + \gamma_3 Treat_i * W_i \\
& + \delta_1 Treat_i * Post_t * W_i + u_{it}
\end{aligned} \tag{2.4}$$

where MH represents a specific standardised mental health factor from principal component factor analysis. $Post$ indicates the year 2006. $Treat$ represents residents who lived below 175 metres altitude. W represents characteristics by which I expect the impact of TGDP to vary as explained in Section 1, including gender, education, marital status, moving out of town versus within town, and moving with the village versus individually. X includes age, gender, household income, whether people were registered as urban or rural residents (1 if urban residents), whether they owned a house, and the number of household members. The parameter of interest is δ_1 , which shows the differential impact on a specific group compared with the base group.

2.5 Results

2.5.1 Average impact

i. Intention to treat

[Table 2.6](#) presents ITT estimates for depression, one of the standardised mental health factors. Four different specifications (columns 1-4) all show the level of depression gets worse. Column 1 shows the results with the full sample (including the individuals available in one wave only). There is a statistically significant change in depression (-0.151), meaning that the depression level increased by 15.1% of the standard deviation. With the reduced sample (only those individuals who were available in both waves), columns 2-4 show that depression became worse by 22.1% of standard deviation, statistically significant at the 5% level, slightly higher than the estimate using the full sample.

As the sample size in the second wave is smaller than that in the first wave, the issue of selection due to attrition could be a concern. Column 3 controls for the hazard rate - inverse Mills ratio ([Mills, 1926](#)) - to correct for attrition. Another approach to correct selection due to attrition is to apply Inverse Probability Weight ([Robins et al., 2000](#)) in the regression analysis. However, this chapter does not use Inverse Probability Weight. This chapter not only focuses on correcting for selection but also aims to understand whether attrition is concerning in the regression. As we do not find the statistical significance of the hazard rate, it demonstrates that selection due to attrition should not be too worrisome. Therefore, no other methods were explored. Furthermore, when comparing columns 2 and 3, no difference is found from the estimation of ITT. While columns 1-3 show Ordinary Least Square (OLS) estimates, column 4 shows the fixed effects (FE) estimate with the reduced sample. Adding FE controls for unobserved differences in changes in sample composition between treatment and control groups can improve precision (as it contributes to explaining the variation in mental health). The magnitude of the effect in column 4 is the same as

Table 2.6: Intention to Treat Effect - Depression

	(1) Full	(2) Reduced	(3) Reduced	(4) Reduced
Treat	-0.098* (0.054)	-0.034 (0.066)	-0.273* (0.164)	
Post	-0.060 (0.066)	-0.072 (0.073)	-0.072 (0.073)	-0.072 (0.068)
Treat*Post	-0.151* (0.082)	-0.221** (0.090)	-0.221** (0.090)	-0.221*** (0.084)
Hazard rate			6.839 (4.343)	
Age	-0.004** (0.002)	-0.003* (0.002)	-0.028* (0.016)	
Male	0.378*** (0.040)	0.360*** (0.044)	0.546*** (0.126)	
Urban Hukou holder	0.054 (0.048)	0.129** (0.054)	-0.760 (0.567)	
House owner	0.001 (0.062)	-0.018 (0.070)	-0.023 (0.070)	
Number of household members	0.033** (0.015)	0.040** (0.017)	0.041** (0.017)	
Household income	-0.0134 (0.040)	-0.046* (0.026)	-0.215* (0.111)	
Dummy for missing value (income)	-0.225* (0.133)	-0.254 (0.171)	-1.942* (1.103)	
Observations	2,533	2,018	2,018	2,018
R-squared	0.067	0.074	0.076	0.035
Fixed Effect	No	No	No	Yes
OLS	Yes	Yes	Yes	No
Number of id	1,524	1,009	1,009	1,009

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Column 1 presents results using the full sample. Columns 2-4 present results using the reduced sample where individuals who appear in the survey for one wave only are dropped.

the estimates in columns 2 and 3; however, the impact is significant at the 1% level, not the 5% level as in columns 2 and 3.

Table 2.7 presents the ITT estimates for the other three standardised mental health factors: anxiety, low mood and low self-esteem (see Tables B1-B3 in Appendix 2B for complete estimates). None of these ITT estimates is statistically significant.

All in all, the designated migrants, no matter whether they moved or not, became more depressed due to TGDP, but no evidence shows that their levels of anxiety, low mood and low self-esteem were significantly affected. Unexpectedly, we do not find a statistically significant impact on anxiety or low mood in the results. A possible explanation could be that the mental health questions corresponding to anxiety and low mood are related to their personality, while the questions corresponding to depression are more about how they feel deep inside. The former stays the same even when people face changes, while the latter does not. As depression has been found to lower economic growth, reduce workers' productivity and lead to the dysfunction of families (Holden, 2000), these consequences are running in the opposite direction of the goal of TGDP. Thus, policymakers should pay attention to preparing designated relocatees psychologically for displacement and ensure a well-communicated resettlement plan for the relocatees.

ii. Treatment effect on the treated

Table 2.8 shows ATT estimates in terms of depression. Columns 1, 3, 5 and 7 are the results for the first stage, and columns 2, 4, 6 and 8 are the results for the second stage. Similar to ITT, Table 2.8 presents four specifications to estimate the model. Columns 1 and 2 present estimates with the full sample (all observations, including those respondents who were available in one wave only). Columns 3 and 4 are estimates with a reduced sample (excluding those who were available in one wave only). As the reduced sample excludes all the respondents who were available in one wave only, the difference in estimates between the reduced sample and the full

Table 2.7: Intention to Treat Effect - Anxiety, Low Mood and Low Self-esteem

	(1) Full	(2) Reduced	(3) Reduced	(4) Reduced
Panel A: Anxiety				
Treat*Post	0.023 (0.084)	-0.052 (0.092)	-0.052 (0.092)	-0.052 (0.084)
Panel B: Low Mood				
Treat*Post	-0.013 (0.085)	-0.042 (0.093)	-0.042 (0.093)	-0.042 (0.083)
Panel C: Low Self-esteem				
Treat*Post	-0.035 (0.084)	-0.063 (0.092)	-0.063 (0.092)	-0.063 (0.086)
N	2,533	2,018	2,018	2,018
Fixed Effect	No	No	No	Yes
OLS	Yes	Yes	Yes	No
Number of id	1,524	1,009	1,009	1,009

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Column 1 presents results using the full sample. Columns 2-4 present results using the reduced sample where individuals who appear in the survey for one wave only are dropped. Control variables include age, gender, whether the individual owns a house, the number of household members, whether the individual is registered as an urban resident or rural resident (1 indicates urban resident) and household income.

sample could come from selection due to attrition. To check whether this selection is significant, columns 5 and 6 add a hazard rate to test and correct for selection. The statistically insignificant hazard rate suggests that selection due to attrition is not a big issue. However, the hazard rate is constructed from observed variables. This is why I have estimated the fixed effects model (columns 7 and 8) with the reduced sample, which controls for unobserved differences in changes in sample composition between treatment and control groups and can improve precision.

The first-stage estimates are very similar in all four specifications, with the estimate in the first specification showing that designated relocatees were 59.4% more likely to move by 2006, slightly lower than the estimate from the other three specifications (59.6% more like to move by 2006). All are significant at the 1% level, showing the instrumental variable ($Treat * Post$) is significantly correlated with displacement.

The second-stage estimates with the full sample (column 2) show increased levels of depression (25.4% of standard deviation) for the people who have already been displaced. This estimate, the smallest among all four specifications, is statistically significant at the 10% level. The estimates for the second (column 4) and third specifications (column 6) are the same: 37.1% of standard deviation higher in the levels of depression for the already displaced, significant at the 5% level. The difference in estimates between the full and reduced samples could be evidence of sample selection; however, as the hazard rate is statistically insignificant (as shown in column 6), this sample selection should not be worrisome. The fixed effects estimates have the same magnitude as columns 4 and 6 but are statistically significant at the 1% level. This confirms that attrition due to unobserved heterogeneity is not a concern.

[Table 2.9](#) presents the ATT estimates for levels of anxiety, low mood and low self-esteem. No evidence shows that levels of anxiety, low mood and low self-esteem were affected for the already displaced (see Tables 2B.4-2B.6 in [Appendix 2B](#) for complete estimates). The explanation could be similar to the results for ITT estimates above. However, as ATT estimates are even higher in absolute value than ITT estimates for

the impact on depression, i.e. people who have already been displaced became more depressed than those who have not, policymakers should pay extra attention to the resettlement of those who were already displaced.

iii. Effect of anticipation

As the individuals in the treatment group were already aware that they would have to be displaced in Wave 1, their mental well-being may be affected even in Wave 1 due to anticipation. To check whether this is the case, I regress standardised mental health factors on the indicator of designated migrants, using first-wave data only. The results are shown in [Table 2.10](#). The coefficient for designated migrants tells us that, before displacement, designated migrants were around 0.105 more anxious (10.5% of the standard deviation) and 0.156 lower in mood (15.6% of the standard deviation) compared with residents who were not expected to be displaced. No significant difference is found for depression or low self-esteem. Higher anxiety and lower mood are understandable because they are facing an unpredictable future, and they may feel frustrated to experience these uncertainties. However, as they have not moved yet, depression and low self-esteem have not come into effect.

Due to the anticipation effect on anxiety and low mood, one issue is worth mentioning for using the DID-IV approach to estimate the ATT effect. As mentioned earlier, this association between being designated migrants and mental health status via anticipation effects leads to invalidity of DID-IV approach, because the variation in the propensity to relocate that is explained by eligibility (being designated migrants) would also be correlated with mental health status. As a result, the exclusion restriction condition is not satisfied. Therefore, I cannot draw a conclusion about the impact on anxiety or low mood but only draw a conclusion about the impact on depression and low self-esteem.

Table 2.8: Treatment Effect on the Treated - Depression

	(1) Propensity to relocate Full	(2) Depression Full	(3) Propensity to relocate Reduced	(4) Depression Reduced	(5) Propensity to relocate Reduced	(6) Depression Reduced	(7) Propensity to relocate Reduced	(8) Depression Reduced
Post	0.000 (0.002)	-0.060 (0.065)	0.000 (0.003)	-0.072 (0.073)	0.000 (0.003)	-0.072 (0.073)	0.000 -	-0.072 (0.068)
Treat	-0.005* (0.003)	-0.100* (0.053)	-0.007* (0.004)	-0.036 (0.065)	0.019 (0.042)	-0.266 (0.162)		
Treat*Post	0.594*** (0.019)		0.596*** (0.019)		0.596*** (0.019)		0.596*** (0.019)	
Predicted Propensity to Relocate		-0.254* (0.137)		-0.371** (0.151)		-0.371** (0.151)		-0.371*** (0.141)
Hazard rate					-0.742 (1.206)	6.563 (4.271)		
Observations	2,533	2,533	2,018	2,018	2,018	2,018	2,018	2,018
R-squared	0.527	0.071	0.51	0.077	0.51	0.078	0.596	
Fixed Effect	No	No	No	No	No	No	Yes	Yes
OLS	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Number of id	1,524	1,524	1,009	1,009	1,009	1,009	1,009	1,009

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Columns 1, 3, 5 and 7 present estimates of the first stage. Columns 2, 4, 6 and 8 present estimates of the second stage. Control variables include age, gender, whether the individual owns a house, the number of household members, whether the individual is registered as an urban resident or rural resident (1 indicates urban resident) and household income.

Table 2.9: Treatment Effect on the Treated - Anxiety, Low Mood and Low Self-esteem

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Propensity to relocate Full	MH factors Full	Propensity to relocate Reduced	MH factors Reduced	Propensity to relocate Reduced	MH factors Reduced	Propensity to relocate Reduced	MH factors Reduced
First Stage								
Treat*Post	0.594*** (0.019)		0.596*** (0.019)		0.596*** (0.019)		0.596*** (0.019)	
Second Stage								
Panel A: Anxiety								
Predicted		0.038		-0.088		-0.088		-0.088
Propensity to relocate		(0.141)		(0.154)		(0.154)		(0.140)
Panel B: Low Mood								
Predicted		-0.022		-0.071		-0.071		-0.071
Propensity to relocate		(0.142)		(0.154)		(0.154)		(0.140)
Panel C: Low Self-esteem								
Predicted		-0.059		-0.106		-0.106		-0.106
Propensity to relocate		(0.141)		(0.154)		(0.154)		(0.144)
N	2,533	2,533	2,018	2,018	2,018	2,018	2,018	2,018
Fixed Effect	No	No	No	No	No	No	Yes	Yes
OLS	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Number of id	1,524	1,524	1,009	1,009	1,009	1,009	1,009	1,009

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Columns 1, 3, 5 and 7 present estimates of the first stage. Columns 2, 4, 6 and 8 present estimates of the second stage. Control variables include age, gender, whether the individual owns a house, the number of household members, whether the individual is registered as an urban resident or rural resident (1 indicates urban resident) and household income.

Table 2.10: Anticipation

	(1) Anxiety	(2) Depression	(3) Low mood	(4) Low self-esteem
Designated migrants	-0.105* (0.056)	-0.076 (0.054)	-0.156*** (0.058)	-0.007 (0.055)
Observations	1,510	1,510	1,510	1,510
R-squared	0.002	0.001	0.005	0

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Data for the first wave only is used for these estimates.

2.5.2 Heterogeneous impact

i. Gender

Table 2.11 shows heterogeneous impacts between genders, with the female as the reference group. For females alone, no significant impact was found in the level of depression, anxiety, low mood and low self-esteem. Compared with the female, no significant difference in impact on males for depression and anxiety was found. However, males' low mood was more negatively affected by displacement than females: 36.1% and 37% of standard deviation worse with the full sample and reduced sample, respectively. Male was also more negatively affected in their self-esteem with the reduced sample: 31.5% lower than female.

These results differ from what we expected, especially for the impact on low mood. A possible explanation is that the main source of distress for men comes from work while the main source of distress for women comes from family (Lai, 1995). Displacement not only affects family life but also affects livelihood, which is a concern for men, especially when the required skillset changes.

Table 2.11: Gender

	(1) Full	(2) Reduced	(3) Reduced
Panel A: Depression			
Treat*Post	-0.131 (0.116)	-0.188 (0.127)	-0.188 (0.119)
Treat*Post*Male	-0.04 (0.162)	-0.082 (0.178)	-0.082 (0.166)
Panel B: Anxiety			
Treat*Post	-0.03 (0.112)	-0.087 (0.123)	-0.087 (0.109)
Treat*Post*Male	0.138 (0.169)	0.091 (0.186)	0.091 (0.169)
Panel C: Low Mood			
Treat*Post	0.145 (0.111)	0.12 (0.121)	0.12 (0.109)
Treat*Post*Male	-0.361** (0.173)	-0.370* (0.189)	-0.370** (0.170)
Panel D: Low Self-esteem			
Treat*Post	0.068 (0.113)	0.072 (0.124)	0.072 (0.110)
Treat*Post*Male	-0.239 (0.169)	-0.315* (0.185)	-0.315* (0.176)
N	2,533	2,018	2,018
Fixed Effect	No	No	Yes
OLS	Yes	Yes	No
Number of id	1,524	1,009	1,009

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. The reference group is female. Control variables include age, whether the individual owns a house, the number of household members, whether the individual is registered as an urban resident or rural resident (1 indicates urban resident) and household income.

ii. Education

[Table 2.12](#) presents estimates of heterogeneous impacts along education level, with those obtained general education as the reference group. For the people with general education, the level of depression was affected badly by displacement: 23.2% and 25.1% of standard deviation worse with full sample and reduced sample, respectively. However, levels of anxiety, low mood and low self-esteem were not affected. People with technical education (technical high school or technical college) or college education were not affected differently in all four standardised mental health factors.

Illiterate individuals' anxiety levels became higher compared with people with general education, as we expected: 53.5% and 57.2% of standard deviation in the full and reduced sample, respectively. However, for the illiterate, the heterogeneous impact on depression was only evident in the full sample (58.1% of standard deviation), meaning the illiterate were not as negatively affected as those who had obtained general education. This significant heterogeneity disappears in the estimates with the reduced sample. After dropping 34% of illiterate respondents with the reduced sample, no significant heterogeneity between the illiterate individuals and the general educated is identified, which suggests sample selection exists.

From these results, heterogeneity in impacts on anxiety exists between the educated and the illiterate, but not much between those with different levels of education. This is in line with the results of [Liu et al. \(2013\)](#) which found that schizophrenia is more likely to develop in the illiterate than in literate individuals. However, anxiety is of course a much milder mental illness than schizophrenia.

iii. Marriage

[Table 2.13](#) shows heterogeneous impacts related to marital status, with the married as the reference group. The married became more depressed (18% and 21.8% of standard deviation in the full and reduced samples, respectively), and these differences are statistically significant at the 5% level, which is in line with the hypothesis. However,

Table 2.12: Education

	(1) Full	(2) Reduced	(3) Reduced
Panel A: Depression			
Treat*Post	-0.232**	-0.251**	-0.251***
	-0.092	-0.1	-0.095
Treat*Post*Illiteracy	0.581**	0.200	0.200
	-0.279	-0.322	-0.285
Treat*Post*Technical Edu	0.279	0.23	0.23
	-0.287	-0.369	-0.334
Treat*Post*College	0.088	-0.074	-0.074
	-0.351	-0.396	-0.239
Panel B: Anxiety			
Treat*Post	0.086	0.027	0.027
	-0.095	-0.105	-0.094
Treat*Post*Illiteracy	-0.535**	-0.572**	-0.572**
	-0.261	-0.277	-0.267
Treat*Post*Technical Edu	-0.197	-0.458	-0.458
	-0.339	-0.368	-0.322
Treat*Post*College	0.629	0.772	0.772
	-0.621	-0.609	-0.507
Panel C: Low Mood			
Treat*Post	-0.056	-0.074	-0.074
	-0.095	-0.104	-0.096
Treat*Post*Illiteracy	0.182	0.364	0.364
	-0.274	-0.287	-0.248
Treat*Post*Technical Edu	0.147	-0.310	-0.310
	-0.333	-0.375	-0.285
Treat*Post*College	0.506	0.390	0.390
	-0.643	-0.684	-0.358
Panel D: Low Self-esteem			
Treat*Post	-0.057	-0.087	-0.087
	-0.092	-0.100	-0.094
Treat*Post*Illiteracy	0.111	0.235	0.235
	-0.317	-0.339	-0.272
Treat*Post*Technical Edu	-0.198	-0.334	-0.334
	-0.312	-0.377	-0.389
Treat*Post*College	0.302	0.547	0.547
	-0.651	-0.684	-0.772
N	2,533	2,018	2,018
Fixed Effect	No	No	Yes
OLS	Yes	Yes	No
Number of id	1,524	1,009	1,009

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. The reference group contains the people whose level of education belongs to general education. Control variables include age, gender, whether the individual owns a house, the number of household members, whether the individual is registered as an urban resident or rural resident (1 indicates urban resident) and household income.

the impacts on levels of anxiety, low mood and low self-esteem are insignificant for the married respondents. The reason could be that anxiety, low mood and low self-esteem do not directly reflect their sadness once they understand their home will no longer exist when the TGDP is completed.

The single and widowed respondents were not significantly different from the married individuals for all four standardised mental health factors. However, the impact on the level of low mood was significantly better for the divorced than the married (136% and 268.9% of standard deviation in the full and reduced samples, respectively); the estimates using the reduced sample is stronger (statistically significant at the 1% level, compared with the 10% level using the full sample). The impact on depression was significantly better for the divorced with the full sample only (114.8% of standard deviation), and the impacts on anxiety (-1.160) and low self-esteem (-1.387) are significantly worse with the reduced sample only.

The different results between the full sample and the reduced sample could come from sample selection. There are 25 divorced respondents in the full sample versus 11 in the reduced sample; that is, more than half of the divorced respondents in the full sample are dropped. The much smaller number of divorced respondents compared with the numbers of married (1,341), single (66) and widowed (77) individuals reflects the reality that divorce was a rare case in China. ([Liao and Heaton, 1992](#); [Zeng and Wu, 2000](#)).

In China, especially in less developed regions, divorce could be a traumatic experience because it not only reflects the failure of the marriage but also damages reputation ([Liao and Heaton, 1992](#)). Thus, when the divorced individuals moved, they had an opportunity to escape the judgement from their peers. However, they also experienced the adverse effects of displacement. The results reflect these mixed experiences of the divorced. Compared with the married, although the impact of TGDP for the divorced became worse on anxiety (116% of standard deviation with the reduced sample) and self-esteem (139% of standard deviation with the reduced

Table 2.13: Marital Status

	(1) Full	(2) Reduced	(3) Reduced
Panel A: Depression			
Treat*Post	-0.180** (0.086)	-0.218** (0.095)	-0.218** (0.087)
Treat*Post*Single	-0.476 (0.330)	-0.543 (0.392)	-0.543 (0.333)
Treat*Post*Divorced	1.448* (0.843)	0.745 (0.882)	0.745 (1.039)
Treat*Post*Widowed	0.557 (0.481)	0.306 (0.545)	0.306 (0.635)
Panel B: Anxiety			
Treat*Post	0.014 (0.089)	-0.053 (0.097)	-0.053 (0.088)
Treat*Post*Single	0.331 (0.397)	-0.048 (0.461)	-0.048 (0.384)
Treat*Post*Divorced	0.753 (0.752)	-1.160* (0.665)	-1.160* (0.704)
Treat*Post*Widowed	0.028 (0.410)	0.282 (0.481)	0.282 (0.336)
Panel C: Low Mood			
Treat*Post	-0.032 (0.089)	-0.070 (0.096)	-0.070 (0.086)
Treat*Post*Single	0.304 (0.415)	0.460 (0.455)	0.460 (0.390)
Treat*Post*Divorced	1.360* (0.786)	2.689*** (0.801)	2.689*** (0.796)
Treat*Post*Widowed	-0.415 (0.423)	0.680 (0.443)	0.680 (0.422)
Panel D: Low Self-esteem			
Treat*Post	-0.035 (0.088)	-0.056 (0.097)	-0.056 (0.090)
Treat*Post*Single	0.115 (0.360)	0.149 (0.389)	0.149 (0.394)
Treat*Post*Divorced	-0.829 (0.688)	-1.387* (0.821)	-1.387** (0.607)
Treat*Post*Widowed	0.041 (0.494)	0.062 (0.547)	0.062 (0.529)
N	2,533	2,018	2,018
Fixed Effect	No	No	Yes
OLS	Yes	Yes	No
Number of id	1,524	1,009	1,009

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. The reference group contains the people who were married. Control variables include age, gender, whether the individual owns a house, the number of household members, whether the individual is registered as an urban resident or rural resident (1 indicates urban resident) and household income.

sample), it became better in depression (145% of standard deviation with the full sample) and low mood (136% of standard deviation with the full sample and 269% with the reduced sample).

iv. Moving distance

Table 2.14 reports estimates of heterogeneity in moving distance for four standardised mental health factors, with the out-of-town movers as the reference group. Out-of-town movers became more anxious (31.6% and 37.8% of standard deviation in the full and reduced samples, respectively) and depressed (69.1% and 68.5% of standard deviation in the full and reduced samples, respectively). Compared with the out-of-town movers, the within-town movers were not affected as badly, as demonstrated by significantly positive estimates for heterogeneous impact in both anxiety (49.1% of standard deviation in the full sample and 46.5% of standard deviation in the reduced sample) and depression (47.1% of standard deviation in the full sample and 48.8% of standard deviation in the reduced sample). These results are in line with the hypothesis. For the levels of low mood and low self-esteem, I do not find evidence of heterogeneity in the impact between out-of-town and within-town movers.

Why does the heterogeneity only appear in anxiety and depression, but not in low mood or low self-esteem? For low mood, it could come from the generality of survey questions. Questions corresponding to low mood are less concrete than questions relating to anxiety and depression. As a consequence, people may have less clear answers to them just by having various moving distances. For low self-esteem, the survey questions “always feel that other people are unfriendly to you” and “always feel that other people do not like you” both focus on others’ perception of oneself, and thus are unrelated to the distance of moving.

Table 2.14: Moving Distance

	(1) Full	(2) Reduced	(3) Reduced
Panel A: Depression			
Treat*Post	-0.691*** (0.141)	-0.685*** (0.146)	-0.685*** (0.144)
Treat*Post*Move within town	0.471*** (0.150)	0.488*** (0.151)	0.488*** (0.149)
Panel B: Anxiety			
Treat*Post	-0.316** (0.139)	-0.378*** (0.143)	-0.378*** (0.133)
Treat*Post*Move within town	0.491*** (0.147)	0.465*** (0.149)	0.465*** (0.137)
Panel C: Low Mood			
Treat*Post	-0.033 (0.133)	-0.015 (0.138)	-0.015 (0.133)
Treat*Post*Move within town	-0.104 (0.142)	-0.109 (0.144)	-0.109 (0.138)
Panel D: Low Self-esteem			
Treat*Post	-0.047 (0.140)	-0.102 (0.143)	-0.102 (0.136)
Treat*Post*Move within town	-0.232 (0.148)	-0.202 (0.148)	-0.202 (0.143)
N	1,639	1,466	1,466
Fixed Effect	No	No	Yes
OLS	Yes	Yes	No
Number of id	906	733	733

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. The reference group contains the people who moved out of town. Control variables include age, gender, whether the individual owns a house, the number of household members, whether the individual is registered as an urban resident or rural resident (1 indicates urban resident) and household income.

v. Moving model

Table 2.15 presents heterogeneous impact for four standardised mental health factors in terms of how people were displaced - displaced together with village or displaced alone by individual households. Individuals who moved alone were not affected in terms of their depression level. Those who were displaced with their village were hit much harder than the individual movers (52.7% of standard deviation with the full sample and OLS estimate; 47.1% of standard deviation with the reduced sample and OLS estimate; 42.4% of standard deviation with the reduced sample and fixed effects estimate). Notice the estimate with the reduced sample is 10% smaller than the estimate with the full sample. This finding is unexpected. Usually, when people face a new environment, they feel less depressed when acquaintances surround them than alone. However, these results suggest the opposite story that depression could be contagious among a group of acquaintances and that it would be difficult to build a new social network when the old social network exists. In addition, having the whole village moved together induced a compounding self-feeding feeling of defeat. Another explanation could be that moving with a village constitutes a more “substantial” move - farther away in terms of distance or social connections. Individuals moving alone were moving closer or into their relatives’ homes or villages.

The impact on the levels of anxiety, low mood and low self-esteem was not significantly different between people who moved together with their village and those who moved alone. These three factors contain mental health questions associated with the environment - for example, questions of hope for the future, and questions relating to things around the respondent - while the depression factor comes from questions more related to inner feelings. This difference in impacts on depression and the other three factors suggests the heterogeneous impacts on group movers, as opposed to individual movers, are related to their inner feelings.

Table 2.15: Moving Model

	(1) Full	(2) Reduced	(3) Reduced
Panel A: Depression			
Treat*Post	-0.034 (0.123)	-0.112 (0.131)	-0.157 (0.121)
Treat*Post*Move with village	-0.527*** (0.138)	-0.471*** (0.145)	-0.424*** (0.147)
Treat*Post*Other	-0.015 (0.167)	0.107 (0.177)	0.147 (0.176)
Panel B: Anxiety			
Treat*Post	-0.019 (0.124)	-0.127 (0.135)	-0.074 (0.115)
Treat*Post*Move with village	-0.048 (0.139)	-0.049 (0.150)	-0.044 (0.140)
Treat*Post*Other	0.000 (0.167)	0.102 (0.177)	0.053 (0.155)
Panel C: Low Mood			
Treat*Post	-0.193 (0.121)	-0.211 (0.130)	-0.179 (0.119)
Treat*Post*Move with village	0.105 (0.131)	0.095 (0.139)	0.102 (0.136)
Treat*Post*Other	0.221 (0.173)	0.295 (0.185)	0.284 (0.189)
Panel D: Low Self-esteem			
Treat*Post	-0.256** (0.127)	-0.301** (0.138)	-0.273** (0.131)
Treat*Post*Move with village	0.169 (0.133)	0.226 (0.145)	0.026 (0.149)
Treat*Post*Other	0.367** (0.171)	0.405** (0.188)	0.264 (0.183)
N	2,108	1,670	1,670
Fixed Effect	No	No	Yes
OLS	Yes	Yes	No
Number of id	1,364	926	926

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. The reference group contains the people who moved by individual households. Control variables include age, gender, whether the individual owns a house, the number of household members, whether the individual is registered as an urban resident or rural resident (1 indicates urban resident) and household income.

2.6 Robustness

As mentioned in section 4, treatment and control groups are not quite balanced in terms of individuals'/households' exogenous characteristics. To address this issue, I perform 10-to-1 matching with calliper 0.01 in order to find a balanced sample between the treatment group and the control group. The matched sample I find is a 2.5% (64 out of 2,533) drop of the unmatched sample, so slightly smaller than the unmatched sample. Although the number of observations dropped is very small, the imbalance between the control and the treatment is reduced and the comparability of the control and treatment groups is improved. [Figure 2.1](#) below provides a visual comparison of the comparability of covariates between the matched and unmatched samples. [Table 2.16](#) is the description of the diagnostic of the matched sample after it is rebalanced; there is no longer an imbalance in exogenous characteristics. Then I reperform the whole analysis, just as I did with the unmatched sample. Comparing the results between the matched and unmatched samples, the results are very similar, and therefore robust (for the main results see [Appendix 2C](#)).

Another issue worth mentioning is that the hypotheses are more likely to be falsely rejected due to multiple hypothesis testing. [Appendix 2D](#) shows the Westfall-Young adjusted p-values which are corrected for multiple hypothesis testing. [Table 2D.1](#) and [Table 2D.2](#) compare unadjusted p-values and Westfall-Young adjusted p-values with the full sample and reduced sample, respectively. Most estimates of the heterogeneity impacts disappear, except for the heterogeneous impact on depression regarding the moving model, however, the Westfall-Young adjusted p-value is higher than the unadjusted p-value, and the level of significance is at 10% instead of 1%.

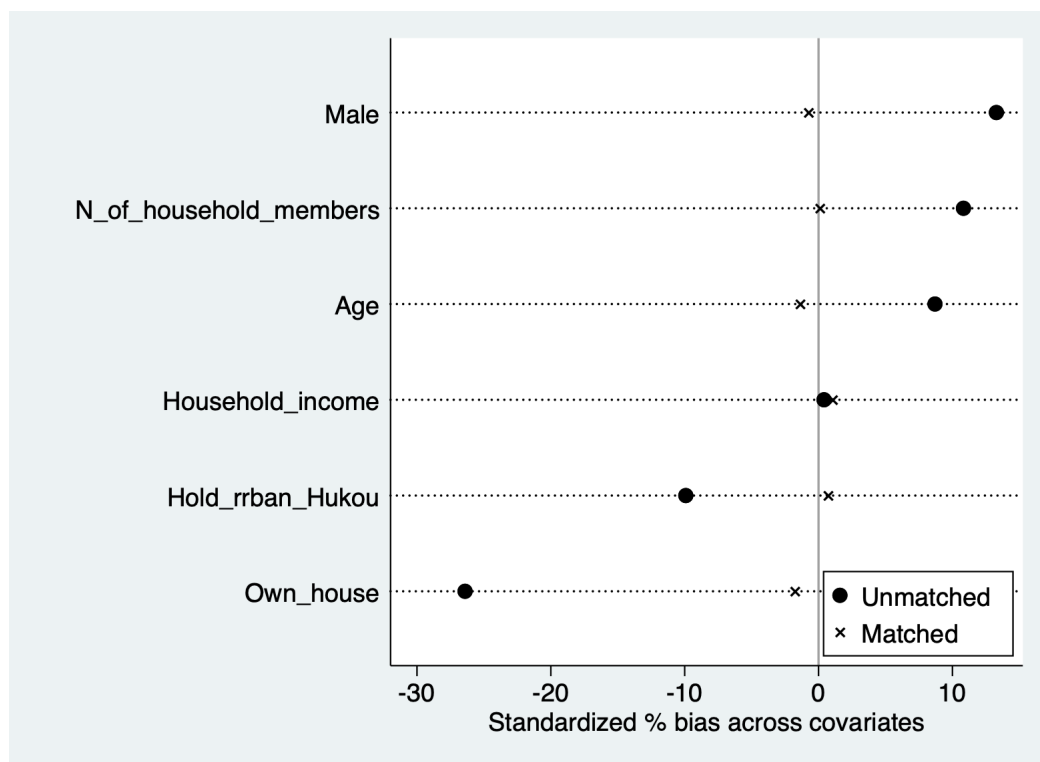


Figure 2.1: Matched versus unmatched sample

Table 2.16: Sample Balance with Matched Sample

	Unmatched Matched	Mean Treated	Control	%bias	%reduct $ bias $	t-test t	$p > t$
Age	U	45.451	44.24	8.7		1.59	0.112
	M	45.232	45.455	-1.6	81.6	-0.36	0.721
Male	U	0.468	0.402	13.3		2.41	0.016
	M	0.463	0.467	-0.9	93.3	-0.2	0.843
House owner	U	0.834	0.920	-26.4		-4.56	0
	M	0.841	0.851	-3	88.6	-0.61	0.542
Number of house members	U	3.917	3.768	10.8		1.89	0.059
	M	3.920	3.948	-2.1	81	-0.45	0.654
Urban Hukou holder	U	0.495	0.544	-9.9		-1.8	0.072
	M	0.496	0.496	0	99.8	0	0.997
Household income	U	0.120	0.119	0.4		0.07	0.947
	M	0.110	0.107	0.9	-109.8	0.35	0.724
Dummy for missing value	U	0.034	0.023	7		1.23	0.219
	M	0.030	0.032	-1.4	80.6	-0.29	0.77

Note: The matched sample is generated by 10-to-1 matching with caliber 0.01.

Several limitations are worth mentioning. First, since I do not have data before 2003, the parallel trend assumption between the treatment group and the control group, which is a typical assumption for the difference-in-differences approach, cannot be tested. If I had another data set consisting of people who also lived along the Yangtze River but were not affected by TGDP - for example, those living at 190 metres above sea level - and the data set had two waves (2003 and 2006), then I could run a placebo test using this data and test whether the change in the mental health status of people living at 190 metres above sea level was statistically different from those living right above 175 metres above sea level (current control group). If no significant difference was found, then making the assumption of a parallel trend between treatment and control would be acceptable. However, I do not have this kind of data, and thus I am unable to do a placebo test.

To reduce the problems of possible biases due to different trends, the model controls for as many exogenous variables as possible, without sacrificing the number of observations, and includes individual FE. I also reduce the differences between the treatment group and the control group with a matched sample to check the results (see above).

Second, the results only speak for the short-term impacts, as the long-term data is not available at the moment. From the qualitative study by [Wilmsen and Van Hulten \(2017\)](#), the long-term overall impact could be milder (or less negative) than the short-term effects, and the heterogeneity analysis could be different.

Third, as mentioned in the previous section, the designated migrants would be affected not only by the displacement itself but also by the anticipation of displacement. As this anticipation effect is negative in anxiety and low mood (see [Table 2.10](#)), it would attenuate the effect of displacement. However, how this anticipation would affect the heterogeneity is unclear.

2.7 Conclusion

This paper explores the impact of forced displacement on people's mental health. First, this study examines the average ITT effect and ATT effect and finds that the ATT effect is much higher than the ITT effect on depression, which suggests that actual moving is more detrimental to levels of depression than the expectation of moving.

Apart from the average impact of the displacement, this study also finds heterogeneous impacts of forced displacement on mental health for different groups in terms of gender, education, marital status, moving distance and moving model. Males got lower mood and lower self-esteem due to TGDP compared with females. Illiterates individuals became more anxious due to TGDP compared with those with general education. The impact on the divorced respondents was not as bad as the impact on the married ones in terms of their levels of depression and low mood but was worse for levels of anxiety and low self-esteem. The reversed directions of heterogeneity for standardised mental health factors may reflect the fact that displacement is a two-faceted experience for the divorced: they face the pressure of their damaged reputation just by being divorced, but they also have the opportunity of a completely new life by moving.

Out-of-town movers became much more depressed than within-town movers. People who moved together with their village became more depressed than those who moved by individual households, which is unexpected. This suggests that moving together with the village could result in a feeling of defeat as a whole, which compounds the depression effect of the group. On the other hand, it could be the case that people who moved with individual households may move closer to their acquaintances while those who moved with the village may have to move further away.

However, most of the heterogeneity impacts found above disappear after adjusting for multiple hypothesis testing. The only impact that remains is that the out-of-town movers became much more depressed than within-town movers, but it is significant

at 10% level instead of 1%.

As this study focuses on short-term effects, an extension for this study is to conduct another interview chasing the same households. Adding this wave of long-term post-displacement allows researchers to identify long-term displacement effects, which is also instructive for policymaking.

These results suggest that the experience of designated migrants varies, given their different characteristics. As displacement policies are usually the same for every migrant, they usually do not fit everyone. From the results, migrants with different levels of education, having different marital statuses, moving with different distances and models were affected differently by TGDP. Therefore, adopting a more tailored policy is the key to the “developmental resettlement” that aims to support a better livelihood after displacement.

Appendix 2A. Principal Component Factor Analysis for mental health

Table 2A.1: Principal Component Factor Analysis - Regression Score

	(1)	(2)	(3)	(4)
	Factor 1	Factor 2	Factor 3	Factor 4
	Anxiety	Depression	Low mood	Low self-esteem
Always feel bothered	0.204	-0.009	0.000	-0.112
Always feel everything you do need great effort	0.233	-0.046	-0.043	-0.068
Always feel that you are not as good as others	-0.333	0.109	0.386	0.090
Always find yourself difficult to focus on what you are doing	0.216	-0.026	-0.124	-0.015
Always feel sad	0.029	0.204	0.019	-0.069
Always feel scared	-0.181	0.457	-0.053	-0.088
Always feel lonely	-0.167	0.348	0.050	0.001
Always cry	-0.129	0.389	-0.047	-0.077
Always feel not like talking	0.062	0.047	-0.009	0.075
Always feel difficult to fall asleep	0.181	0.008	-0.060	-0.053
Do not always feel life is happy	-0.052	-0.056	0.388	0.015
Always feel hard to get rid of bad mood, even with help from families and friends	0.162	-0.035	0.006	0.024
Always feel your life is unsuccessful	0.248	-0.155	0.026	-0.007
Do not always feel happy	-0.024	-0.020	0.354	-0.028
Always get frustrated no matter what you do	0.213	-0.090	0.028	0.009
Do not always feel your future is promising	-0.018	-0.063	0.353	-0.091
Always feel that other people are unfriendly to you	-0.114	-0.097	-0.003	0.573
Always feel not like eating	0.059	0.068	-0.077	0.101
Always feel in a bad mood	0.151	-0.021	0.017	0.056
Always feel that other people do not like you	-0.105	-0.087	-0.020	0.557

Note: The names of four factors are given based on the mental health questions which have higher scores (in absolute value), as the higher the score in absolute value, the higher correlation between the factor and mental health question.

Appendix 2B

Table 2B.1: Intention to Treat Effect - Anxiety

	(1) Full	(2) Reduced	(3) Reduced	(4) Reduced
Treat	-0.090 (0.056)	-0.021 (0.068)	-0.198 (0.161)	
Post	-0.037 (0.067)	0.053 (0.074)	0.053 (0.074)	0.053 (0.069)
Treat*Post	0.023 (0.084)	-0.052 (0.092)	-0.052 (0.092)	-0.052 (0.084)
Hazard rate			5.079 (4.071)	
Age	-0.002 (0.002)	-0.003 (0.002)	-0.021 (0.015)	
Male	0.026 (0.042)	0.001 (0.046)	0.139 (0.122)	
Urban Hukou holder	-0.183*** (0.047)	-0.190*** (0.054)	-0.851 (0.535)	
House owner	0.054 (0.062)	0.012 (0.070)	0.008 (0.070)	
Number of household members	0.001 (0.014)	0.007 (0.016)	0.008 (0.016)	
Household income	0.153* (0.083)	0.122 (0.076)	-0.003 (0.126)	
Dummy for missing value (income)	-0.240* (0.124)	-0.242* (0.132)	-1.495 (1.022)	
Observations	2,533	2,018	2,018	2,018
R-squared	0.031	0.026	0.026	0.001
Fixed Effect	No	No	No	Yes
OLS	Yes	Yes	Yes	No
Number of id	1,524	1,009	1,009	1,009

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Column 1 presents results using the full sample. Columns 2-4 present results using the reduced sample where individuals who appear in the survey for one wave only are dropped.

Table 2B.2: Intention to Treat Effect - Low Mood

	(1) Full	(2) Reduced	(3) Reduced	(4) Reduced
Treat	-0.136** (0.058)	-0.102 (0.068)	-0.331** (0.166)	
Post	-0.163** (0.070)	-0.174** (0.076)	-0.174** (0.076)	-0.174** (0.068)
Treat*Post	-0.013 (0.085)	-0.042 (0.093)	-0.042 (0.093)	-0.042 (0.083)
Hazard rate			6.572 (4.229)	
Age	-0.001 (0.002)	-0.002 (0.002)	-0.025 (0.015)	
Male	-0.050 (0.042)	-0.056 (0.045)	0.122 (0.123)	
Urban Hukou holder	-0.195*** (0.048)	-0.165*** (0.053)	-1.020* (0.550)	
House owner	0.072 (0.065)	0.024 (0.071)	0.019 (0.071)	
Number of household members	0.052*** (0.014)	0.054*** (0.016)	0.054*** (0.016)	
Household income	0.178*** (0.067)	0.166*** (0.063)	0.004 (0.122)	
Dummy for missing value (income)	-0.487*** (0.135)	-0.541*** (0.159)	-2.163** (1.062)	
Observations	2,533	2,018	2,018	2,018
R-squared	0.047	0.050	0.051	0.025
Fixed Effect	No	No	No	Yes
OLS	Yes	Yes	Yes	No
Number of id	1,524	1,009	1,009	1,009

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Column 1 presents results using the full sample. Columns 2-4 present results using the reduced sample where individuals who appear in the survey for one wave only are dropped.

Table 2B.3: Intention to Treat Effect - Low Self-esteem

	(1) Full	(2) Reduced	(3) Reduced	(4) Reduced
Treat	0.009 (0.055)	0.040 (0.066)	0.112 (0.173)	
Post	0.000 (0.069)	0.022 (0.074)	0.022 (0.074)	0.022 (0.068)
Treat*Post	-0.035 (0.084)	-0.063 (0.092)	-0.063 (0.092)	-0.063 (0.086)
Hazard rate			-2.064 (4.598)	
Age	0.008*** (0.002)	0.008*** (0.002)	0.015 (0.017)	
Male	-0.179*** (0.042)	-0.190*** (0.046)	-0.247* (0.133)	
Urban Hukou holder	0.159*** (0.048)	0.179*** (0.055)	0.447 (0.600)	
House owner	0.001 (0.064)	0.009 (0.074)	0.010 (0.074)	
Number of household members	0.011 (0.014)	0.012 (0.016)	0.011 (0.016)	
Household income	0.014 (0.053)	0.003 (0.056)	0.054 (0.127)	
Dummy for missing value (income)	0.177 (0.123)	0.145 (0.148)	0.655 (1.166)	
Observations	2,533	2,018	2,018	2,018
R-squared	0.039	0.04	0.04	0.001
Fixed Effect	No	No	No	Yes
OLS	Yes	Yes	Yes	No
Number of id	1,524	1,009	1,009	1,009

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Column 1 presents results using the full sample. Columns 2-4 present results using the reduced sample where individuals who appear in the survey for one wave only are dropped.

Table 2B.4: Treatment Effect on the Treat - Anxiety

	(1) Propensity to relocate Full	(2) Anxiety Full	(3) Propensity to relocate Reduced	(4) Anxiety Reduced	(5) Propensity to relocate Reduced	(6) Anxiety Reduced	(7) Propensity to relocate Reduced	(8) Anxiety Reduced
Post	0.000 (0.002)	-0.037 (0.067)	0.000 (0.003)	0.053 (0.074)	0.000 (0.003)	0.053 (0.074)	0.000 -	0.053 (0.069)
Treat	-0.005* (0.003)	-0.090 (0.055)	-0.007* (0.004)	-0.021 (0.067)	0.019 (0.042)	-0.197 (0.161)		
Treat*Post	0.594*** (0.019)		0.596*** (0.019)		0.596*** (0.019)		0.596*** (0.019)	
Predicted Propensity to Relocate		0.038 (0.141)		-0.088 (0.154)		-0.088 (0.154)		-0.088 (0.140)
Hazard rate					-0.742 (1.206)	5.014 (4.052)		
Observations	2,533	2,533	2,018	2,018	2,018	2,018	2,018	2,018
R-squared	0.527	0.029	0.51	0.029	0.51	0.03	0.596	
Fixed Effect	No	No	No	No	No	No	Yes	Yes
OLS	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Number of id	1,524	1,524	1,009	1,009	1,009	1,009	1,009	1,009

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Columns 1, 3, 5 and 7 present estimates of the first stage. Columns 2, 4, 6 and 8 present estimates of the second stage. Control variables include age, gender, whether the individual owns a house, the number of household members, whether the individual is registered as an urban resident or rural resident (1 indicates urban resident) and household income.

Table 2B.5: Treatment Effect on the Treat - Low Mood

	(1) Propensity to relocate Full	(2) Low mood Full	(3) Propensity to relocate Reduced	(4) Low mood Reduced	(5) Propensity to relocate Reduced	(6) Low mood Reduced	(7) Propensity to relocate Reduced	(8) Low mood Reduced
Post	0.000 (0.002)	-0.163** (0.069)	0.000 (0.003)	-0.174** (0.075)	0.000 (0.003)	-0.174** (0.075)	0.000 -	-0.174** (0.068)
Treat	-0.005* (0.003)	-0.136** (0.057)	-0.007* (0.004)	-0.102 (0.067)	0.019 (0.042)	-0.330** (0.166)		
Treat*Post	0.594*** (0.019)		0.596*** (0.019)		0.596*** (0.019)		0.596*** (0.019)	
Predicted Propensity to Relocate		-0.022 (0.142)		-0.071 (0.154)		-0.071 (0.154)		-0.071 (0.140)
Hazard rate					-0.742 (1.206)	6.519 (4.199)		
Observations	2,533	2,533	2,018	2,018	2,018	2,018	2,018	2,018
R-squared	0.527	0.047	0.51	0.052	0.51	0.053	0.596	
Fixed Effect	No	No	No	No	No	No	Yes	Yes
OLS	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Number of id	1,524	1,524	1,009	1,009	1,009	1,009	1,009	1,009

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Columns 1, 3, 5 and 7 present estimates of the first stage. Columns 2, 4, 6 and 8 present estimates of the second stage. Control variables include age, gender, whether the individual owns a house, the number of household members, whether the individual is registered as an urban resident or rural resident (1 indicates urban resident) and household income.

Table 2B.6: Treatment Effect on the Treated - Low Self-esteem

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Propensity to Relocate Full	Low self- esteem Full	Propensity to Relocate Reduced	Low self- esteem Reduced	Propensity to Relocate Reduced	Low self- esteem Reduced	Propensity to Relocate Reduced	Low self- esteem Reduced
Post	0.000 (0.002)	0.000 (0.069)	0.000 (0.003)	0.022 (0.074)	0.000 (0.003)	0.022 (0.074)	0.000 -	0.022 (0.068)
Treat	-0.005* (0.003)	0.008 (0.054)	-0.007* (0.004)	0.039 (0.065)	0.019 (0.042)	0.114 (0.173)		
Treat*Post	0.594*** (0.019)		0.596*** (0.019)		0.596*** (0.019)		0.596*** (0.019)	
Predicted Propensity to Relocate		-0.059 (0.141)		-0.106 (0.154)		-0.106 (0.154)		-0.106 (0.144)
Hazard rate					-0.742 (1.206)	-2.143 (4.600)		
Observations	2,533	2,533	2,018	2,018	2,018	2,018	2,018	2,018
R-squared	0.527	0.04	0.51	0.04	0.51	0.04	0.596	
Fixed Effect	No	No	No	No	No	No	Yes	Yes
OLS	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Number of id	1,524	1,524	1,009	1,009	1,009	1,009	1,009	1,009

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Columns 1, 3, 5 and 7 present estimates of the first stage. Columns 2, 4, 6 and 8 present estimates of the second stage. Control variables include age, gender, whether the individual owns a house, the number of household members, whether the individual is registered as an urban resident or rural resident (1 indicates urban resident) and household income.

Appendix 2C. Results with Matched Sample

Table 2C.1: Intention to Treat Effect - Depression (matched sample)

	(1) Full	(2) Reduced	(3) Reduced	(4) Reduced
Treat	-0.113** (0.054)	-0.057 (0.066)	-0.498** (0.218)	
Post	-0.066 (0.067)	-0.085 (0.074)	-0.085 (0.074)	-0.085 (0.069)
Treat*Post	-0.149* (0.083)	-0.202** (0.091)	-0.202** (0.091)	-0.202** (0.086)
Hazard rate			9.259** (4.369)	
Age	-0.004** (0.002)	-0.003* (0.002)	-0.030** (0.013)	
Male	0.383*** (0.040)	0.365*** (0.045)	0.617*** (0.127)	
Urban Hukou holder	0.051 (0.049)	0.126** (0.054)	-1.123* (0.591)	
House owner	-0.001 (0.065)	-0.018 (0.073)	-0.019 (0.073)	
Number of household members	0.0246 (0.016)	0.034** (0.017)	0.035** (0.017)	
Household income	0.249** (0.104)	0.137 (0.161)	-0.137 (0.206)	
Dummy for missing value (income)	-0.223 (0.143)	-0.269 (0.190)	-2.263** (0.980)	
Observations	2,469	1,976	1,976	1,976
R-squared	0.071	0.076	0.078	0.034
Fixed Effect	No	No	No	Yes
OLS	Yes	Yes	Yes	No
Number of id	1,481	988	988	988

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Column 1 presents results using the full sample. Columns 2-4 present results using the reduced sample where individuals who appear in the survey for one wave only are dropped.

Table 2C.2: Intention to Treat Effect - Anxiety (matched sample)

	(1) Full	(2) Reduced	(3) Reduced	(4) Reduced
Treat	-0.089 (0.057)	-0.018 (0.068)	-0.169 (0.219)	
Post	-0.027 (0.069)	0.052 (0.076)	0.052 (0.076)	0.052 (0.070)
Treat*Post	0.019 (0.086)	-0.053 (0.094)	-0.053 (0.094)	-0.053 (0.086)
Hazard rate			3.158 (4.295)	
Age	-0.002 (0.002)	-0.002 (0.002)	-0.012 (0.013)	
Male	0.022 (0.042)	0.003 (0.047)	0.089 (0.127)	
Urban Hukou holder	-0.201*** (0.048)	-0.220*** (0.054)	-0.646 (0.584)	
House owner	0.047 (0.064)	-0.014 (0.072)	-0.014 (0.072)	
Number of household members	-0.010 (0.014)	-0.004 (0.016)	-0.004 (0.016)	
Household income	0.571*** (0.147)	0.579*** (0.215)	0.485** (0.245)	
Dummy for missing value (income)	-0.192 (0.133)	-0.147 (0.139)	-0.827 (0.940)	
Observations	2,469	1,976	1,976	1,976
R-squared	0.034	0.029	0.029	0.001
Fixed Effect	No	No	No	Yes
OLS	Yes	Yes	Yes	No
Number of id	1,481	988	988	988

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Column 1 presents results using the full sample. Columns 2-4 present results using the reduced sample where individuals who appear in the survey for one wave only are dropped.

Table 2C.3: Intention to Treat Effect - Low Mood (matched sample)

	(1) Full	(2) Reduced	(3) Reduced	(4) Reduced
Treat	-0.119** (0.058)	-0.091 (0.068)	-0.293 (0.215)	
Post	-0.160** (0.071)	-0.168** (0.076)	-0.168** (0.076)	-0.168** (0.069)
Treat*Post	-0.025 (0.086)	-0.056 (0.093)	-0.056 (0.093)	-0.056 (0.085)
Hazard rate			4.232 (4.221)	
Age	-0.001 (0.002)	-0.002 (0.002)	-0.014 (0.013)	
Male	-0.046 (0.042)	-0.056 (0.045)	0.059 (0.123)	
Urban Hukou holder	-0.221*** (0.049)	-0.182*** (0.053)	-0.753 (0.572)	
House owner	0.041 (0.066)	0.004 (0.073)	0.003 (0.073)	
Number of household members	0.043*** (0.015)	0.041** (0.016)	0.042** (0.016)	
Household income	0.601*** (0.202)	0.809*** (0.183)	0.684*** (0.221)	
Dummy for missing value (income)	-0.385*** (0.136)	-0.439*** (0.158)	-1.351 (0.909)	
Observations	2,469	1,976	1,976	1,976
R-squared	0.046	0.052	0.052	0.026
Fixed Effect	No	No	No	Yes
OLS	Yes	Yes	Yes	No
Number of id	1,481	988	988	988

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Column 1 presents results using the full sample. Columns 2-4 present results using the reduced sample where individuals who appear in the survey for one wave only are dropped.

Table 2C.4: Intention to Treat Effect - Low Self-esteem (matched sample)

	(1) Full	(2) Reduced	(3) Reduced	(4) Reduced
Treat	0.013 (0.055)	0.051 (0.067)	-0.027 (0.247)	
Post	0.015 (0.070)	0.047 (0.075)	0.047 (0.075)	0.047 (0.069)
Treat*Post	-0.046 (0.085)	-0.079 (0.093)	-0.079 (0.093)	-0.079 (0.086)
Hazard rate			1.620 (5.004)	
Age	0.008*** (0.002)	0.008*** (0.002)	0.003 (0.015)	
Male	-0.168*** (0.042)	-0.181*** (0.047)	-0.137 (0.144)	
Urban Hukou holder	0.164*** (0.049)	0.184*** (0.055)	-0.035 (0.677)	
House owner	0.001 (0.065)	0.009 (0.075)	0.009 (0.075)	
Number of household members	0.011 (0.014)	0.015 (0.016)	0.016 (0.016)	
Household income	-0.005 (0.136)	-0.144 (0.195)	-0.192 (0.245)	
Dummy for missing value (income)	0.158 (0.128)	0.143 (0.157)	-0.206 (1.111)	
Observations	2,469	1,976	1,976	1,976
R-squared	0.039	0.04	0.04	0.001
Fixed Effect	No	No	No	Yes
OLS	Yes	Yes	Yes	No
Number of id	1,481	988	988	988

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Column 1 presents results using the full sample. Columns 2-4 present results using the reduced sample where individuals who appear in the survey for one wave only are dropped.

Table 2C.5: Treatment Effect on the Treated - Depression (matched sample)

	(1) Propensity to Relocate Full	(2) Depression Full	(3) Propensity to Relocate Reduced	(4) Depression Reduced	(5) Propensity to Relocate Reduced	(6) Depression Reduced	(7) Propensity to Relocate Reduced	(8) Depression Reduced
Post	0.001 (0.003)	-0.066 (0.067)	0.000 (0.003)	-0.085 (0.074)	0.000 (0.003)	-0.085 (0.074)	0.000 -	-0.085 (0.069)
Treat	-0.004 (0.003)	-0.114** (0.053)	-0.006 (0.004)	-0.059 (0.065)	0.041 (0.057)	-0.484** (0.216)		
Treat*Post	0.592*** (0.019)		0.592*** (0.019)		0.592*** (0.019)		0.592*** (0.019)	
Propensity to Relocate		-0.252* (0.140)		-0.341** (0.153)		-0.341** (0.153)		-0.341** (0.145)
Hazard rate					-0.989 (1.194)	8.922** (4.315)		
Observations	2,469	2,469	1,976	1,976	1,976	1,976	1,976	1,976
R-squared	0.526	0.075	0.505	0.079	0.505	0.081	0.592	
Fixed Effect	No	No	No	No	No	No	Yes	Yes
OLS	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Number of id	1,481	1,481	988	988	988	988	988	988

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Columns 1, 3, 5 and 7 present estimates of the first stage. Columns 2, 4, 6 and 8 present estimates of the second stage. Control variables include age, gender, whether the individual owns a house, the number of household members, whether the individual is registered as an urban resident or rural resident (1 indicates urban resident) and household income.

Table 2C.6: Treatment Effect on the Treated - Anxiety (matched sample)

	(1) Propensity to Relocate Full	(2) Anxiety Full	(3) Propensity to Relocate Reduced	(4) Anxiety Reduced	(5) Propensity to Relocate Reduced	(6) Anxiety Reduced	(7) Propensity to Relocate Reduced	(8) Anxiety Reduced
Post	0.001 (0.003)	-0.027 (0.069)	0.000 (0.003)	0.052 (0.076)	0.000 (0.003)	0.052 (0.076)	0.000 -	0.052 (0.070)
Treat	-0.004 (0.003)	-0.089 (0.056)	-0.006 (0.004)	-0.019 (0.067)	0.041 (0.057)	-0.165 (0.220)		
Treat*Post	0.592*** (0.019)		0.592*** (0.019)		0.592*** (0.019)		0.592*** (0.019)	
Propensity to Relocate		0.033 (0.144)		-0.089 (0.157)		-0.089 (0.157)		-0.089 (0.144)
Hazard rate					-0.989 (1.194)	3.070 (4.276)		
Observations	2,469	2,469	1,976	1,976	1,976	1,976	1,976	1,976
R-squared	0.526	0.033	0.505	0.032	0.505	0.033	0.592	
Fixed Effect	No	No	No	No	No	No	Yes	Yes
OLS	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Number of id	1,481	1,481	988	988	988	988	988	988

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Columns 1, 3, 5 and 7 present estimates of the first stage. Columns 2, 4, 6 and 8 present estimates of the second stage. Control variables include age, gender, whether the individual owns a house, the number of household members, whether the individual is registered as an urban resident or rural resident (1 indicates urban resident) and household income.

Table 2C.7: Treatment Effect on the Treated - Low Mood (matched sample)

	(1) Propensity to Relocate Full	(2) Low mood Full	(3) Propensity to Relocate Reduced	(4) Low mood Reduced	(5) Propensity to Relocate Reduced	(6) Low mood Reduced	(7) Propensity to Relocate Reduced	(8) Low mood Reduced
Post	0.001 (0.003)	-0.160** (0.071)	0.000 (0.003)	-0.168** (0.076)	0.000 (0.003)	-0.168** (0.076)	0.000 -	-0.168** (0.069)
Treat	-0.004 (0.003)	-0.119** (0.057)	-0.006 (0.004)	-0.091 (0.067)	0.041 (0.057)	-0.289 (0.215)		
Treat*Post	0.592*** (0.019)		0.592*** (0.019)		0.592*** (0.019)		0.592*** (0.019)	
Propensity to Relocate		-0.042 (0.145)		-0.094 (0.156)		-0.094 (0.156)		-0.094 (0.144)
Hazard rate					-0.989 (1.194)	4.14 (4.188)		
Observations	2,469	2,469	1,976	1,976	1,976	1,976	1,976	1,976
R-squared	0.526	0.047	0.505	0.054	0.505	0.055	0.592	
Fixed Effect	No	No	No	No	No	No	Yes	Yes
OLS	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Number of id	1,481	1,481	988	988	988	988	988	988

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Columns 1, 3, 5 and 7 present estimates of the first stage. Columns 2, 4, 6 and 8 present estimates of the second stage. Control variables include age, gender, whether the individual owns a house, the number of household members, whether the individual is registered as an urban resident or rural resident (1 indicates urban resident) and household income.

Table 2C.8: Treatment Effect on the Treated - Low Self-esteem (matched sample)

	(1) Propensity to Relocate Full	(2) Low self-esteem Full	(3) Propensity to Relocate Reduced	(4) Low self-esteem Reduced	(5) Propensity to Relocate Reduced	(6) Low self-esteem Reduced	(7) Propensity to Relocate Reduced	(8) Low self-esteem Reduced
Post	0.001 (0.003)	0.015 (0.070)	0.000 (0.003)	0.047 (0.075)	0.000 (0.003)	0.047 (0.075)	0.000 -	0.047 (0.069)
Treat	-0.004 (0.003)	0.0131 (0.055)	-0.006 (0.004)	0.050 (0.066)	0.041 (0.057)	-0.021 (0.248)		
Treat*Post	0.592*** (0.019)		0.592*** (0.019)		0.592*** (0.019)		0.592*** (0.019)	
Propensity to Relocate		-0.078 (0.144)		-0.133 (0.156)		-0.133 (0.156)		-0.133 (0.145)
Hazard rate					-0.989 (1.194)	1.488 (5.007)		
Observations	2,469	2,469	1,976	1,976	1,976	1,976	1,976	1,976
R-squared	0.526	0.039	0.505	0.04	0.505	0.04	0.592	
Fixed Effect	No	No	No	No	No	No	Yes	Yes
OLS	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Number of id	1,481	1,481	988	988	988	988	988	988

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Columns 1, 3, 5 and 7 present estimates of the first stage. Columns 2, 4, 6 and 8 present estimates of the second stage. Control variables include age, gender, whether the individual owns a house, the number of household members, whether the individual is registered as an urban resident or rural resident (1 indicates urban resident) and household income.

Appendix 2D. P-values corrected for multiple hypothesis testing

Table 2D.1: Unadjusted p-value versus Westfall-Young adjusted p-value - full sample

	Depression		Anxiety		Low mood		Low self-esteem	
	Unadjusted p-value	Westfall- Young adjusted p-value	Unadjusted p-value	Westfall- Young adjusted p-value	Unadjusted p-value	Westfall- Young adjusted p-value	Unadjusted p-value	Westfall- Young adjusted p-value
Gender								
Treat*Post*Male	0.616	0.630	0.754	0.990	0.039	0.350	0.222	0.840
Education								
Treat*Post*Illiteracy	0.031	0.400	0.466	0.980	0.214	0.820	0.435	0.960
Treat*Post*Technical Edu	0.270	0.560	0.045	0.540	0.307	0.940	0.134	0.740
Treat*Post*College	0.098	0.560	0.384	0.980	0.709	1.000	0.883	0.990
Marital Status								
Treat*Post*Single	0.190	0.560	0.385	0.980	0.752	1.000	0.585	0.970
Treat*Post*Divorced	0.092	0.560	0.510	0.980	0.012	0.220	0.441	0.960
Treat*Post*Widowed	0.096	0.560	0.750	0.990	0.319	0.940	0.883	0.990
Moving Distance								
Treat*Post*Move within town	0.002	0.140	0.001	0.100	0.464	0.940	0.118	0.740
Moving Model								
Treat*Post*Move with village	0.000	0.070	0.704	0.990	0.861	1.000	0.734	0.990

Note: All the regressions are estimated with OLS.

Table 2D.2: Unadjusted p-value versus Westfall-Young adjusted p-value - reduced sample

	Depression		Anxiety		Low mood		Low self-esteem	
	Unadjusted p-value	Westfall- Young adjusted p-value	Unadjusted p-value	Westfall- Young adjusted p-value	Unadjusted p-value	Westfall- Young adjusted p-value	Unadjusted p-value	Westfall- Young adjusted p-value
Gender								
Treat*Post*Male	0.952	0.900	0.882	0.990	0.083	0.520	0.258	0.920
Education								
Treat*Post*Illiteracy	0.153	0.650	0.699	0.990	0.072	0.770	0.413	0.950
Treat*Post*Technical Edu	0.246	0.750	0.100	0.750	0.112	0.490	0.159	0.880
Treat*Post*College	0.069	0.570	0.224	0.900	0.808	0.970	0.822	0.970
Marital Status								
Treat*Post*Single	0.146	0.650	0.474	0.940	0.600	0.970	0.687	0.970
Treat*Post*Divorced	0.331	0.750	0.947	0.990	0.002	0.110	0.399	0.950
Treat*Post*Widowed	0.307	0.750	0.289	0.900	0.346	0.920	0.915	0.970
Moving Distance								
Treat*Post*Move within town	0.001	0.140	0.002	0.130	0.449	0.920	0.173	0.880
Moving Model								
Treat*Post*Move with village	0.000	0.070	0.504	0.940	0.885	0.970	0.529	0.970

Note: All the regressions are estimated with OLS.

References

- Araya, M., Chotai, J., Komproe, I.H. and de Jong, J.T.V.M. (2007). Gender differences in traumatic life events, coping strategies, perceived social support and sociodemographic among postconflict displaced persons in Ethiopia. *Social Psychiatry Psychiatr Epidemiol*, 42, 307-315.
- Booth, A. and Amato, P. (1991). Divorce and psychological. *Journal of Health and Social Behavior*, 32(4), 396-407.
- Cao, Y., Hwang, S. S., Xi, J., (2012). Project-induced Displacement, Secondary Stressors, and Health. *Social Science & Medicine*, 74(7), 1130-1138.
- de Chaisemartin, C. (2010). A Note on the Assumptions Underlying Instrumented Difference in Differences.
- de Chaisemartin, C. and D'Haultfoeuille, X. (2018). Fuzzy Differencesin-Differences. *Review of Economic Studies*, 85(2), 999–1028.
- Downes, D.J. et al., (2021) Identification of LZTFL1 as a candidate effector gene at a COVID-19 risk locus. *Nature Genetics*, 53, 1606-1615.
- Duan, Y. and Wilmsen, B. (2012). Addressing the resettlement challenges at the Three Gorges Project. *International Journal of Environmental Studies*, 69(3), 461-474. doi:10.1080/00207233.2012.676374
- Glenn, N. D. and Weaver, C. N. (1981). Education's Effects on Psychological Well-Being. *The Public Opinion Quarterly*, 45(1), 22-39.
- Eisenman, D.P., Gelberg, L., Liu, H., Shapiro, M. F. (2003). Mental health and health-related quality of life among adult Latino primary care patients living in the United States with previous exposure to political violence. *Journal of the American Medical Association*, 290(5), 627-634.

- Fazel, M., Wheeler, J., Danesh, J. (2005). Prevalence of serious mental disorder in 7000 refugees resettled in western countries: a systematic review. *Lancet*, 365(9467), 1309-1314.
- Feng, X. (2004). “Touch the ground and take the root?” – Social adaptation of Three Gorges Dam rural relocatees. *Sociological*, 5, 19-27. (in Chinese)
- Fussell E., Lowe S. R. (2014). The impact of housing displacement on the mental health of low-income parents after Hurricane Katrina. *Social Science & Medicine*, 113(2014), 137-144.
- Glenn, N. D. (1975). The Contribution of Marriage to the Psychological Well-Being of Males and Females. *Journal of Marriage and Family*, 37(3), 594-600.
- Holden, C. (2000). Global Survey Examines Impact of Depression *Science*, 288(5463), 39-40. doi: 10.1126/science.288.5463.39
- Hudson, S., Hull, P. and Liebersohn, J. (2017). Interpreting Instrumented Difference-in-Differences. Notes. <http://www.mit.edu/~hull/DDIV.pdf>
- Hwang, SS., Xi, J., Cao, Y., Feng, X., Qiao, X. (2007). Anticipation of migration and psychological stress and the Three Gorges Dam project, China. *Social science & medicine*, 65(5), 1012-1024.
- Hwang, SS., Cao, Y., Xi, J. (2010). Project-induced migration and depression: A panel analysis. *Social science & medicine*, 70(11), 1765-1772.
- Hwang, SS., Cao, Y. and Xi, J. (2011). The Short-Term Impact of Involuntary Migration in China’s Three Gorges: A Prospective Study. *Social Indicators Study*, 101(1), 73-92.
- Jackson, S., Sleight, A. (2000). Resettlement for China’s Three Gorges Dam: socio-economic impact and institutional tensions. *Communist and Post-Communist Studies*, 33(2), 223-241.

- Julian, L. J., Gregorich, S. E., Tonner, C., Yazdany, J., Trupin, L., Criswell, L. A., Yelin, E. and Katz, P.P (2011). Using the center for epidemiologic studies depression scale to screen for depression in systemic lupus erythematosus. *Arthritis Care & Research*, 63(6), 884-890.
- Kasteler, J. M., Gray, R. M., Carruth, M. L. (1968). Involuntary relocation of the elderly. *The Gerontologist*, 8(4), 276-279.
- Kessler, R. C. (1982). A Disaggregation of the Relationship between Socioeconomic Status and Psychological Distress. *American Sociological Review*, 47(6), 752-764.
- Lai, G. (1995). Work and family roles and psychological well-being in urban China. *Journal of Health and Social Behavior*, 36(1), 11-37.
- Lee I-H, Lee J-W, Kong, S.W. (2020). A survey of genetic variants in SARS-CoV-2 interesting domains of ACE2, TMPRSS2 and TLR3/7/8 across populations. *Infection, Genetics and Evolution*, 85,104507.
- Li, H. (1994). A preliminary study on the social-ecological types of migrants in the Three Gorges Reservoir area. *Sociological*, 3, 100-103. (in Chinese)
- Li, H., Waley, P. and Rees, P. (2001). Reservoir Resettlement in China: Past Experience and the Three Gorges Dam. *The Geographical Journal*, 167(3), 195-212.
- Li, Y. (2009). December. Part of Three Gorges Dam relocatees return to Chongqing. *China Newsweek*. (in Chinese)
- Liao, C. and Heaton, T. B. (1992). Divorce Trends and Differentials in China. *Journal of Comparative Family Studies*, 23(3), 413-429.
- Liu, T., Song, X., Chen, G., Buka, S. L., Zhang, L., Pang, L. and Zheng, X. (2013). Illiteracy and schizophrenia in China: a population-based survey. *Social Psychiatry and Psychiatric Epidemiology*, 48(3), 455-464.

Mills, J.P. 1926. Table of the Ratio: Area to Bounding Ordinate, for Any Portion of Normal Curve. *Biometrika*, 18(3/4), 395–400.

Morgan, W. Genetics is not why more BAME people die of coronavirus: structural racism is. *The Guardian*, 4 June 2020.

Mujeeb, A. (2015). Mental health of internally displaced persons in Jalozi camp, Pakistan. *International Journal of Social Psychiatry*, 61(7), 653–659.

Padovani, F. (2016). Analysis of the Social Implications of the Three Gorges Dam Project. In F. Padovani (Eds.), *Development-Induced Displacement in India and China* (pp. 169–190). Lexington Books.

Peng, Z. (2011), August. The Duckweed that lost root: Three Gorges Dam relocatees return to their hometown due to inadaptation to new life. *Southern Metropolis Weekly*.

Randell, H. (2016). The short-term impacts of development-induced displacement on wealth and subjective well-being in the Brazilian Amazon. *World Development*, 87, 385–400.

Porter, M., Haslam, N. (2005). Predisplacement and postdisplacement factors associated with mental health of refugees and internally displaced persons: a meta-analysis. *Journal of the American Medical Association*, 294(5), 602–612.

Robins, J.M., Hernan, M.A. and Brumback, B. (2000). Marginal Structural Models and Causal Inference in Epidemiology. *Epidemiology*, 11(5), 550–560.

Salah, T.T.M., Abdelrahman, A., Lien, L., Eide, A.H., Martinez, P. and Hauff, E. (2012). The mental health of internally displaced persons: An epidemiological study of adults in two settlements in Xentral Sudan. *International Journal of Social Psychiatry*, 59(8), 782–788.

Sastry, N. and VanLandingham, M. (2009). One Year Later: Mental Illness Prevalence and Disparities Among New Orleans Residents Displaced by Hurricane Katrina. *American Journal of Public Health*, 99(Suppl 3), S725-S731.

Shevlin, M., Hyland, P., Vallieres, F., Bisson, J., Makhashvili, N., Javakhishvili, J., Shpiker, M. and Roberts, B. (2018). A comparison of DSM-5 and ICD-11 PTSD prevalence, comorbidity and disability: an analysis of the Ukrainian Internally Displaced Person's Mental Health Survey. *Acta Psychiatrica Scandinavica*, 137, 138-147.

Stone, R. (2008). Three Gorges Dam: Into the Unknown. *Science*, New Series, 321(5889), 628-632.

Taha, P.H., Sijbrandij, M. (2021). Gender Differences in Traumatic Experiences, PTSD, and Relevant Symptoms among the Iraqi Internationally Displaced Persons. *International Journal of Environmental Research and Public Health*, 18, 9779.

Tan, Y., Hugo, G. and Potter, L. (2005). Rural Women, Displacement and the Three Gorges Project. *Development and Change*, 36(4), 711-734.

Wade T. J. and Pevalin D. J. (2004). Marital Transitions and Mental Health. *Journal of Health and Social Behavior*, 45(2), 155-170.

McDonald-Wilmsen, B. (2009). Development-Induced Displacement and Resettlement: Negotiating Fieldwork Complexities at the Three Gorges Dam, China. *The Asia Pacific Journal of Anthropology*, 10(4), 283-300. doi:10.1080/14442210903271320

Wilmsen B. (2016). After the deluge: a longitudinal study of resettlement at the Three Gorges Dam, China. *World Development*, 84, 41-54.

Wilmsen, B. and Van Hulten, A. (2017). Following resettled people over time: the value of longitudinal data collection for understanding the livelihood impacts of the Three Gorges Dam, China. *Impact Assessment and Project Appraisal*, 35, 94-105.

Wilmsen, B., Webber, M. and Duan, Y.F. (2011). Development for Whom? Ru-

ral to Urban Resettlement at the Three Gorges Dam, China. *Asian Studies Review*, 35, 21-42.

Wilmsen, B. (2018). Damming China's rivers to expand its cities: the urban livelihoods of rural people displaced by the Three Gorges Dam. *Urban Geography*, 39(3), 345-366. doi:10.1080/02723638.2017.1328578

Xi, J. and Hwang, SS. (2011). Unmet expectations and symptoms of depression among the Three Gorges Project resettlers. *Social Science Research*, 40(1), 245-256.

Xi, J. and Hwang, SS. (2011). Relocation Stress, Coping, and Sense of Control Among Resettlers Resulting from China's Three Gorges Dam Project. *Social Indicators Research*, 104(3), 507-522.

Xi, J., Hwang, SS., Feng, X., QIAO, X. and Cao, Y. (2007). Perceived Risks and Benefits of the Three Gorges Project. *Sociological Perspectives*, 50(2), 323-337.

Xi, J., Hwang, S.S. and Drentea, P. (2012). Experiencing a Forced Relocation at Different Life States: The Effects of China's Three Gorges Project-induced Relocation on Depression. *Society and Mental Health*, 3(1), 59-76.

Xi, J. (2016). Types of integration and depressive symptoms: A latent class analysis on the resettled population for the Three Gorges dam project, China. *Social Science & Medicine*, 157, 78-86.

Zeng, Y., and Wu, D. (2000). Regional Analysis of Divorce in China since 1980. *Demography*, 37(2), 215-219.

Chapter 3

Impact of shielding policies during COVID-19 pandemic on BAME versus non-BAME high-risk population

3.1 Introduction

One of the most popular discussions about the consequences of COVID-19 has been the differentiated impact on people from a Black, Asian, and Minority Ethnic (BAME) background compared with their non-BAME counterparts ([Tai et al., 2020](#); [Platt and Warwick, 2020](#); [Butcher and Massey, 2020](#); [Alsan et al., 2021](#)). The main factor of concern regarding COVID-19 is its high mortality rate, which is much higher than other comparable respiratory diseases like the flu. The UK government has resorted to national lockdown and shielding to reduce infection rates, especially for those who have critical conditions, so that the National Health Service (NHS) will not be overwhelmed and mortality will be reduced. Reports from the Office for National Statistics in the UK¹ show that ethnic minorities account for a disproportionate

¹<https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/articles/updatingethniccontrastsindeathsinvolvingthecoronaviruscovid19englandandwales/>

number of COVID-19-related deaths. From previous studies and anecdotal evidence, BAME were more affected by COVID-19 pandemic compared with their non-BAME counterparts. This could be due to two reasons: first, BAME individuals are more likely to work in the sectors where face-to-face interactions are necessary (Platt and Warwick, 2020); second, the fact that a higher proportion of BAME population dying of COVID-19 compared with the white population might negatively affect the mental health of the high-risk BAME population. Another explanation from the literature is genetic difference. Research found that 60% of people with South Asian ancestry carry a gene called LZTFL1 which was found likely to double the risk of ‘respiratory failure’ (Downes et al., 2021), although earlier articles (Morgan, 2020; Lee et al., 2020) claimed that genetics is not to blame for the high mortality rate of BAME as a whole. However, due to the inaccessibility of the relevant dataset to explore this angle, this chapter does not address this explanation. Due to the fact that pre-existing health conditions are a key factor leading to mortality due to COVID-19, to have a well-tailored policy for the BAME, it might be more impactful to focus on the clinically extremely vulnerable patients from a BAME background rather than looking at the BAME population as a whole².

This chapter studies whether the change in shielding policies would affect the high-risk differently between BAME and non-BAME populations over time in the COVID-19 pandemic. I exploit the first change in shielding policy, the lift of shielding requirement, as a threshold to explore whether the lift of shielding requirement would not be beneficial for the high-risk BAME in terms of their COVID-19 infection rates, and mental well-being-related outcomes. As the first wave surged in March 2020, together with the lockdown, the high-risk individuals were required to stay at home

24january2020to31march2021

²Just to make terminology easier to follow, from now on, I refer to individuals who are clinically extremely vulnerable (by NHS definition) as “high-risk”, and those who are not clinically extremely vulnerable as “low-risk”. I refer to individuals from a BAME background as BAME individuals, and those who are not from a BAME background as non-BAME individuals. For example, a person who is clinically extremely vulnerable and from a BAME background is simply “high-risk BAME individuals”.

(“shielding”). When the cases in the first wave went down, the shielding requirement started to be relaxed and was eventually lifted in August 2020. It would be helpful to understand how this may have affected the high-risk, and in particular, whether the impact, if any, differed between BAME and non-BAME individuals as BAME were more likely to work in the sectors where face-to-face interactions are necessary and the mortality rate among BAME was higher than the white population. In addition, this COVID-19 pandemic will not be the last pandemic, and for any pandemic, we may face in the future, there will always be a period when no vaccine is available and shielding is the only solution for keeping high-risk people safe. Understanding the shielding policy provides us with insights that can guide us through similar situations in the future.

This chapter hypothesizes that the BAME population was more adversely affected by lifting the shielding policy during the COVID-19 pandemic. The BAME population is more likely to be employed in the service industry that requires face-to-face contact or on-site work, the high-risk BAME are, therefore, more likely to be infected and feel more at risk of catching COVID-19. In addition, the fact that the BAME population has a much higher mortality rate than their non-BAME counterpart due to COVID-19 makes the high-risk BAME individuals more likely to experience mental distress and thus more likely to be diagnosed with mental ill-health. All these impacts are expected to be higher when a new wave of the pandemic happens. These outcome measures are essential to capture people’s physical and mental experiences during the COVID-19 pandemic.

To the best of my knowledge, this study is the first one to quantify the dynamic impact of lifting the shielding requirement on the possible disparity of the vulnerable people between BAME and non-BAME. This chapter employs eight waves of the Understanding Society COVID-19 survey in the UK, which allows us to follow up on the effects over time.

Applying the event study approach, no evidence is found suggesting that the high-

risk from a BAME background were more likely to have COVID-19 symptoms once the shielding requirement was relaxed. The impact on mental ill-health diagnosis became statistically significantly higher after the shielding was relaxed in July 2020. They felt more at risk of catching COVID-19 from November 2020 to January 2021 when the daily infections became higher and the shielding was not required, while in March 2021, they became less depressed and anxious. Besides well-being, I also explore people's behaviour during the pandemic including their use of transportation and the frequency of their in-person contacts. The results show that although the high-risk BAME individuals were more cautious when the COVID-19 pandemic became worse (less likely to meet friends and families outside their household, and reduced travelling by train), they were less likely to have alternative transport options.

This study makes two main contributions. First, it documents the impact of the COVID-19 pandemic and the policies to combat it on mental well-being. Mental well-being was brought to attention in the early months of the pandemic. [Daly et al. \(2020\)](#) found people's mental health deteriorated in April of 2020 compared with years before the pandemic. [Pierce et al. \(2020\)](#) found similar results and also pointed out that the deterioration in the mental health of young people, women and people with young children was the largest. However, these studies did not pay attention to the high-risk group. Using survey data from April of 2020, [Iob et al. \(2020\)](#) looked at the mental health of high-risk individuals and found that they are more likely to suffer from mental health issues. However, no significant distinction was found between BAME and non-BAME populations, which makes intuitive sense as in April of 2020, all high-risk individuals were shielding, regardless of race.

Now shift our attention to the policy that directly targets the high-risk group - shielding. While some studies focus on how the shielding requirement harms the mental well-being of shielding patients with specific health conditions ([Antoun et al., \(2021\)](#) [Taylor et al., \(2021\)](#)), [Kemp et al. \(2020\)](#) mentioned a possible threat to psychological well-being for the high-risk group as a whole when shielding was

lifted because the high-risk people would have to face the risk of catching COVID-19. However, the study did not provide any evidence for this concern. This chapter fills the gap and takes it one step further to quantify the heterogeneous impacts of lifting the shielding together with the epidemiological situation during COVID-19 on high-risk people.

Secondly, this study builds on existing research measuring the impacts of the pandemic on people from a BAME background. We have seen a higher mortality rate in the BAME population compared with the non-BAME population from media releases and survey reports. Researchers have been trying to disentangle the factors at work in making ethnic minorities suffer significantly more. These factors include but are not limited to, living in crowded areas, being employed in sectors requiring face-to-face contact and being affected by the high mortality rate within the BAME population. The complexity of the association between these factors makes it harder to identify, especially when these factors come from a more fundamental social issue. [Nazroo et al. \(2020\)](#) indicated that all these factors are rooted in a more fundamental factor: racism. However, racism is difficult to measure and this chapter is concerned with providing information to policymakers on how to mitigate the impact of shielding policies. [Hu \(2020\)](#) found that COVID-19 has enlarged the socio-economic differences between people from a BAME background, either natives or migrants, and the natives from a non-BAME background. By comparing the pre-and post-COVID period (2017-2019 vs. April 2020), [Proto and Quintana-Domeque \(2020\)](#) found that the differential impact of COVID-19 on the mental health of BAME versus non-BAME is associated with gender, and they also found heterogeneity across BAME groups. However, none of these studies paid special attention to the people from a BAME background who are at high risk, which could potentially present a group with the highest risk of dying from COVID-19.

To understand more about this issue, this paper pays attention to the high-risk BAME individuals and explores the heterogeneous impacts of the changes in shielding

policies and the epidemiological situation in the UK on the high-risk people between BAME and their non-BAME counterparts. I want to understand whether the lifting of the shielding requirement and the constantly changing epidemiological situation during COVID-19 would affect the high-risk group differently from the low-risk group in the BAME population regarding their physical and mental health compared with the difference in their non-BAME counterparts.

For the rest of this chapter, Section 3.2 introduces background information about this study, and Section 3.3 describes the data and variables. Section 3.4 demonstrates empirical methods. Section 3.5 presents the main results and discusses vaccines and ethnic groups. Section 3.6 discusses robustness. Finally, Section 3.7 concludes.

3.2 Background

3.2.1 How are the “high-risk” individuals defined?

In March 2020, the UK experienced a sharp rise in the number of COVID-19 cases. As a precautionary measure, the UK urged high-risk individuals to shield, meaning they were required to stay at home at all times and avoid face-to-face contact for at least 12 weeks ([Shielding Patient letter, April 2020](#)). The high-risk individuals were informed of the requirement for shielding by a letter from health officials notifying them that they were clinically extremely vulnerable based on their health records. The letter could be shown to their employer (if the person was employed) and used as proof that they should stay at home or work from home if they were able to.

3.2.2 How have shielding policies been changing over time?

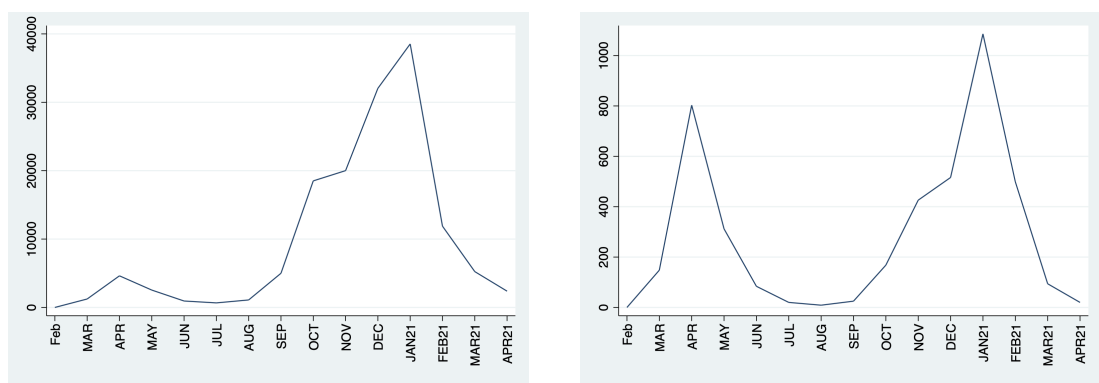
As the first wave of the pandemic improved, the UK started to relax the shielding requirement in June 2020 and completely lifted the shielding requirement in August 2020. The shielding requirements resumed on the 4th of January in 2021 due to

a surge in infection triggered by the Alpha variant of coronavirus from the UK, and a new national lockdown was adopted to prevent hospitals from running out of capacity. However, during the months between August and December 2020, the high-risk individuals were not required to shield.

3.2.3 Lockdown and COVID-19 pandemic

Shielding is a policy only relevant to those who are at high risk. Lockdown, however, affects everyone. The first national lockdown happened in March 2020 and lasted for more than three months. The situation improved as the number of daily new infections went down. By the end of October 2020, the number of daily new infections had been rising, and the UK government implemented a four-tier coronavirus alert system, which ranked each region by its COVID-19 data into four tiers: Tier 1 is “Medium Alert”; Tier 2 is “High Alert”; Tier 3 is “Very High Alert”; Tier 4 is “Stay at home”. Each tier had its own rules. The four-tier system was implemented until it was replaced by another national lockdown in January 2021, when hospitals were almost out of capacity.

[Figure 3.1](#) shows the average daily new COVID-19 infections and deaths in the UK. The number of deaths reached its peak in April 2020 during the first wave. The daily infections in the second wave were much higher than in the first wave. In July 2020, infections reached their lowest point, while deaths reached their lowest in August 2020. Both infections and deaths reached their peak in January 2021. [Figure 3A.1](#) (a) and (b) in [Appendix 3A](#) show the mortality rates among various ethnic groups for males and females, respectively. It is clear from the graph that ethnic minority groups are more affected by the pandemic. However, due to the small proportion of ethnic minority groups in the dataset, this chapter will look at the BAME population as a whole for the main analysis rather than at each ethnic minority group individually.



(a) Average daily new COVID-19 infections

(b) Average daily new deaths due to COVID-19

Figure 3.1: Average daily new COVID-19 infections and deaths

Note: Data is from <https://coronavirus.data.gov.uk/>. Instead of using the daily entry, these plots take the monthly average.

3.3 Data

The data used for this study is Understanding Society COVID-19 Survey in the UK.³ The survey was conducted monthly from April to July 2020, and then bimonthly from July 2020 to March 2021. Up to now, nine waves of surveys have been publicly available (2020: April, May, June, July, September and November; 2021: January, March and September).

For this study, I restrict the sample to waves 1-8 (April 2020 to March 2021). Although data are available for September 2021, I exclude them from the analysis due to the prevalence of vaccination. For example, by the end of September 2021 (wave 9), the first-dose vaccination rate in the UK was around 73%, and the double-dose vaccination rate was around 67%. The fact that vaccines became available provided people, especially high-risk individuals, with protection confounds the results related to COVID-19.

This study uses the main data file in which the respondents are 16 or above, and

³University of Essex, Institute for Social and Economic Research. (2021). Understanding Society: COVID-19 Study, 2020-2021. [data collection]. 9th Edition. UK Data Service. SN: 8644, 10.5255/UKDA-SN-8644-9.

only uses variables from household-level data to control for household characteristics. This study also uses web surveys only and not telephone surveys (available for wave 2 and wave 6 only) because the questionnaires were not exactly the same between web and telephone surveys for a specific wave.⁴

The data is longitudinal. It does aim to keep track of the same group of people over time; however, not everyone responds to all waves of the survey. So it is an unbalanced panel data. After cleaning the data and dropping observations with missing data in outcome variables, the final number of observations is 100,978⁵.

In terms of demographic data, the survey provides information including age, gender and ethnicity⁶. In addition, the survey also provides baseline information (January-February of 2020) for some variables such as baseline household composition, employment status and marital status; please see [Table 3.1](#) for more details. In the sample, 8.7% are from a BAME background and 7.5% are at high risk, so the high-risk BAME category represents an even smaller proportion - 0.8%. The average age of the sample is 51, and 47% are male.

In the survey, the high-risk respondents were identified through a survey question asking whether they have received notification from the medical officials saying that they should shield, meaning staying at home and having no contact with anyone outside their household. (See Section 3.2.1)

⁴Institute for Social and Economic Research (2021) Understanding Society COVID-19 User Guide. Version 8.0, May 2021. Colchester: University of Essex.

⁵The observations with missing data in control variables are not dropped. Instead, I include a dummy variable indicating whether the corresponding control variable has missing data for the observation. Then all the missing data for that variable is filled with the same number.

⁶BAME in this study represents any ethnic groups that are not white, including the mixed races between white and other ethnic groups.

Table 3.1: Summary statistics for individual characteristics

	No. of Obs.	Mean	Std. Dev.	Min	Max
BAME	84,677	0.085	0.279	0	1
High-risk (NHS Shielded patients)	84,677	0.074	0.262	0	1
Age	84,677	50.755	17.887	17	98
16-35 years	84,677	0.234	0.423	0	1
36-48 years	84,677	0.207	0.405	0	1
49-65 years	84,677	0.327	0.469	0	1
65+ years	84,677	0.232	0.422	0	1
Male	84,677	0.470	0.499	0	1
Employed at baseline	84,529	0.617	0.486	0	1
Years of education at baseline	60,739	12.833	2.418	8	16
Married/Civil partnered at baseline	82,445	0.551	0.497	0	1
Baseline household composition					
# of people in age 0-5	81,693	0.134	0.419	0	3
# of people in age 6-15	81,693	0.326	0.701	0	5
# of people in age 16-pensioner	81,693	1.888	1.328	0	9
# of people who are pensioner	81,693	0.423	0.721	0	3

Note: the summary statistics are weighted using cross-sectional weights that are provided by Understanding Society COVID-19 data.

Table 3.2: Summary statistics for outcomes of interest

	Waves	No. of Obs.			Mean		
		All	BAME	non-BAME	All	BAME	non-BAME
Health condition: An emotional, nervous or psychiatric problem	All	84,677	7,995	76,682	0.029	0.047	0.028
Mental well-being factors							
Depression	All	84,677	7,995	76,682	0.078 (1.024)	0.278 (1.053)	0.060 (1.019)
Anxiety	All	84,677	7,995	76,682	0.010 (1.046)	-0.071 (1.160)	0.018 (1.034)
COVID-19 illness-related variables							
Had COVID-19 symptoms	All	84,677	7,995	76,682	0.053	0.062	0.053
Felt at risk of catching COVID-19 in the future	All except for Apr 2020	71,988	6,700	65,288	0.068	0.134	0.062
Frequency of face-to-face contact	Jun 2020 and Nov 2020	20,387	1,858	18,529	0.122	0.086	0.125
Use of transportation							
Frequency of taking tube	Jun 2020, Nov 2020 and Mar 2021	30,350	2,793	27,557	0.018	0.052	0.015
Frequency of taking bus	Jun 2020, Nov 2020 and Mar 2021	30,350	2,793	27,557	0.065	0.136	0.058
Frequency of taking train	Jun 2020, Nov 2020 and Mar 2021	30,350	2,793	27,557	0.023	0.062	0.020
Frequency of taking car	Jun 2020, Nov 2020 and Mar 2021	30,350	2,793	27,557	0.779	0.633	0.793
Frequency of riding bike	Jun 2020, Nov 2020 and Mar 2021	30,350	2,793	27,557	0.096	0.102	0.096
Frequency of walking	Jun 2020, Nov 2020 and Mar 2021	30,350	2,793	27,557	0.657	0.655	0.658
Frequency of using public transportation	Jun 2020, Nov 2020 and Mar 2021	30,350	2,793	27,557	0.082	0.172	0.073
Vaccination							
Want vaccine	Nov 2020, Jan 2021 and Mar 2021	26,569	2,436	24,133	0.896	0.756	0.909
Had at least one dose of vaccine	Jan 2021 and Mar 2021	19,136	1,736	17,400	0.426	0.342	0.434

Note: the summary statistics are weighted using cross-sectional weights that are provided by Understanding Society survey data. Due to these weights used for this table, the standardised mental well-being factors - depression and anxiety - do not show a mean of 0 and a standard deviation of 1. All variables are binary except for two - depression and anxiety whose standard deviations are presented in the parentheses underneath the corresponding means.

In terms of the outcome variables of interest, this survey asks about COVID-19 symptoms and health conditions including mental ill-health diagnosis and 12 general health questions that focus on mental well-being. The survey also provides a rich set of variables reflecting people’s behaviour including whether they use specific types of transportation and whether they meet friends and families outside their household in person; please see [Table 3.2](#) for details. The definition of these variables is provided below.

The rich information collected for each wave of the survey across time allows me to apply the event study approach to explore how the impacts of the policy and COVID-19 have evolved over time.

Outcomes:

i. COVID-19 related variables

Has had COVID-19-like symptoms: 1 if the respondent has likely ever experienced COVID-19 symptoms (if first time completing the survey), or has experienced symptoms since the last time they completed the survey (if not the first time completing the survey); 0 otherwise.

Risk of contracting COVID-19: 1 if the respondent feels they are likely to contract COVID-19 in the next month; 0 if unlikely.

ii. Mental health

Depression (Mental well-being factor 1): Standardised mental well-being factor from 12 general health questions (with higher weights on the loss of sleep, constantly under strain, problems overcoming difficulties, unhappy or depressed, losing confidence, believing worthless). Please see [Table 3.3](#) for the score coefficients.

Anxiety (Mental well-being factor 2): Standardised mental well-being factor from 12 general health questions (with higher weights on low concentration, not feeling playing a useful role, incapable of making decisions, unable to enjoy day-to-day activities,

unable to face problems, unhappy in general). Please see [Table 3.4](#) for the score coefficients.

Table 3.3: Scoring coefficients from principal component factor analysis

	Depression (Factor 1)	Anxiety (Factor 2)
Hard to concentrate	-0.05120	0.23585
Loss of sleep	0.28056	-0.16995
Unable to play a useful role	-0.16925	0.35825
Incapable of making decisions	-0.14087	0.34169
Constantly under strain	0.29905	-0.17923
Problem overcoming difficulties	0.27034	-0.13195
Unable to enjoy day-to-day activities	-0.13802	0.30785
Unable to face problems	-0.06600	0.25981
Unhappy or depressed	0.23009	-0.07658
Losing confidence	0.23622	-0.08546
Believe worthless	0.19380	-0.05129
Unhappy in general	-0.02789	0.21724
Note: The numbers in red demonstrate in which factor the corresponding mental well-being questions (12 in total) are assigned a higher weight.		

Mental ill-health diagnosis: 1 if the respondent has ever been diagnosed with any emotional, nervous or psychiatric problem (if first time completing the survey) or has been diagnosed since the last time the respondent completed the survey (if the respondent has completed at least one previous monthly COVID-19 survey); 0 otherwise.

iii. Other variables

Face-to-face contact outside household: 1 if the respondent meets frequently with friends and families outside their household; 0 if not frequently. (Frequent: at least several times per week).

Use any public transportation: 1 if the respondent frequently uses any public transportation including tube, bus and train; 0 if infrequent. (Frequent: at least three times a week; infrequent: twice a week or less).

Use train: 1 if the respondent frequently uses trains; 0 if infrequent. (Frequent: at

least three times a week; infrequent: twice a week or less).

Use car: 1 if the respondent frequently uses cars; 0 if infrequent. (Frequent: at least three times a week; infrequent: twice a week or less).

Use bus: 1 if the respondent frequently uses buses; 0 if infrequent. (Frequent: at least three times a week; infrequent: twice a week or less).

Use tube: 1 if the respondent frequently uses tube; 0 if infrequent. (Frequent: at least three times a week; infrequent: twice a week or less).

Use bike: 1 if the respondent frequently uses a bike; 0 if infrequent. (Frequent: at least three times a week; infrequent: twice a week or less).

Walk: 1 if the respondent frequently walks; 0 if infrequent. (Frequent: at least three times a week; infrequent: twice a week or less).

3.4 Methods

3.4.1 Overview

The requirement for the high-risk to stay at home started in March 2020 after evidence emerged that patients with specific underlying health conditions are more likely to become seriously ill once infected with COVID-19. We have seen from Section 3.2 that UK shielding policies have changed several times since June 2020. This means that an individual's eligibility to be shielding was determined by their physical health condition and temporal variation pre- and post-change of policy. With these two sources of variation, I will test whether the impact of shielding policies on high-risk individuals is different between BAME and non-BAME individuals using the triple differences model. The estimate is an intent-to-treat effect of shielding policies because I cannot observe whether the high-risk patients behaved as they were required to.

3.4.2 Estimating the Intent-to-Treat Effects of Shielding Policies Change

I want to investigate whether the change in shielding policy and the epidemiological situation during the COVID-19 pandemic would potentially harm high-risk people from a BAME background. The rationale behind this is that a higher percentage of them work in industries where face-to-face interactions are necessary, which makes them more sensitive to a change in policy. Also, the higher mortality rate among the BAME population than the white population adds extra stress to their mental health and well-being. To tackle this question, I exploit an event study approach, which allows us to explore the dynamics of the impact of the COVID-19 situation on the high-risk BAME population.

The primary regression analyses adjust for individual and socioeconomic characteristics. For example, age is a risk factor for developing severe conditions due to COVID-19. Research also shows that gender plays a vital role in the pandemic: women's mental health is more adversely affected by the COVID-19 pandemic than men's ([Etheridge and Spantig, 2020](#); [Liu et al., 2020](#)). Baseline employment and education reflect the social status and resources the individuals have to reduce the risk of infection or to develop severe conditions after infection. Marital status and household composition are related to the level of support once the individuals caught COVID-19 but also reflect the potential everyday contact they may have.

I estimate the heterogeneous impacts over time using the following event study

model:

$$\begin{aligned}
Y_{it} = & \alpha_0 + \alpha_1 Highrisk_i + \alpha_2 BAME_i + \alpha_3 Highrisk_i * BAME_i + \sum_{\tau=-3}^4 \beta_\tau Month_\tau \\
& + \sum_{\tau=-3}^4 \gamma_\tau Month_\tau * Highrisk_i + \sum_{\tau=-3}^4 \omega_\tau Month_\tau * BAME_i \\
& + \sum_{\tau=-3}^4 \delta_\tau Month_\tau * Highrisk_i * BAME_i + X'_{it}\sigma + u_{it}
\end{aligned} \tag{3.1}$$

where Y_{it} represents outcomes of interest for individual i at month t . Outcomes include COVID-19-related variables and mental health-related variables. Please see Section 3.3 for the definitions of all the outcome variables. $Highrisk_i$ is a binary indicator, equal to one if the individual has ever received a letter that requires them to be shielding. $BAME_i$ is a binary indicator, equal to one if the individual is from any minority ethnic background that is not white, including the mixed races between white and other ethnic groups. $Month_\tau$ demonstrates the month of the survey. The omitted month is June 2020 ($\tau = -1$). X_{it} is a vector of individual-specific covariates, including age group indicators, gender, baseline employment status, baseline marital status and baseline household composition. All standard errors are adjusted for clustering at the individual level.

The main parameter of interest, δ_τ , represents the heterogeneous impact of ending of shielding on outcomes Y_{it} for BAME versus non-BAME high-risk individuals. The interaction $Month_\tau * Highrisk_i$ controls for time-varying differences between high- and low-risk individuals. $Highrisk_i * BAME_i$ controls for permanent ethnicity-specific differences between high- and low-risk individuals. $Month_\tau * BAME_i$ controls for ethnicity-specific trends that are invariant to their health conditions. After controlling for the interaction terms, the remaining variation in Y_{it} contains only the time-varying differences between high- and low-risk individuals from a BAME background compared with their non-BAME counterparts. This event study approach

allows the effect of $Highrisk_i * BAME_i$ to vary separately for each wave and thus allows us to explore how the effects of shielding policies evolve over time.

Table 3.3 presents the descriptive statistics for sub-group individual characteristics. The four sub-groups are high-risk non-BAME, low-risk non-BAME, high-risk BAME and low-risk BAME. Within the non-BAME population, the high-risk are less likely to be in any of the age groups 16-35, 36-48 and 49-65 compared with the low-risk. Within the BAME population, the high-risk are less likely to be in the age group 16-35 but are as likely to be in age groups 36-48 and 49-65 as the low-risk. This suggests that the BAME population may develop critical health conditions at a younger age than the non-BAME population.

For the baseline COVID-19 employment status, in both BAME and non-BAME populations, the high-risk are less likely to be employed compared with the low-risk. Within the high-risk population, BAME individuals are as likely to be employed as non-BAME individuals. However, BAME is less likely to be employed within the low-risk population than non-BAME, which reflects a less stable economic situation in the BAME group.

For the baseline household composition, the BAME respondents, both the high-risk and the low-risk, reported a higher number of household members in all age groups compared with the non-BAME respondents, which implies a potentially higher risk of catching COVID-19.

Table 3.4: Summary statistics for sub-group individual characteristics

	Non-BAME		BAME		Non-BAME	BAME	HR	LR
	HR	LR	HR	LR	HR vs. LR	HR vs. LR	BAME vs. non-BAME	BAME vs. non-BAME
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age groups								
Age 16-35	0.069 (0.253)	0.170 (0.376)	0.205 (0.405)	0.309 (0.462)	-0.064*** (0.011)	-0.060** (0.029)	0.164*** (0.048)	0.162*** (0.016)
Age 36-48	0.110 (0.313)	0.216 (0.411)	0.292 (0.456)	0.323 (0.468)	-0.037*** (0.010)	-0.026 (0.028)	0.128*** (0.048)	0.153*** (0.016)
Age 49-65	0.344 (0.475)	0.373 (0.483)	0.366 (0.483)	0.304 (0.460)	-0.017** (0.008)	-0.013 (0.027)	0.093** (0.043)	0.112*** (0.015)
Age 65+	0.477 (0.500)	0.242 (0.428)	0.137 (0.345)	0.064 (0.245)				
Male	0.46 (0.499)	0.416 (0.493)	0.441 (0.498)	0.39 (0.488)	0.009** (0.004)	0.016 (0.011)	0.021 (0.020)	0.003 (0.005)
Employed (baseline)	0.350 (0.477)	0.643 (0.479)	0.545 (0.500)	0.700 (0.458)	-0.049*** (0.005)	-0.039*** (0.013)	-0.015 (0.022)	-0.053*** (0.006)
Education (baseline)	10.951 (3.083)	11.636 (3.110)	11.901 (3.113)	12.368 (3.104)	0.001 (0.001)	-0.002 (0.003)	0.011** (0.005)	0.011*** (0.001)
Married/Civil partnered (baseline)	0.609 (0.488)	0.635 (0.481)	0.629 (0.485)	0.609 (0.488)	-0.018*** (0.005)	-0.008 (0.014)	-0.014 (0.023)	-0.010* (0.006)
Baseline household composition								
# of people in age 0-5	0.061 (0.281)	0.127 (0.409)	0.148 (0.429)	0.239 (0.540)	0.001 (0.007)	-0.019 (0.014)	0.008 (0.027)	0.024*** (0.005)
# of people in age 6-15	0.147 (0.486)	0.31 (0.678)	0.634 (0.934)	0.62 (0.911)	-0.007 (0.004)	0.008 (0.007)	0.067*** (0.015)	0.035*** (0.003)
# of people in age 16-pensioner	1.251 (1.251)	1.758 (1.226)	2.493 (1.649)	2.557 (1.413)	0.001 (0.003)	0.005 (0.005)	0.048*** (0.009)	0.043*** (0.002)
# of people who are pensioner	0.802 (0.853)	0.454 (0.749)	0.31 (0.632)	0.186 (0.487)	-0.002 (0.005)	0.012 (0.015)	0.050** (0.025)	0.047*** (0.008)
N	958	14,505	161	2,018	15,463	2,179	1,119	16,523

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. This table shows the descriptive statistics for four sub-groups including high-risk non-BAME (Column 1), low-risk non-BAME (Column 2), high-risk BAME (Column 3) and low-risk BAME (Column 4). Note that the number of observations for baseline employment status, baseline marital status and baseline household composition is fewer than the “N” shown on the table as there are missing values for these variables, and “N” represents the number of observations for the variables that do not have missing values (age and gender). Columns (5), (6), (7) and (8) represent marginal effects by regressing indicators of high-risk non-BAME (vs. low-risk non-BAME), high-risk BAME (vs. low-risk BAME), high-risk BAME (vs. high-risk non-BAME) and low-risk BAME (vs. low-risk non-BAME), respectively, on age groups, gender, baseline employment status, baseline education, baseline marital status and household composition plus the missing value control variables. All regressions are estimated by logit.

3.5 Results and Discussion

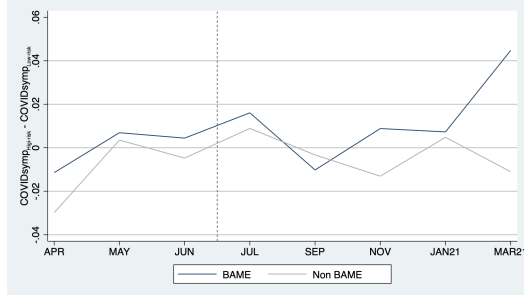
3.5.1 Visuals: BAME versus non-BAME

Figure 3.2 shows a visual depiction of the variation captured by the heterogeneous event study model. The graph plots trends in differences between high- and low-risk individuals in the outcomes of interest, including having likely COVID-19 symptoms, mental ill-health diagnosis, feeling at risk of catching COVID-19, and depression and anxiety. Following the idea in Zuo (2021), I divide the data into two series: one represents BAME and one represents non-BAME. Different from the method used in Zuo (2021), the predicted outcomes are obtained by predicting the corresponding event study model (equation (3.1)) and then taking the mean of the predicted outcomes within each group for each survey month. The difference between the two lines for a month subtracting the difference between the two lines for June 2020 represents the triple differences estimates of δ for that month in equation (3.1).

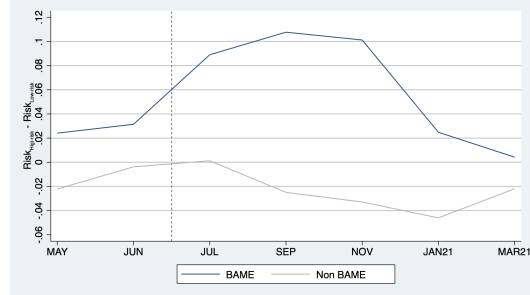
For COVID-19 symptoms (see Figure 3.2(a)), the two series fluctuate within a very small range (5 percentage points). This suggests that the impact on the infection for BAME is similar to that for non-BAME.

For outcomes of feeling at risk of catching COVID-19 and mental ill-health diagnosis (see Figure 3.2(b) and Figure 3.2(c)), the trajectory of the gap between high- and low-risk groups in the BAME population takes off after the lifting of the shielding requirement in July 2020, while for non-BAME the difference is not that obvious before and after the shielding lift.

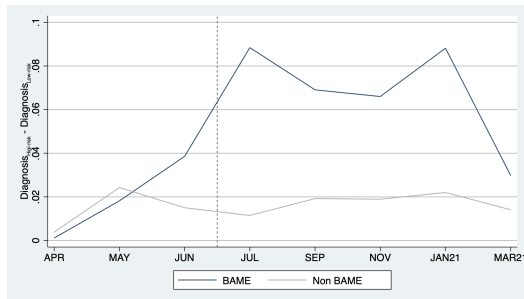
For anxiety (see Figure 3.2(e)), the gap between high- and low-risk groups in the BAME population reaches its peak in June 2020 even before the lifting of shielding, while it remains steady for non-BAME. This could be explained by higher concerns for the high-risk BAME in both economic conditions and health risks.



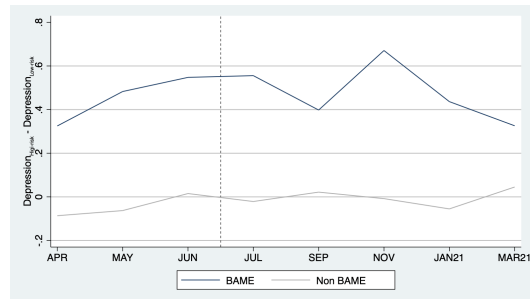
(a) Having had COVID-19-like symptoms



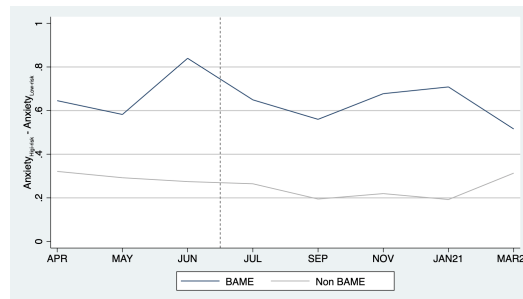
(b) Feeling at risk of contracting COVID-19



(c) Mental ill-health diagnosis



(d) Depression



(e) Anxiety

Figure 3.2: BAME versus non-BAME - Main Outcomes

Note: For each outcome of interest, the blue line shows the difference in the predicted outcomes between high-risk and low-risk for BAME, and the grey line shows the difference in the predicted outcome between high-risk and low-risk for non-BAME.

3.5.2 Main Results

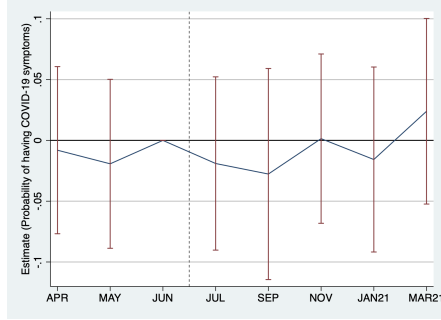
In [Figure 3.3](#), I present intent-to-treat estimates (δ 's in equation (3.1)) of change in shielding policies on having COVID-19 symptoms, feeling at risk of catching COVID-19 in the next month, mental ill-health diagnosis, depression and anxiety. Each estimate represents the difference in the impact of shielding policies on high-risk BAME as compared with high-risk non-BAME. Please see [Table 3A.1](#) for the complete estimates.

In [Figure 3.3](#)⁷, all of the effect sizes in pre-treatment years are statistically insignificant except for anxiety. The point estimates in post-treatment waves fluctuate in COVID-19 symptoms, depression and anxiety. It increases over time for feeling at risk of catching COVID-19 and decreases over time for mental ill-health diagnosis. [Table 3.5](#) provides the precise values of marginal effects depicted in [Figure 3.3](#) and corresponding standard errors.

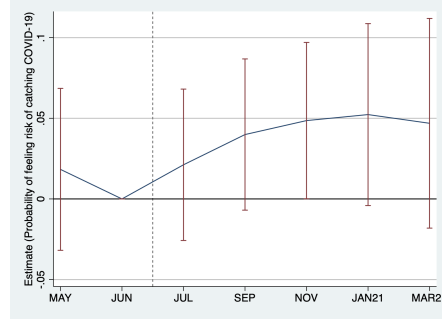
i. COVID-19 symptoms

[Figure 3.3\(a\)](#) presents the estimates for having COVID-19 symptoms. The COVID-19 symptom effect in July 2020 is -0.019 and fluctuates through time till March 2021. However, unlike our hypothesis, no effect is statistically significant. As the confidence intervals are quite wide, I cannot rule out either positive or native differences. So although studies found that BAME as a whole is more likely to be infected by COVID-19 compared with non-BAME ([De Lusignan et al., 2020](#); [Pan et al., 2020](#)), no evidence shows that this is also true for the high-risk subgroup. The results do not provide evidence that shows the high-risk BAME are more likely to have COVID-19 symptoms compared with the high-risk non-BAME as a result of lifting the shielding requirement. Of course, having COVID-19 symptoms does not mean the symptoms

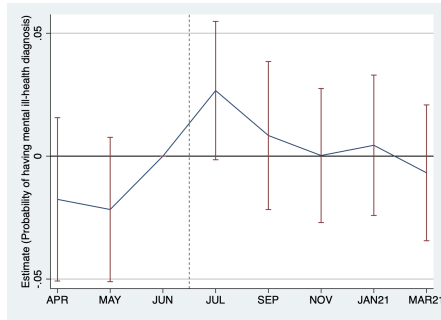
⁷All the graphs in [Figure 3.2](#) are compatible with the corresponding graphs in [Figure 3.3](#). For example, to obtain the event study estimate for November 2020 in [Figure 3.3\(a\)](#), you need to subtract the gap between BAME and non-BAME in June 2020 from the gap between BAME and non-BAME in November 2020 in [Figure 3.2\(a\)](#).



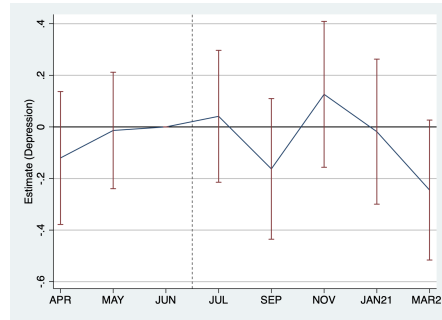
(a) Having had COVID-19-like symptoms



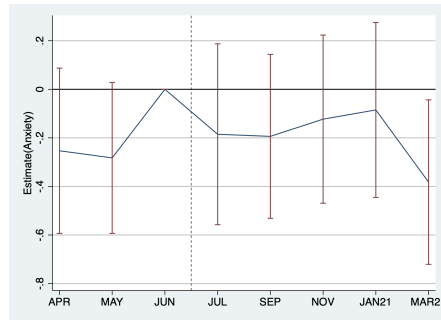
(b) Feeling at risk of contracting COVID-19



(c) Mental ill-health diagnosis



(d) Depression



(e) Anxiety

Figure 3.3: Event Study Estimates

Table 3.5: Event Study of Shielding Requirement Lift and Health Outcomes

	(1) Have had COVID-19 symptoms	(2) Feel at risk of catching COVID-19	(3) Mental ill- health diagnosis	(4) Depression	(5) Anxiety
BAME*High risk*April2020	-0.008 (0.035)		-0.018 (0.017)	-0.120 (0.131)	-0.254 (0.174)
BAME*High risk*May2020	-0.019 (0.035)	0.018 (0.026)	-0.022 (0.015)	-0.014 (0.115)	-0.282* (0.158)
BAME*High risk*July2020	-0.019 (0.036)	0.022 (0.024)	0.027* (0.014)	0.042 (0.130)	-0.186 (0.190)
BAME*High risk*September2020	-0.029 (0.044)	0.041* (0.024)	0.008 (0.015)	-0.163 (0.139)	-0.193 (0.172)
BAME*High risk*November2020	0.002 (0.035)	0.048** (0.025)	0.000 (0.014)	0.126 (0.144)	-0.123 (0.177)
BAME*High risk*January2021	-0.016 (0.039)	0.052* (0.029)	0.004 (0.015)	-0.018 (0.143)	-0.086 (0.184)
BAME*High risk*March2021	0.023 (0.039)	0.047 (0.033)	-0.007 (0.014)	-0.245* (0.138)	-0.382** (0.173)
Control variables	Yes	Yes	Yes	Yes	Yes
N	100,978	85,530	100,978	100,978	100,978

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Columns 1-5 present the event study triple differences estimates for COVID-19 symptoms, feeling at risk of catching COVID-19, mental ill-health diagnosis, depression and anxiety, respectively. All the outcomes but feeling at risk of catching COVID-19 are observed for all eight waves. Feeling at risk of catching COVID-19 is observed in all waves but the first one (April 2020). Columns 1-3 are estimated using logit. Columns 4-5 are estimated using OLS. Control variables include age groups, gender, baseline employment status, baseline education, baseline marital status and baseline household composition.

are necessarily caused by COVID-19, as the symptoms are shared by other illnesses such as flu. But a more plausible explanation is risk compensation that was first raised by Sam Peltzman ([Peltzman, 1975](#)). When people felt they are facing higher risk than before, their behaviour will change to compensate. So the high-risk BAME group may stay at home as much as possible and less frequently expose themselves to crowds in order to reduce their risk of infection. This compensates for the risk of infection increased by lifting shielding requirements.

ii. Feeling at risk of catching COVID-19 in the next month

Figure [3.3\(b\)](#) presents the estimate for feeling at risk of catching COVID-19. The effect on feeling at risk of catching COVID-19 provides a direct measure of the stress that COVID-19 brings to people. This stress could be harmful to their mental well-being. The effect on feeling at risk of catching COVID-19 starts small in July 2020 (0.021), then becomes large and statistically significant in September 2020 (0.040), and continues to grow through January 2021 (0.052). These results are in line with our hypothesis. This could reflect the fact that the COVID-19 infection rate was rising from September 2020 to January 2021, but only in January was shielding required again. The high-risk BAME population face higher risk as they are more likely to work in sectors requiring in-person contact compared with non-BAME and they are no longer required to be shielding even when the infection rate started to soar in November 2020.

iii. Mental ill-health diagnosis

Figure [3.3\(c\)](#) presents the estimate for mental ill-health diagnosis. The effect on mental ill-health diagnosis starts statistically significant in July 2020 at 0.027, then drops to very small and statistically insignificant in later months. From [Figure 3.2\(c\)](#), this peak in July 2020 is driven by an increase in the impact on the high-risk BAME. Although the respondents in the survey may report a mental ill-health diagnosis more

than once⁸, it still suggests two possibilities: first, during the first shielding period, the high-risk BAME population were obeying the rules and trying to postpone seeking professional help for mental ill-health, and the peak in July 2020 is an evidence of delayed diagnosis; second, mental ill-health accumulated during the first three months for the high-risk BAME as they were more affected by the very strict first lockdown and shielding, shortage of food and necessities, and misinformation about the virus and protection measures (Westcott et al., 2021; Hubbard et al., 2021). These results suggest the cost of shielding is that the shielded population has a disruption in the normal process of seeking help for mental ill-health, and the end of shielding means an end of such disruption.

iv. Depression and anxiety

Figures 3.3(d) and 3.3(e) present the estimates of δ 's in equation (3.1) for depression and anxiety factors, respectively. The effect on depression starts statistically insignificant in July 2020 at 0.041 (positive estimate means more depressed). Then it fluctuates through time and eventually becomes statistically significant by March 2021 at -0.245, meaning the impact on the high-risk BAME becomes less depressed by 0.245 points than the impact on high-risk non-BAME. Similarly, the impact on anxiety starts statistically insignificant in July 2020 at -0.185 (negative estimate means less anxious) and it fluctuates through time until it reaches a nadir in March 2021 at -0.382, meaning the impact on the BAME becomes 0.382 less anxious than the impact on non-BAME. From Figure 3.2(d) and 3.2(e), the statistically significant estimate in March 2021 is mainly driven by a drop in the impact on depression and anxiety, respectively, for high-risk BAME. The impact on high-risk non-BAME has been relatively stable over time.

Although these results are not in line with the hypothesis, it is understandable

⁸For example, the respondents of the survey were told by a doctor that they had mental ill-health in March 2020, and when they revisited the doctor in July 2020, they were told that the mental ill-health still existed. In this case, I observe twice mental ill-health diagnosis for this respondent.

given that mental well-being reflects not only the direct effects of the COVID-19 pandemic but other aspects of life as well (part of which could be differentially affected among BAME and non-BAME individuals by COVID-19 pandemic such as employment, travelling, socialising, etc.), they are not direct measures of the shielding policy impact. This could be the reason why the impact from July 2020 to January 2021 is statistically insignificant.

3.5.3 Supplementary results

Trying to explain the main results above, in this section, I present the results for two behavioural outcomes: face-to-face contact with friends and families from outside the household and the use of transportation. The former information is available in the surveys conducted in June and November 2020, while the latter is available in the surveys conducted in June 2020, November 2020 and March 2021. June 2020 is the reference month and thus is omitted in the regression.

i. Face-to-face contact with friends and families who do not live together

Table 3.6 presents marginal effect estimates from a logit regression for face-to-face contact with friends and families from outside the household. It shows that the high-risk BAME was 0.124 percentage points less likely to have frequent face-to-face contact with friends and families outside their households compared with the high-risk non-BAME in November 2020 when the COVID-19 pandemic became worse. The estimate is statistically significant at 10% level.

ii. Use of transportation

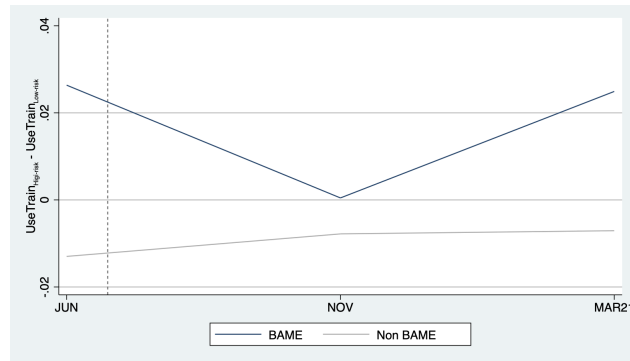
Table 3.7 shows the estimates of the impact of using different types of transportation. Column 1 demonstrates the effect on using any public transportation, including tube, train and bus. Then each column (2-7) represents one means of transportation. I find that the impact on train usage for high-risk BAME was 3.8% less in November 2020

Table 3.6: Face-to-face contact with families and friends who do not live with them

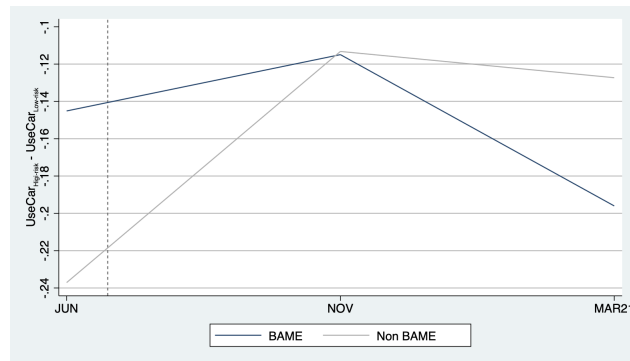
	Frequent contact (1)
BAME*High risk*November2020	-0.124* (0.073)
N	24,288

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. This binary outcome is estimated by logit. Frequent contact means meeting families or friends who do not live together at least several times per week. The model is controlled by age, gender, baseline employment status, baseline education, baseline marital status, and household composition. This variable is available in two waves of survey data (June 2020 and November 2020). June 2020 is omitted in the regression.

than for high-risk non-BAME, while in March 2021 this statistically significant difference disappeared. On the other hand, the impact on car usage for high-risk BAME was 8.1% less than for high-risk non-BAME in March 2021. From [Figure 3.4\(a\)](#), the drop in the difference in train usage was mainly driven by the reduced usage of trains in the BAME group, while from [Figure 3.4\(b\)](#), the reduced car usage in March 2021 was driven by a drop in car usage in high-risk BAME individuals. No significant difference is found between high-risk BAME and high-risk non-BAME in other ways of commuting. (For more complete results, see [Table 3A.2](#).)



(a) Use Train



(b) Use Car

Figure 3.4: BAME versus non-BAME - Use of Transportation

Note: For each outcome of interest, the blue line shows the difference in the predicted outcomes between high-risk and low-risk for BAME, and the grey line shows the difference in the predicted outcome between high-risk and low-risk for non-BAME.

Table 3.7: Use of Transportation

	Use Any Public Transportation (1)	Use Tube (2)	Use Bus (3)	Use Train (4)	Use Car (5)	Use Bike (6)	Walk (7)
BAME*High risk*November2020	-0.016 (0.026)	-0.021 (0.017)	0.013 (0.026)	-0.038* (0.022)	-0.045 (0.040)	0.021 (0.062)	-0.032 (0.066)
BAME*High risk*March2021	-0.006 (0.026)	0.006 (0.022)	0.028 (0.025)	-0.030 (0.022)	-0.081** (0.040)	0.001 (0.071)	-0.037 (0.066)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	35,870	35,870	35,870	35,870	35,870	35,870	35,870

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. All regressions are estimated by logit. The model is controlled by age, gender, baseline employment status, baseline education, baseline marital status, and household composition. This variable is available in three waves of survey data (June 2020, November 2020 and March 2021). June 2020 is omitted in the regression.

These results suggest that high-risk BAME individuals were more careful when the COVID-19 pandemic became worse as they reduced train usage in November 2020 when infections surged. However, the results also suggest that although the government recommended that people, especially those at high risk, should avoid using public transportation as much as possible, those from a BAME background were less likely to have alternative options. Their reduced car usage in March 2021 suggests that they may have gone back to using public transportation. Whether they bought a car, rented a car, or carpooled with other people during this time would be interesting to explore in the future.

3.5.4 Does the vaccine help?

Although this paper does not aim to provide a full analysis of the COVID-19 vaccination programme that started on 8 December 2020 in the UK, I want to check whether vaccination can explain the results above. In particular, I want to know whether the attitudes towards vaccines are different between high- and low-risk BAME compared with non-BAME, by estimating the model below using the data from November 2020, January 2021 and March 2021 waves:

$$Y_{it} = \alpha_0 + \alpha_1 Highrisk_i + \alpha_2 BAME_i + \alpha_3 Highrisk_i * BAME_i + \beta_1 Jan2021 + \beta_2 Mar2021 + X'_{it}\sigma + u_{it} \quad (3.2)$$

where Y_{it} represents outcomes of vaccination for individual i at month t . Outcomes include whether the individuals want the vaccine and whether the individuals had received at least one dose of the vaccine:

Want vaccine: 1 if the respondent has received at least one dose of the COVID-19 vaccine or has no vaccine yet but wants it when it is available; 0 if never had the vaccine and does not want vaccine even when it is available.

Had vaccine: 1 if the respondent has received at least one dose of COVID-19 vaccine;

0 if never had COVID-19 vaccine.

$Highrisk_i$ is a binary indicator, equal to one if the individual has ever received a letter that requires them to be shielding. $BAME_i$ is a binary indicator, equal to one if the individual is from a minority ethnic background. $Jan2021$ and $Mar2021$ are two survey month indicators that capture time trends. X_{it} is a vector of individual-specific covariates, including age group indicators, gender, baseline employment status, baseline marital status and baseline household composition.⁹ All standard errors are adjusted for clustering at the individual level.

Table 3.8 presents results for vaccination. Column 1 shows estimates for the outcome “whether the respondent wanted the vaccine” from a logit regression based on equation (3.2). This outcome is available in the survey months of November 2020, January 2021 and March 2021. I only find a significant difference within low-risk individuals, that is the low-risk BAME is 7.9 percentage points less likely to want vaccine than the low-risk non-BAME. And I do not find that this difference would be statistically significantly larger among the high-risk individuals. Significant effect on time indicators demonstrates a growing willingness for accepting the COVID-19 vaccine.

Column 2 shows the results for whether the respondent has already had at least one dose of vaccine. It is from a logit regression based on equation (3.2). This outcome information is available in January 2021 and March 2021. I only find a significant difference within non-BAME individuals, that is the high-risk non-BAME is 9.8 percentage points more likely to have had at least one dose of the COVID-19 vaccine than the low-risk non-BAME. And I do not find that this difference would be statistically significantly smaller among the BAME individuals. Significant effect on time indicators demonstrates a growing number of people who have received at least one dose of the COVID-19 vaccine.

⁹The outcome “Want vaccine” is available in three survey months, November 2020, January 2021 and March 2021, so there are two month dummies in the model: January 2021 and March 2021. The outcome “Had vaccine” is available in two survey months, January 2021 and March 2021, so there is one month dummy in the model: March 2021.

Table 3.8: Vaccination

	Want Vaccine (1)	Had Vaccine (2)
BAME	-0.084*** (0.005)	0.001 (0.011)
High risk	0.008 (0.010)	0.099*** (0.011)
BAME * High risk	-0.018 (0.019)	-0.021 (0.033)
January2021	0.079*** (0.003)	-
March2021	0.110*** (0.004)	0.456*** (0.002)
Control variables	Yes	Yes
N	31,301	22,425

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. All regressions are estimated by logit. The model is controlled by age, gender, baseline employment status, baseline education, baseline marital status, and household composition. The dependent variable in Column 1 is available in three waves of survey data (November 2020, January 2021 and March 2021). The dependent variable in Column 2 is available in two waves of survey data (January 2021 and March 2021)

It makes sense that the high-risk individuals are more likely to have at least one dose of the COVID-19 vaccine than the low-risk as the high-risk individuals were prioritized for the vaccine in the UK. And it is an established and widely discussed fact that within low-risk individuals, BAME are less likely to take the vaccine compared with non-BAME. [De Figueiredo et al. \(2020\)](#) found a significant decrease in the belief in vaccine safety in Asian countries such as the Philippines and Indonesia, which could potentially affect the ethnic minority groups in the UK through peer effects. They also found that minority religious groups tend to have lower probabilities of accepting the vaccine. Of course, as more data are becoming available, it would be beneficial to explore and understand the impact of the vaccine on physical and mental well-being in the future. Since no significant difference is found in the attitude toward vaccination between high-risk BAME and high-risk non-BAME, these results do not provide an explanation for the main results discussed above.

3.5.5 Ethnic groups (Black and Asian)

Up to now, the analysis and results show the average for the impacts on all ethnic minority groups (BAME). However, from [Alsan et al. \(2021\)](#), each ethnic minority has had a very different experience during the COVID-19 pandemic, so looking at all the ethnic groups as a whole is very likely to miss important effects on a particular ethnic group.

This section briefly discusses the heterogeneity of the results for subgroups of ethnic minorities including Black and Asian. The tables for these results can be found in [Appendix 3A: Table 3A.3-3A.7](#).

The higher rate of mental ill-health diagnosis in July 2020 was mainly driven by people of Black ethnicity; the decreased impact on anxiety in March 2021 was also driven by the Black group. The only impact that was significant for the Asians was feeling at risk of catching COVID-19 in November 2020 and January 2021. Since mental ill-health and anxiety could be related to factors other than COVID-19 illness,

this tells us that the Black population may face more other issues than catching COVID-19 during the pandemic compared with Asians, while the Asians seemed to be much more concerned about COVID-19 illness and also more careful. For example, the decrease in car usage in March 2021 was driven by the Black population rather than Asians. And the lower probability of having frequent contact with families and friends who do not live together in November 2020 when the COVID-19 pandemic became worse is driven by the Asian population.

3.6 Robustness Check

All regressions are run with several versions: no control variables, with age group controls only, with gender control only, with baseline employment status control only, with baseline marital status only, and with household composition controls only. The results (not presented, but available upon request) are robust with all these specifications.

In case the outcomes of people from the same area may not be independent, for example, people are more likely to have COVID-19 symptoms in an area with a higher COVID-19 infection rate compared with people from other areas, I check the robustness by clustering the standard error at primary sampling unit (PSU) level, instead of individual level. Please see [Appendix 3B](#) for the results. Although some standard errors of the estimates change slightly, the levels of significance do not change.

One concern of this study is multiple hypothesis testing. With the five main outcomes, 34 ($4 * 7 + 1 * 6 = 34$) hypotheses are tested. If we test each hypothesis separately, the probability of having false rejections is $1 - 0.95^{34} = 83\%$ for a critical value of 0.05. [Table 3B.3](#) compares the p-values adjusted for multiple hypothesis testing, the Westfall-Young adjusted p-values¹⁰, with the unadjusted p-values. And

¹⁰Based on the free step-down resampling methodology ([Westfall et al., 1993](#)). Stata command ‘wyoung’ is used to generate this adjusted p-value ([Jones et al., 2019](#)).

it shows all the Westfall-Young adjusted p-values are greater than 0.1. Thus, after adjusting for multiple hypothesis testing, all the estimates in the main regressions are statistically insignificant.

3.7 Conclusion

Using this rich COVID-19 survey data by Understanding Society in the UK, this paper addresses the main questions in a dynamic setting: how are the well-being of high-risk BAME and non-BAME people affected differently by shielding policies and the epidemiological situation during the COVID-19 pandemic?

Leveraging an event study approach, I do not find significant differences between BAME and non-BAME in the high-risk group in terms of having COVID-19 symptoms once the shielding requirement was relaxed. However, since the estimates are very imprecise, I cannot rule out any possibilities. I find that the high-risk BAME did feel more at risk of catching COVID-19 from November 2020 to January 2021 when the daily infection rate became higher and shielding was not required. This feeling could contribute to distress during the COVID-19 pandemic.

Regarding the diagnosis of mental ill-health, the impact on the diagnosis of mental ill-health is significantly higher for the high-risk BAME than their non-BAME counterpart after the shielding was relaxed in July 2020. For the mental health factors, depression and anxiety, the impact on depression and anxiety for the high-risk BAME dropped significantly compared with the high-risk non-BAME in March 2021. However, after adjusting the p-values for multiple hypothesis testing of all the main regressions above, none of the estimates remains statistically significant.

In terms of behaviours, although the high-risk BAME population reduced their use of trains when COVID-19 infections surged, they had fewer options for transportation compared with the high-risk non-BAME. In November 2020 when the pandemic became worse, the high-risk BAME had significantly minimised face-to-face contact

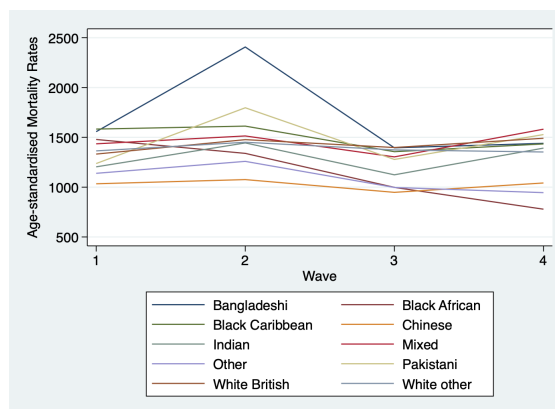
compared with their non-BAME counterpart.

In addition to taking the BAME population as a whole, this study also looks at two individual ethnic minority groups: Black and Asian. In general, Asians were more directly affected by the COVID-19 pandemic itself, while the Black population were more affected by factors that could be associated with the pandemic such as mental health and ways of travelling or commuting.

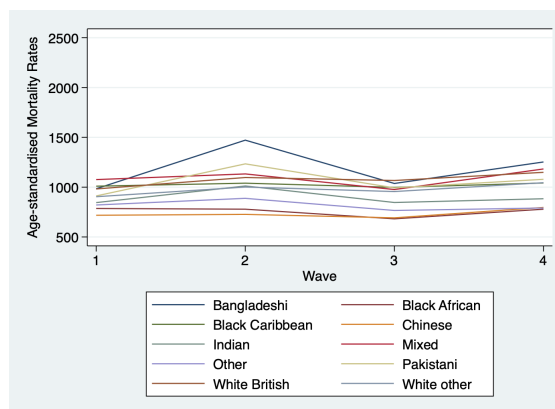
It is unexpected not to find a significant impact of lifting shielding on catching COVID-19 for high-risk BAME individuals. However, since the proxy of having COVID-19 is whether or not the respondents have had “symptoms that could be coronavirus”, respondents may report “yes” even when they had flu instead of coronavirus, as they share some common symptoms. Therefore, to better answer this research question in the future, it would be helpful to distribute antibody tests to the suspected patients. This will improve the data collected by identifying patients who were COVID-positive but only had mild symptoms and therefore received no PCR test.

These results imply two points. First, since lifting the shielding could trigger insecurity felt by the high-risk individuals, especially those from a BAME background, it would be helpful to provide the high-risk individuals with more options (e.g. self-decision of continuing shielding or an official notice of some adjustments on shielding) when the shielding was lifted. Second, due to the significantly higher mental ill-health diagnosis in the high-risk BAME population when the shielding was relaxed, it is likely that the mental ill-health occurred during the shielding period. Therefore, mental health support for the high-risk BAME group is essential, especially during shielding.

Appendix 3A



(a) Male



(b) Female

Figure 3A.1: Mortality across different ethnic groups over waves of COVID-19 pandemic in the UK

Source: Office for National Statistics.

<https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/articles/updatingethniccontrastsindeathsinvolvingthecoronaviruscovid19englandandwales/10january2022to16february2022>

Table 3A.1: Event Study of Shielding Requirement Lift and Health Outcomes

	(1) Have had COVID-19 symptoms	(2) Feel at risk of catching COVID-19	(3) Mental ill- health diagnosis	(4) Depression	(5) Anxiety
BAME	0.002 (0.009)	0.059*** (0.006)	-0.004 (0.006)	0.051* (0.030)	0.004 (0.033)
High risk	0.001 (0.014)	0.016 (0.012)	0.014*** (0.006)	0.161*** (0.038)	0.190*** (0.041)
April2020	0.086*** (0.003)		0.032*** (0.002)	-0.065*** (0.008)	0.092*** (0.011)
May2020	0.020*** (0.004)	0.022*** (0.003)	0.005** (0.002)	-0.039*** (0.008)	0.039*** (0.010)
July2020	-0.007 (0.004)	-0.013*** (0.003)	0.001 (0.002)	-0.066*** (0.008)	-0.081*** (0.009)
September2020	0.021*** (0.004)	0.042*** (0.003)	0.003 (0.002)	0.003 (0.008)	-0.102*** (0.010)
November2020	0.041*** (0.004)	0.047*** (0.003)	0.004* (0.002)	0.072*** (0.009)	0.028** (0.011)
January2021	0.059*** (0.004)	0.044*** (0.003)	0.004* (0.002)	0.022** (0.009)	0.116*** (0.011)
March2021	0.036*** (0.004)	-0.008** (0.004)	0.003 (0.002)	0.031*** (0.009)	-0.022** (0.011)
BAME*High risk	0.012 (0.032)	0.007 (0.020)	0.016 (0.014)	0.476*** (0.132)	0.617*** (0.167)
BAME*April2020	0.005 (0.009)		-0.009 (0.006)	-0.012 (0.031)	0.044 (0.038)
BAME*May2020	0.013 (0.010)	-0.023*** (0.006)	0.006 (0.006)	0.045 (0.028)	-0.058 (0.036)
BAME*July2020	0.017 (0.012)	-0.004 (0.007)	-0.008 (0.007)	-0.023 (0.029)	-0.037 (0.037)
BAME*September2020	0.006 (0.011)	-0.025*** (0.006)	0.000 (0.006)	-0.077*** (0.030)	-0.106*** (0.040)
BAME*November2020	0.013 (0.011)	-0.039*** (0.007)	0.008 (0.006)	-0.023 (0.033)	-0.088** (0.041)
BAME*January2021	0.009 (0.010)	-0.048*** (0.007)	0.006 (0.007)	-0.047 (0.033)	-0.144*** (0.043)
BAME*March2021	0.005 (0.011)	-0.005 (0.007)	0.001 (0.007)	0.001 (0.033)	-0.097** (0.041)
High risk*April2020	-0.002 (0.014)		-0.014*** (0.005)	-0.085** (0.034)	0.041 (0.045)
High risk*May2020	0.017 (0.016)	-0.022 (0.015)	0.004 (0.006)	-0.071** (0.031)	0.019 (0.042)
High risk*July2020	0.031* (0.016)	0.009 (0.014)	-0.003 (0.006)	-0.033 (0.031)	-0.010 (0.040)
High risk*September2020	0.007 (0.017)	-0.016 (0.013)	0.002 (0.006)	0.015 (0.035)	-0.082* (0.045)
High risk*November2020	-0.003 (0.016)	-0.024* (0.013)	0.002 (0.006)	-0.021 (0.036)	-0.055 (0.046)
High risk*January2021	0.016 (0.016)	-0.047*** (0.017)	0.003 (0.006)	-0.070* (0.036)	-0.083 (0.051)
High risk*March2021	-0.003 (0.018)	-0.057** (0.024)	-0.002 (0.007)	0.021 (0.037)	0.039 (0.047)
BAME*High risk*April2020	-0.008 (0.035)		-0.018 (0.017)	-0.120 (0.131)	-0.254 (0.174)
BAME*High risk*May2020	-0.019 (0.035)	0.018 (0.026)	-0.022 (0.015)	-0.014 (0.115)	-0.282* (0.158)
BAME*High risk*July2020	-0.019 (0.036)	0.022 (0.024)	0.027* (0.014)	0.042 (0.130)	-0.186 (0.190)
BAME*High risk*September2020	-0.029 (0.044)	0.041* (0.024)	0.008 (0.015)	-0.163 (0.139)	-0.193 (0.172)
BAME*High risk*November2020	0.002 (0.035)	0.048* (0.025)	0.000 (0.014)	0.126 (0.144)	-0.123 (0.177)
BAME*High risk*January2021	-0.016 (0.039)	0.052* (0.029)	0.004 (0.015)	-0.018 (0.143)	-0.086 (0.184)
BAME*High risk*March2021	0.023 (0.039)	0.047 (0.033)	-0.007 (0.014)	-0.245* (0.138)	-0.382** (0.173)
Control variables	Yes	Yes	Yes	Yes	Yes
N	100,978	85,530	100,978	100,978	100,978

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Columns 1-5 present the event study estimates for COVID-19 symptoms, feeling at risk of catching COVID-19, mental ill-health diagnosis, depression and anxiety, respectively. All the outcomes but feeling at risk of catching COVID-19 are observed for all eight waves. Feeling at risk of catching COVID-19 is observed in all waves but the first one (April 2020). Columns 1-3 are estimated using logit. Columns 4-5 are estimated using OLS. Control variables include age groups, gender, baseline employment status, baseline education, baseline marital status and baseline household composition.

Table 3A.2: Transportation - BAME (complete)

	Use Tube (1)	Use Bus (2)	Use Train (3)	Use Car (4)	Use Bike (5)	Walk (6)	Use Any Public Transportation (7)
BAME	0.023*** (0.003)	0.074*** (0.006)	0.017*** (0.003)	-0.088*** (0.010)	-0.032*** (0.009)	-0.032** (0.015)	0.082*** (0.006)
November2020	0.003* (0.002)	0.025*** (0.003)	0.003 (0.002)	0.038*** (0.004)	-0.040*** (0.003)	-0.035*** (0.005)	0.029*** (0.003)
March2021	-0.001 (0.002)	0.019*** (0.003)	0.002 (0.002)	0.035*** (0.004)	-0.037*** (0.003)	-0.027*** (0.005)	0.021*** (0.003)
BAME*November2020	0.000 (0.003)	-0.009 (0.006)	0.003 (0.003)	-0.030*** (0.012)	0.011 (0.011)	0.012 (0.019)	-0.005 (0.006)
BAME*March2021	0.000 (0.003)	-0.010* (0.006)	0.002 (0.004)	-0.042*** (0.012)	0.009 (0.011)	0.033* (0.019)	-0.009 (0.007)
High risk	-0.028* (0.015)	-0.017 (0.011)	-0.036** (0.018)	-0.151*** (0.011)	-0.048*** (0.013)	-0.177*** (0.017)	-0.029** (0.014)
High risk*November2020	0.021 (0.017)	0.014 (0.010)	0.030 (0.020)	0.066*** (0.014)	0.003 (0.017)	0.085*** (0.020)	0.019 (0.012)
High risk*March2021	0.024 (0.017)	0.017 (0.011)	0.031 (0.020)	0.057*** (0.014)	0.013 (0.017)	0.073*** (0.021)	0.028** (0.013)
BAME*High risk	0.003 (0.022)	-0.009 (0.024)	0.046** (0.020)	0.061* (0.033)	-0.039 (0.047)	0.063 (0.052)	0.022 (0.024)
BAME*High risk*November2020	-0.021 (0.017)	0.013 (0.026)	-0.038* (0.022)	-0.045 (0.040)	0.021 (0.062)	-0.032 (0.066)	-0.016 (0.026)
BAME*High risk*March2021	0.006 (0.022)	0.028 (0.025)	-0.030 (0.022)	-0.081** (0.040)	-0.001 (0.071)	-0.037 (0.066)	-0.006 (0.026)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	35,870	35,870	35,870	35,870	35,870	35,870	35,870

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. All regressions are estimated by logit. The model is controlled by age, gender, baseline employment status, baseline education, baseline marital status, and household composition. This variable is available in three waves of survey data (June 2020, November 2020 and March 2021). June 2020 is omitted in the regression.

Table 3A.3: Event Study of Shielding Requirement Lift and Health Outcomes - Black

	(1) Have had COVID-19 symptoms	(2) Feel at risk of catching COVID-19	(3) Mental ill- health diagnosis	(4) Depression	(5) Anxiety
Black	0.029** (0.014)	0.070*** (0.010)	0.001 (0.011)	-0.114* (0.067)	-0.107 (0.071)
High risk	0.001 (0.014)	0.016 (0.011)	0.015*** (0.006)	0.164*** (0.038)	0.188*** (0.041)
April2020	0.083*** (0.003)		0.032*** (0.002)	-0.066*** (0.008)	0.093*** (0.011)
May2020	0.020*** (0.004)	0.020*** (0.003)	0.005** (0.002)	-0.039*** (0.008)	0.039*** (0.010)
July2020	-0.006 (0.004)	-0.012*** (0.003)	0.001 (0.002)	-0.066*** (0.008)	-0.081*** (0.009)
September2020	0.020*** (0.004)	0.039*** (0.003)	0.003 (0.002)	0.004 (0.008)	-0.102*** (0.010)
November2020	0.040*** (0.004)	0.043*** (0.003)	0.004* (0.002)	0.073*** (0.009)	0.028** (0.011)
January2021	0.057*** (0.003)	0.041*** (0.003)	0.004* (0.002)	0.023** (0.009)	0.116*** (0.011)
March2021	0.035*** (0.004)	-0.008** (0.003)	0.003 (0.002)	0.031*** (0.009)	-0.022** (0.011)
Black*High risk	-0.000 (0.050)	-0.081 (0.054)	0.029 (0.019)	0.212 (0.327)	0.820*** (0.266)
Black*April2020	-0.013 (0.016)		-0.019* (0.011)	0.028 (0.073)	-0.105 (0.080)
Black*May2020	-0.005 (0.018)	-0.031*** (0.011)	-0.028 (0.018)	0.027 (0.066)	-0.101 (0.077)
Black*July2020	0.001 (0.021)	-0.009 (0.013)	-0.025 (0.018)	0.035 (0.067)	-0.020 (0.076)
Black*September2020	-0.018 (0.020)	-0.042*** (0.012)	-0.022 (0.018)	-0.073 (0.066)	-0.055 (0.081)
Black*November2020	-0.041** (0.021)	-0.038*** (0.013)	-0.009 (0.013)	-0.029 (0.075)	0.019 (0.096)
Black*January2021	-0.009 (0.018)	-0.042*** (0.014)	-0.036 (0.022)	0.010 (0.075)	-0.021 (0.095)
Black*March2021	-0.022 (0.019)	-0.005 (0.013)	-0.010 (0.016)	-0.045 (0.076)	-0.096 (0.086)
High risk*April2020	-0.002 (0.014)		-0.015*** (0.005)	-0.085** (0.034)	0.041 (0.045)
High risk*May2020	0.016 (0.015)	-0.021 (0.014)	0.004 (0.006)	-0.071** (0.031)	0.018 (0.042)
High risk*July2020	0.030* (0.015)	0.008 (0.013)	-0.003 (0.006)	-0.034 (0.031)	-0.010 (0.040)
High risk*September2020	0.007 (0.016)	-0.015 (0.012)	0.002 (0.006)	0.014 (0.035)	-0.082* (0.045)
High risk*November2020	-0.003 (0.016)	-0.023* (0.012)	0.001 (0.006)	-0.022 (0.036)	-0.054 (0.046)
High risk*January2021	0.015 (0.015)	-0.043*** (0.015)	0.003 (0.006)	-0.070* (0.036)	-0.083 (0.051)
High risk*March2021	-0.003 (0.017)	-0.052** (0.022)	-0.002 (0.007)	0.020 (0.037)	0.039 (0.047)
Black*High risk*April2020	-0.072 (0.068)		-0.011 (0.022)	0.218 (0.296)	-0.513* (0.308)
Black*High risk*May2020	-0.033 (0.070)	0.029 (0.073)	0.007 (0.021)	-0.090 (0.229)	-0.218 (0.272)
Black*High risk*July2020	-0.012 (0.069)		0.034* (0.019)	0.125 (0.321)	-0.418 (0.431)
Black*High risk*September2020	-0.001 (0.071)	0.116* (0.063)	0.024 (0.022)	-0.121 (0.360)	-0.582* (0.338)
Black*High risk*November2020	-				
Black*High risk*January2021	-0.037 (0.070)		0.041* (0.024)	0.180 (0.346)	-0.399 (0.347)
Black*High risk*March2021	0.038 (0.062)		-0.006 (0.021)	-0.322 (0.296)	-0.570* (0.339)
Control variables	Yes	Yes	Yes	Yes	Yes
N	92,982	187,850	93,000	93,000	93,000

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Columns 1-5 present the event study estimates for COVID-19 symptoms, feeling at risk of catching COVID-19, mental ill-health diagnosis, depression and anxiety, respectively. All the outcomes but feeling at risk of catching COVID-19 are observed for all eight waves. Feeling at risk of catching COVID-19 is observed in all waves but the first one (April 2020). Columns 1-3 are estimated using logit. Columns 4-5 are estimated using OLS. Control variables include age groups, gender, baseline employment status, baseline education, baseline marital status and baseline household composition.

Table 3A.4: Event Study of Shielding Requirement Lift and Health Outcomes - Asian

	(1) Have had COVID-19 symptoms	(2) Feel at risk of catching COVID-19	(3) Mental ill- health diagnosis	(4) Depression	(5) Anxiety
Asian	-0.010 (0.012)	0.062*** (0.007)	-0.014 (0.009)	0.101*** (0.037)	-0.002 (0.042)
High risk	0.001 (0.014)	0.016 (0.011)	0.014*** (0.005)	0.164*** (0.038)	0.191*** (0.041)
April2020	0.084*** (0.003)		0.032*** (0.002)	-0.065*** (0.008)	0.093*** (0.011)
May2020	0.020*** (0.004)	0.022*** (0.003)	0.005** (0.002)	-0.039*** (0.008)	0.039*** (0.010)
July2020	-0.006 (0.004)	-0.013*** (0.003)	0.001 (0.002)	-0.066*** (0.008)	-0.081*** (0.009)
September2020	0.020*** (0.004)	0.041*** (0.003)	0.003 (0.002)	0.003 (0.008)	-0.102*** (0.010)
November2020	0.041*** (0.004)	0.046*** (0.003)	0.004* (0.002)	0.072*** (0.009)	0.028** (0.011)
January2021	0.058*** (0.004)	0.043*** (0.003)	0.004* (0.002)	0.022** (0.009)	0.116*** (0.011)
March2021	0.035*** (0.004)	-0.008** (0.004)	0.003 (0.002)	0.031*** (0.009)	-0.022** (0.011)
Asian*High risk	0.029 (0.038)	0.025 (0.022)	0.000 (0.027)	0.600*** (0.145)	0.540** (0.226)
Asian*April2020	0.008 (0.013)		-0.006 (0.008)	-0.037 (0.039)	0.120** (0.048)
Asian*May2020	0.021 (0.014)	-0.020*** (0.007)	0.009 (0.008)	0.012 (0.036)	-0.056 (0.046)
Asian*July2020	0.015 (0.016)	-0.006 (0.008)	0.005 (0.009)	-0.074** (0.037)	0.004 (0.047)
Asian*September2020	0.015 (0.015)	-0.022*** (0.007)	0.011 (0.008)	-0.095** (0.037)	-0.090* (0.049)
Asian*November2020	0.031** (0.014)	-0.039*** (0.007)	0.013 (0.008)	-0.040 (0.041)	-0.102** (0.051)
Asian*January2021	0.016 (0.014)	-0.046*** (0.008)	0.014 (0.010)	-0.087** (0.042)	-0.126** (0.052)
Asian*March2021	0.015 (0.015)	-0.003 (0.008)	0.007 (0.010)	-0.012 (0.041)	-0.072 (0.051)
High risk*April2020	-0.002 (0.014)		-0.014*** (0.005)	-0.085** (0.034)	0.041 (0.045)
High risk*May2020	0.017 (0.015)	-0.022 (0.014)	0.004 (0.006)	-0.072** (0.031)	0.018 (0.042)
High risk*July2020	0.030* (0.015)	0.008 (0.014)	-0.003 (0.006)	-0.034 (0.031)	-0.011 (0.040)
High risk*September2020	0.007 (0.017)	-0.016 (0.013)	0.002 (0.006)	0.014 (0.035)	-0.082* (0.045)
High risk*November2020	-0.003 (0.016)	-0.024* (0.013)	0.002 (0.006)	-0.022 (0.036)	-0.055 (0.046)
High risk*January2021	0.016 (0.016)	-0.046*** (0.016)	0.003 (0.006)	-0.070* (0.036)	-0.083 (0.051)
High risk*March2021	-0.003 (0.018)	-0.055** (0.023)	-0.002 (0.007)	0.021 (0.037)	0.039 (0.047)
Asian*High risk*April2020	0.008 (0.041)		0.005 (0.032)	-0.172 (0.157)	-0.058 (0.226)
Asian*High risk*May2020	-0.017 (0.039)	0.012 (0.028)		0.034 (0.136)	-0.319 (0.201)
Asian*High risk*July2020	-0.013 (0.040)	0.025 (0.026)	0.020 (0.032)	-0.075 (0.152)	-0.107 (0.234)
Asian*High risk*September2020	-0.072 (0.062)	0.021 (0.025)	0.005 (0.032)	-0.212 (0.161)	-0.080 (0.215)
Asian*High risk*November2020	-0.010 (0.040)	0.047* (0.026)	-0.010 (0.036)	0.075 (0.177)	-0.061 (0.229)
Asian*High risk*January2021	-0.066 (0.052)	0.051* (0.030)	0.006 (0.032)	-0.133 (0.171)	-0.075 (0.233)
Asian*High risk*March2021	0.016 (0.046)	0.053 (0.034)	0.013 (0.021)	-0.210 (0.158)	-0.337 (0.209)
Control variables	Yes	Yes	Yes	Yes	Yes
N	97,132	184,310	97,069	97,132	97,132

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Columns 1-5 present the event study estimates for COVID-19 symptoms, feeling at risk of catching COVID-19, mental ill-health diagnosis, depression and anxiety, respectively. All the outcomes but feeling at risk of catching COVID-19 are observed for all eight waves. Feeling at risk of catching COVID-19 is observed in all waves but the first one (April 2020). Columns 1-3 are estimated using logit. Columns 4-5 are estimated using OLS. Control variables include age groups, gender, baseline employment status, baseline education, baseline marital status and baseline household composition.

Table 3A.5: Face-to-face contact with families and friends who do not live with them
- Asian and Black

	Frequent contact (1)
Panel A: Asian	
Asian*High risk*November2020	-0.161* (0.096)
N	23,399
Panel B: Black	
Black*High risk*November2020	-0.007 (0.141))
N	22,441

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. This binary outcome is estimated by logit. Frequent contact means meeting families or friends who do not live together at least several times per week. The model is controlled by age, gender, baseline employment status, baseline education, baseline marital status, and household composition. This variable is available in two waves of survey data (June 2020 and November 2020). June 2020 is omitted in the regression.

Table 3A.6: Transportation - Asian and Black

	Use Any Public Transportation (1)	Use Tube (2)	Use Bus (3)	Use Train (4)	Use Car (5)	Use Bike (6)	Walk (7)
Panel A: Asian							
Asian*High risk*November2020	-0.023 (0.029)	-	0.008 (0.030)	-0.030 (0.020)	-0.047 (0.045)	0.067 (0.066)	0.046 (0.083)
Asian*High risk*March2021	-0.021 (0.032)	0.176*** (0.020)	0.020 (0.030)	-0.031 (0.021)	-0.033 (0.045)	0.034 (0.077)	0.022 (0.078)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	34,551	34,503	34,551	34,551	34,551	34,551	34,551
Panel B: Black							
Black*High risk*November2020	0.009 (0.038)	-0.019 (0.015)	0.013 (0.039)	-0.042 (0.027)	-0.044 (0.080)	-	-0.224 (0.140)
Black*High risk*March2021	0.024 (0.033)	-0.014 (0.017)	0.035 (0.032)	-0.029 (0.022)	-0.188** (0.091)	-	-0.207 (0.148)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	33,139	33,139	33,139	33,139	33,139	33,100	33,139

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. All regressions are estimated by logit. The model is controlled by age, gender, baseline employment status, baseline education, baseline marital status, and household composition. This variable is available in three waves of survey data (June 2020, November 2020 and March 2021). June 2020 is omitted in the regression.

Table 3A.7: Vaccination - Asian and Black

	Want Vaccine (1)	Had Vaccine (2)
Panel A: Asian		
Asian	-0.064*** (0.006)	0.027** (0.013)
High risk	0.007 (0.010)	0.099*** (0.011)
Asian * High risk	-0.001 (0.024)	-0.059 (0.039)
January2021	0.075*** (0.003)	
March2021	0.104*** (0.004)	0.458*** (0.002)
Control variables	Yes	Yes
N	30,148	21,605
Panel B: Black		
Black	-0.135*** (0.009)	-0.068*** (0.024)
High risk	0.007 (0.009)	0.099*** (0.011)
Black * High risk	-0.016 (0.024)	0.081 (0.081)
January2021	0.073*** (0.003)	
March2021	0.102*** (0.004)	0.460*** (0.002)
Control variables	Yes	Yes
N	28,938	20,747

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. All regressions are estimated by logit. The model is controlled by age, gender, baseline employment status, baseline education, baseline marital status, and household composition. The dependent variable in Column 1 is available in three waves of survey data (November 2020, January 2021 and March 2021). The dependent variable in Column 2 is available in two waves of survey data (January 2021 and March 2021)

Appendix 3B. Robustness

Table 3B.1: Event Study of Shielding Requirement Lift and Health Outcomes with Standard Errors Clustered at PSU Level

	(1) Have had COVID-19 symptoms	(2) Feel at risk of catching COVID-19	(3) Mental ill- health diagnosis	(4) Depression	(5) Anxiety
BAME	0.002 (0.009)	0.059*** (0.006)	-0.004 (0.006)	0.051 (0.032)	0.004 (0.034)
High risk	0.001 (0.014)	0.016 (0.011)	0.014*** (0.005)	0.161*** (0.038)	0.190*** (0.041)
April2020	0.086*** (0.003)		0.032*** (0.002)	-0.065*** (0.009)	0.092*** (0.011)
May2020	0.020*** (0.004)	0.022*** (0.003)	0.005** (0.002)	-0.039*** (0.008)	0.039*** (0.009)
July2020	-0.007 (0.004)	-0.013*** (0.003)	0.001 (0.002)	-0.066*** (0.008)	-0.081*** (0.009)
September2020	0.021*** (0.004)	0.042*** (0.003)	0.003 (0.002)	0.003 (0.008)	-0.102*** (0.011)
November2020	0.041*** (0.004)	0.047*** (0.003)	0.004* (0.002)	0.072*** (0.009)	0.028** (0.011)
January2021	0.059*** (0.004)	0.044*** (0.003)	0.004* (0.002)	0.022** (0.009)	0.116*** (0.011)
March2021	0.036*** (0.004)	-0.008** (0.004)	0.003 (0.002)	0.031*** (0.009)	-0.022* (0.011)
BAME*High risk	0.012 (0.032)	0.007 (0.020)	0.016 (0.014)	0.476*** (0.137)	0.617*** (0.170)
BAME*April2020	0.005 (0.009)		-0.009 (0.006)	-0.012 (0.032)	0.044 (0.038)
BAME*May2020	0.013 (0.010)	-0.023*** (0.006)	0.006 (0.006)	0.045 (0.029)	-0.058 (0.038)
BAME*July2020	0.017 (0.011)	-0.004 (0.007)	-0.008 (0.007)	-0.023 (0.030)	-0.037 (0.038)
BAME*September2020	0.006 (0.011)	-0.025*** (0.006)	0.000 (0.006)	-0.077** (0.031)	-0.106** (0.042)
BAME*November2020	0.013 (0.011)	-0.039*** (0.007)	0.008 (0.006)	-0.023 (0.034)	-0.088** (0.042)
BAME*January2021	0.009 (0.010)	-0.048*** (0.007)	0.006 (0.007)	-0.047 (0.034)	-0.144*** (0.043)
BAME*March2021	0.005 (0.011)	-0.005 (0.008)	0.001 (0.007)	0.001 (0.032)	-0.097** (0.042)
High risk*April2020	-0.002 (0.014)		-0.014*** (0.005)	-0.085** (0.034)	0.041 (0.045)
High risk*May2020	0.017 (0.016)	-0.022 (0.015)	0.004 (0.006)	-0.071** (0.031)	0.019 (0.042)
High risk*July2020	0.031** (0.016)	0.009 (0.014)	-0.003 (0.006)	-0.033 (0.031)	-0.010 (0.040)
High risk*September2020	0.007 (0.017)	-0.016 (0.013)	0.002 (0.006)	0.015 (0.035)	-0.082* (0.045)
High risk*November2020	-0.003 (0.016)	-0.024* (0.013)	0.002 (0.006)	-0.021 (0.036)	-0.055 (0.046)
High risk*January2021	0.016 (0.016)	-0.047*** (0.017)	0.003 (0.006)	-0.070** (0.035)	-0.083* (0.050)
High risk*March2021	-0.003 (0.018)	-0.057** (0.024)	-0.002 (0.007)	0.021 (0.038)	0.039 (0.047)
BAME*High risk*April2020	-0.008 (0.035)		-0.018 (0.017)	-0.120 (0.132)	-0.254 (0.173)
BAME*High risk*May2020	-0.019 (0.035)	0.018 (0.026)	-0.022 (0.015)	-0.014 (0.113)	-0.282* (0.161)
BAME*High risk*July2020	-0.019 (0.035)	0.022 (0.023)	0.027* (0.014)	0.042 (0.132)	-0.186 (0.196)
BAME*High risk*September2020	-0.029 (0.044)	0.041* (0.024)	0.008 (0.015)	-0.163 (0.139)	-0.193 (0.169)
BAME*High risk*November2020	0.002 (0.035)	0.048** (0.024)	0.000 (0.014)	0.126 (0.140)	-0.123 (0.179)
BAME*High risk*January2021	-0.016 (0.039)	0.052* (0.028)	0.004 (0.015)	-0.018 (0.145)	-0.086 (0.198)
BAME*High risk*March2021	0.023 (0.039)	0.047 (0.034)	-0.007 (0.014)	-0.245* (0.137)	-0.382** (0.181)
Control variables	Yes	Yes	Yes	Yes	Yes
N	100,978	85,530	100,978	100,978	100,978

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Columns 1-5 present the event study estimates for COVID-19 symptoms, feeling at risk of catching COVID-19, mental ill-health diagnosis, depression and anxiety, respectively. All the outcomes but feeling at risk of catching COVID-19 are observed for all eight waves. Feeling at risk of catching COVID-19 is observed in all waves but the first one (April 2020). Columns 1-3 are estimated using logit. Columns 4-5 are estimated using OLS. Control variables include age groups, gender, baseline employment status, baseline education, baseline marital status and baseline household composition. Standard errors are clustered at the primary sampling unit level.

Table 3B.2: Transportation with Standard Errors Clustered at PSU Level

	Use Tube (1)	Use Bus (2)	Use Train (3)	Use Car (4)	Use Bike (5)	Walk (6)	Use Any Public Transportation (7)
BAME	0.023*** (0.003)	0.074*** (0.006)	0.017*** (0.003)	-0.088*** (0.011)	-0.032*** (0.009)	-0.032** (0.016)	0.082*** (0.007)
November2020	0.003* (0.002)	0.025*** (0.003)	0.003 (0.002)	0.038*** (0.004)	-0.040*** (0.003)	-0.035*** (0.006)	0.029*** (0.003)
March2021	-0.001 (0.002)	0.019*** (0.003)	0.002 (0.002)	0.035*** (0.004)	-0.037*** (0.003)	-0.027*** (0.006)	0.021*** (0.003)
BAME*November2020	0.000 (0.003)	-0.009 (0.006)	0.003 (0.003)	-0.030*** (0.012)	0.011 (0.011)	0.012 (0.019)	-0.005 (0.006)
BAME*March2021	0.000 (0.003)	-0.010* (0.006)	0.002 (0.004)	-0.042*** (0.012)	0.009 (0.011)	0.033* (0.019)	-0.009 (0.007)
High risk	-0.028* (0.015)	-0.017 (0.011)	-0.036** (0.018)	-0.151*** (0.011)	-0.048*** (0.013)	-0.177*** (0.017)	-0.029** (0.014)
High risk*November2020	0.021 (0.017)	0.014 (0.010)	0.030 (0.020)	0.066*** (0.014)	0.003 (0.017)	0.085*** (0.020)	0.019 (0.012)
High risk*March2021	0.024 (0.017)	0.017 (0.010)	0.031 (0.020)	0.057*** (0.014)	0.013 (0.017)	0.073*** (0.021)	0.028** (0.013)
BAME*High risk	0.003 (0.021)	-0.009 (0.023)	0.046** (0.019)	0.061* (0.033)	-0.039 (0.047)	0.063 (0.052)	0.022 (0.024)
BAME*High risk*November2020	-0.021 (0.017)	0.013 (0.026)	-0.038* (0.022)	-0.045 (0.039)	0.021 (0.062)	-0.032 (0.066)	-0.016 (0.025)
BAME*High risk*March2021	0.006 (0.022)	0.028 (0.025)	-0.030 (0.022)	-0.081** (0.039)	-0.001 (0.072)	-0.037 (0.064)	-0.006 (0.026)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	35,870	35,870	35,870	35,870	35,870	35,870	35,870

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. All regressions are estimated by logit. The model is controlled by age, gender, baseline employment status, baseline education, baseline marital status, and household composition. This variable is available in three waves of survey data (June 2020, November 2020 and March 2021). June 2020 is omitted in the regression. Standard errors are clustered at the primary sample unit level.

Table 3B.3: Unadjusted p-value versus Westfall-Young adjusted p-value

	Have had COVID-19 symptoms		Feel at risk of catching COVID-19		Mental ill-health diagnosis		Depression		Anxiety	
	Unadjusted p-value	Westfall-Young adjusted p-value	Unadjusted p-value	Westfall-Young adjusted p-value	Unadjusted p-value	Westfall-Young adjusted p-value	Unadjusted p-value	Westfall-Young adjusted p-value	Unadjusted p-value	Westfall-Young adjusted p-value
BAME*High risk*April2020	0.813	1.000			0.298	1.000	0.362	1.000	0.144	0.867
BAME*High risk*May2020	0.592	1.000	0.480	1.000	0.147	1.000	0.905	1.000	0.075	0.733
BAME*High risk*July2020	0.593	1.000	0.368	1.000	0.063	0.933	0.748	1.000	0.327	1.000
BAME*High risk*September2020	0.515	1.000	0.089	0.867	0.594	1.000	0.241	1.000	0.261	0.967
BAME*High risk*November2020	0.964	1.000	0.050	0.767	0.985	1.000	0.381	1.000	0.486	1.000
BAME*High risk*January2021	0.685	1.000	0.069	0.867	0.758	1.000	0.902	1.000	0.641	1.000
BAME*High risk*March2021	0.548	1.000	0.155	0.967	0.627	1.000	0.077	0.867	0.027	0.167

References

- Alsan, M., Chandra, A. and Simon, K. (2021). The Great Unequalizer: Initial Health Effects of COVID-19 in the United States. *Journal of Economic Perspectives*, 35(3), 25-46.
- Antoun, J., Brown, D.J., Jones, D.J.W., Sangala, N.C., Lewis, R.J., Shepherd, A.I., McNarry, M.A., Mackintosh, K.A.; Mason, L.; Corbett, J. and Saynor, Z.L. (2021). Understanding the Impact of Initial COVID-19 Restrictions on Physical Activity, Wellbeing and Quality of Life in Shielding Adults with End-Stage Renal Disease in the United Kingdom Dialysing at Home versus In-Centre and Their Experiences with Telemedicine. *International Journal of Environmental Research and Public Health*, 18(6), 3144. <https://doi.org/10.3390/ijerph18063144>
- Bhala, N., Curry, G., Martineau, A. R., Agyeman, C., Bhopal, R. (2020). Sharpening the global focus on ethnicity and race in the time of COVID-19. *Lancet*, 395, 1673–1676.
- Butcher, B. and Massey, J. (19 June 2020). Why are more people from BAME backgrounds dying from coronavirus? *BBC News*. <https://www.bbc.co.uk/news/uk-52219070>
- Clark, A., Jit, M., C. Warren-Gash, C., Guthrie, B., Wang, H.H.X., Mercer, S.W., Sanderson, C., Mckee, M., Troeger, C., Ong, K.L., Checchi, F., Perel, P., Joseph, S., Gibbs, H.P., Banerjee, A., Eggo, R.M., Nightingale, E.S., O'reilly, K., Jombart, T., Edmunds, W.J., Rosello, A., Sun, F.Y., Atkins, K.E., Bosse, N.I., Clifford, S., Russell, T.W., Deol, A.K., Liu, Y., Procter, S.R., Leclerc, Q.J., Medley, G., Knight, G., Munday, J.D., Kucharski, A.J., Pearson, C.a.B., Klepac, P., Prem, K., Houben, R.M.G.J., Endo, A., Flasche, S., Davies, N.G., Diamond, C., Van Zandvoort, K., Funk, S., Auzenberg, M., Rees, E.M., Tully, D.C., Emery, J.C., Quilty, B.J., Abbott, S., Villabona-Arenas, C.J., Hué, S., Hellewell, J., Gimma, A., Jarvis, C.I. (2020). Global, regional, and national estimates of the population at increased risk of severe COVID-19 due to underlying health conditions in 2020: a modelling study. *The Lancet Global Health*, 8, e1003–e1017.

Daly, M., Sutin, A., Robinson, E. (2020). Longitudinal changes in mental health and the COVID-19 pandemic: Evidence from the UK Household Longitudinal Study. *Psychological Medicine*, 1-10. doi:10.1017/S0033291720004432

De Figueiredo, A., Simas, C., Karafillakis, E., Paterson, P. and Larson, H. J. (2020). Mapping global trends in vaccine confidence and investigating barriers to vaccine uptake: a large-scale retrospective temporal modelling study. *Lancet*, 396, 898-908.

De Lusignan, S., Dorward, J., Correa, A., Jones, N., Akinyemi, O., Amirthalingam, G., Andrews, N., Byford, R., Dabrera, G., Elliot, A., Ellis, J., Ferreira, F., Bernal, J.L., Okusi, C., Ramsay, M., Sherlock, J., Smith, G., Williams, J., Howsam, G., Zambon, M., Joy, M., and Hobbs, F.D.R. (2020). Articles Risk factors for SARS-CoV-2 among patients in the Oxford Royal College of General Practitioners Research and Surveillance Centre primary care network: a cross-sectional study. *Lancet Infect Dis*, 20(9), 1034-1042. doi: 10.1016/S1473-3099(20)30371-6.

Etheridge, B. and Spantig, L. (2020). The gender gap in mental well-being during the Covid-19 outbreak: Evidence from the UK. *ISER Working Paper Series, No. 2020-08*. University of Essex, Institute for Social and Economic Research (ISER), Colchester.

Hu, Y. (2020). Intersecting ethnic and native–migrant inequalities in the economic impact of the COVID-19 pandemic in the UK. *Research in social stratification and mobility*, 68, 100528.

Hubbard, G., Den Dass, C., Johnston, M., Murchie, P., Thompson, C.W. and Dixon, D. (2021). Are Rurality, Area Deprivation, Access to Outside Space, and Green Space Associated with Mental Health during the COVID-19 Pandemic? A Cross Sectional Study (CHARIS-E). *International Journal of Environmental Research and Public Health*, 18(8), 3869. doi: 10.3390/ijerph18083869

Hvide, Hans.K. and Johnsen, J. (2021). COVID-19 and Mental Health: a Longitudinal Population Study from Norway. *CEPR Discussion Paper 16195*. <https://cepr.org>

g/publications/dp16195

Iob, E., Frank, P., Steptoe, A. and Fancourt, D. (2020). Levels of Severity of Depressive Symptoms Among At-Risk Groups in the UK During the COVID-19 Pandemic. *JAMA Netw Open*, 3(10), e2026064. doi:10.1001/jamanetworkopen.2020.26064

Jones, D., Molitor, D., and Reif, J. (2019). What Do Workplace Wellness Programs Do? Evidence from the Illinois Workplace Wellness Study. *Quarterly Journal of Economics*, 134(4), 1747-1791.

Kemp, O., Horne, G.A., and Soutar, R. (2020). The psychological impact of COVID-19 on a shielding high-risk cohort. *Scottish Medical Journal*, 65(4), 120–122.

Laurencin, C. T., McClinton, A. J. (2020). The COVID-19 pandemic: A call to action to identify and address racial and ethnic disparities. *Racial and Ethnic Health Disparities*, 7(3), 398–402. <https://doi.org/10.1007/s40615-020-00756-0>

Liu, N., Zhang, F., Wei, C., Jia, Y., Shang, Z., Sun, L., Wu, L., Sun, Z., Zhou, Y., Wang, Y. and Liu, W. (2020). Prevalence and predictors of PTSS during COVID-19 outbreak in China hardest-hit areas: Gender differences matter. *Psychiatry Research*, 287:112921. doi: 10.1016/j.psychres.2020.112921.

Nazroo, J., Murray, K., Taylor, H., Bécaries, L., Field, Y., Kapadia, D. and Rolston, Y. (2020). Rapid Evidence Review: Inequalities in relation to COVID-19 and their effects on London. <https://data.london.gov.uk/dataset/rapid-evidence-review-inequalities-in-relation-to-covid-19-and-their-effects-on-london>

Office for National Statistics. Updating ethnic contrasts in deaths involving the coronavirus (COVID-19), England and Wales: deaths occurring 2 March to 28 July 2020. 16 October 2020. <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/articles/updatingethniccontrastsindeathsinvolvingthecoronaviruscovid19englandandwales/deathsoccurring2marchto28july2020>

Pan, D., Sze, S., Minhas, J.S., Bangash, M.N., Pareek, N., Divall, P., Williams,

C.M.L., Oggioni, M.R., Squire, I.B., Nellums, L.B., Hanif, W., Khunti, K. and Pareek, M. (2020). The impact of ethnicity on clinical outcomes in COVID-19: a systematic review. *EClinicalMedicine*, 23, 100404. doi:10.1016/j.eclinm.2020.100404

Peltzman, S. (1975). The Effects of Automobile Safety Regulation. *Journal of Political Economy*, 83(4), 677-725.

Pierce, M., Hope, H., Ford, T., Hatch, S., Hotopf, M., John, A., Kontopantelis, E., Webb, R., Wessely, S., McManus, S. and Abel, K.M. (2020). Mental health before and during the COVID-19 pandemic: A longitudinal probability sample survey of the UK population. *The Lancet Psychiatry*, 7, 883-892. doi:10.1016/S2215-0366(20)30308-4

Platt, L. and Warwick, R. (2020). Are some ethnic groups more vulnerable to COVID-19 than others? <https://ifs.org.uk/inequality/chapter/are-some-ethnic-groups-more-vulnerable-to-covid-19-than-others/>

Proto, E. and Quintana-Domeque, C. (2020). COVID-19 and Mental Health Deterioration among BAME Groups in the UK. *IZA DP No. 13503*. <https://www.econstor.eu/bitstream/10419/223945/1/dp13503.pdf>

Shielding Patient letter, April 2020 Advice and Actions. <https://southreddishmedicalcentre.co.uk/shielding-patient-letter-april-2020-advice-and-actions-05-04-2020/>

Tai, D.B.G., Shah, A., Doubeni, C.A., Sia, I.G. and Wieland, M.L. (2021). The Disproportionate Impact of COVID-19 on Racial and Ethnic Minorities in the United States. *Clinical Infectious Diseases*, 72(4), 703-706. doi: 10.1093/cid/ciaa815

Taylor, B., Tod, A., Gardiner, C., Creech, L., Forde, F. and Darlison, L. (2021). Effects of the COVID-19 pandemic on people with mesothelioma and their carers. *Cancer Nursing Practice*, 20 (2), 22-28. doi:10.7748/cnp.2021.e1773

Westcott KA., Wilkins F., Chancellor A., Anderson, A., Doe, S., Echevarria, C., and Bourke, S.J. (2021). The impact of COVID-19 shielding on the wellbeing, mental health and treatment adherence of adults with cystic fibrosis. *Future Healthcare*

Journal, 8(1), e47-e49.

Westfall, P.H., Young, S.S. and Wright, S.P. (1993). On Adjusting P-Values for Multiplicity. *Biometrics*, 49(3), 941-945.

Zuo, G.W. (2021). Wired and Hired: Employment Effects of Subsidized Broadband Internet for Low-Income Americans. *American Economic Journal: Economic policy*, 13 (3), 447-482.