5-2018

Essays in the economics of health and ageing

Yi Jin TAN
Singapore Management University, yijin.tan.2013@phdecons.smu.edu.sg

Essays in the Economics of Health and Ageing

by
Yi Jin Tan

Submitted to the School of Economics in partial fulfilment
of the requirements for the Degree of Doctor of Philosophy in Economics

Dissertation Committee:

Jing Li (Chair/Advisor)
Assistant Professor of Economics
Singapore Management University

Tomoki Fujii
Associate Professor of Economics
Singapore Management University

Christine Ho
Assistant Professor of Economics
Singapore Management University

Rhema Vaithianathan
Professor of Economics
Director, Singapore Life Panel
Co-Director, Centre for Social Data Analytics at AUT

Singapore Management University
2018

Contents

Abstract ... i
Acknowledgements ... iii
1 Introduction ... 4
2 The effect of environmental disamentities on electricity consumption: Evidence from dengue outbreaks in Singapore ... 9
2.1 Introduction ... 9
2.2 Background ... 13
2.3 Data and Variables ... 15
2.3.1 Data Sources and Construction of the Combined Dataset ... 15
2.3.2 Construction of Key Variables ... 16
2.3.3 Summary Statistics ... 16
2.4 Identification Strategy and Empirical Model ... 16
2.5 Main Empirical Specifications and Results ... 18
2.5.1 Effect of Being Near a Dengue Cluster ... 18
2.5.2 Cut-off Distances ... 18
2.5.3 Effect of Proximity to Dengue Cluster on Electricity Consumption ... 20
2.5.4 Dynamic Response to Being Near a Dengue Cluster ... 21
2.5.5 Effect of Risk Classification Within a Cluster ... 21
2.6 Threats to Identification and Robustness Checks ... 23
2.7 Costs of Protective Behaviour ... 25
2.8 Conclusion ... 26
References ... 28
Appendix A: Detailed information on dataset construction ... 41
Appendix B: Tables of coefficients associated with Figure 2.3 ... 43
3 Effects of In-utero Exposure to Maternal Fasting during Ramadan on Subjective Well-being and Health in Old Age ... 46
3.1 Introduction ... 46
3.2 Background and Literature Review ... 49
3.2.1 Ramadan and Fasting ... 49
3.2.2 Mechanisms linking in-utero Ramadan exposure to adult outcomes ... 50
3.2.3 Famines as a natural experiment for studying the effects of foetal malnutrition ... 52
3.2.4 Long-term impact of in-utero exposure to Ramadan fasting ... 52
3.3 Data and Variables ... 54
3.3.1 Data source and sample definition ... 54
3.3.2 Variables ... 54

3.3.3 Summary Statistics 56
3.4 Identification Strategy and Empirical Specifications 59
3.5 Results 61
3.5.1 Life satisfaction and mental well-being 61
3.5.2 Social/Family and Daily activities satisfaction 62
3.5.3 Economic satisfaction 62
3.5.4 Self-rated health 63
3.5.5 Diagnosed medical conditions 63
3.5.6 Anthropometric measures 65
3.6 Validity of identifying assumption and robustness checks 66
3.6.1 Selective timing of pregnancy 66
3.6.2 Differing survey participation behaviours between treatment and control 68
3.6.3 Other regression specifications 68
3.7 Conclusion 68
References 70
Appendix: Additional Tables 85
4 Income and subjective well-being: Evidence from Singapore’s first national non-contributory pension 92
4.1 Introduction 92
4.2 Literature review on the impact of income on subjective well-being 96
4.3 Background on the Silver Support Scheme 98
4.4 Data and variables 100
4.4.1 Data source 100
4.4.2 Variables 101
4.5 Identification strategy and empirical model 103
4.5.1 Identification strategy 103
4.5.2 Empirical specifications 106
4.6 Results 110
4.6.1 Overall effects 110
4.6.2 Heterogeneity by financial preparedness for retirement and assets 111
4.6.3 Heterogeneity by identity of recipients 114
4.7 Robustness Checks 115
4.8 Conclusion 116
References 118
Appendix A 131
5 A pragmatic, randomized controlled trial of a cardiac hospital-to-home transitional care program 138

5.1 Introduction ... 138
5.2 Data and Methods ... 139
5.2.1 CareHub versus Usual Care ... 139
5.2.2 Study Population and Randomisation ... 139
5.2.3 Data sources ... 140
5.2.4 Outcome Measures and Covariates ... 141
5.2.5 Statistical Analyses ... 141
5.2.5.1 Main analyses ... 141
5.2.5.2 Subgroup analyses ... 142
5.2.6 Study Approval ... 143
5.3 Results ... 143
5.3.1 Summary Statistics and Balance Checks ... 143
5.3.2 Primary Outcomes ... 143
5.3.3 Heterogeneous Treatment Effects ... 144
5.3.4 Secondary Outcomes: Survey Results and Clinical Indicators ... 145
5.4 Discussion ... 146
5.4.1 Limitations ... 146
5.4.2 Potential Reasons for Improved Outcomes ... 147
5.4.3 Implications ... 148
5.5 Conclusion ... 149
References ... 150
Supplementary Materials ... 159
Robustness Checks for Main Results ... 159
Decomposition of Total Post-Discharge Inpatient Days ... 163
Results from Cox Models ... 166
Patient Survey ... 169
Methodology ... 169
Results ... 171
Clinical indicators – diastolic and systolic blood pressure ... 175
Comparison of NUH vs non-NUH utilisation for CareHub and non-CareHub groups ... 176

List of tables

Table 2.1: Summary statistics[1] 32
Table 2.2: Effect of being near a dengue cluster on electricity consumption 33
Table 2.3: Decay of persistence 34
Table 2.4: Effect of risk classification on electricity consumption 35
Table 2.5: Robustness Checks 36
Table 2.6a: Total additional electricity consumption and carbon dioxide release[1] 37
Table 3.1: Summary statistics for basic demographics and major outcome variables 73
Table 3.2: Effect of in-utero exposure to Ramadan on overall life satisfaction and mental well-being index 75
Table 3.3: Effect of in-utero exposure to Ramadan on social/family life satisfaction and satisfaction with daily activities 76
Table 3.4: Effect of in-utero exposure to Ramadan on economic satisfaction 77
Table 3.5: Effect of in-utero exposure to Ramadan on general self-rated health 78
Table 3.6: Effect of in-utero exposure to Ramadan on cardiovascular conditions and diabetes 79
Table 3.7: Effect of in-utero exposure to Ramadan on body mass index 80
Table 3.8: Parental characteristics and in-utero exposure to Ramadan 81
Table 3.9: Robustness check for Family Planning Policy (born before 1966) 82
Table 3.10: Relationship between probability of being in-utero during Ramadan and age 83
Table 4.1: Summary statistics for baseline demographics 121
Table 4.2: Summary statistics for pre-announcement dependent variables[1] 122
Table 4.3: Overall impact of the Silver Support Scheme on subjective well-being 123
Table 4.4: Silver Support payouts and subjective well-being – heterogeneous effects by financial preparation for retirement 124
Table 4.5: Silver Support payouts and subjective well-being – heterogeneous effects by baseline assets quintile 125
Table 4.6: Silver Support payouts and subjective well-being – heterogeneous effects by recipient identity 126
Table A1: Summary statistics for baseline demographics 131
Table A2: Silver Support payouts and subjective well-being – heterogeneous effects by flat-type and financial preparedness for retirement 132
Table A3: Robustness checks for overall life satisfaction 133
Table A4: Robustness checks for social and family life satisfaction 134
Table A5: Robustness checks for household income satisfaction 135
Table A6: Robustness checks for economic satisfaction 136
Table A7: Overall impact of the Silver Support Scheme on subjective well-being for full sample 137
Table 5.1: Comparison of study intervention and usual care 152
Table 5.2: Summary Statistics and Balance Check 155
Table 5.3: Effect of CareHub on Key Outcomes 156

Table 5.4: Heterogeneity Analyses 157
Table 5.5: Survey Results 158
Table S1: Robustness Checks for No. of Unplanned, Cardiac-Related Readmissions 160
Table S2: Robustness Checks for Total Number of Days Readmitted for Unplanned, Cardiac-Related Reasons 160
Table S3: Robustness Checks for No. of Cardiac-Related Readmissions 161
Table S4: Robustness Checks for Total Number of Days Readmitted for Cardiac-Related Reasons 161
Table S5: Robustness Checks for Total Number of Cardiac-Related Outpatient Specialist Visits 162
Table S6: Robustness Checks for Total Number of Emergency Department Visits 162
Table S7: Summary Statistics for Variables Used in Computation 165
Table S8: Relative Contribution of Each Effect 165
Table S9: Results from Cox Model Studying First Unplanned, Cardiac-Related Readmission 167
Table S10: Results from Cox Model Studying First Cardiac-Related Readmission 167
Table S11: Results from Cox Model Studying First Cardiac-Related Specialist Outpatient Visit 168
Table S12: Results from Cox Model Studying First Emergency Department Visit 168
Table S13: Comparison of Survey Outcomes for CareHub vs Non-CareHub Groups 173
Table S14: Comparison of Non-NUH Healthcare for CareHub and Non-CareHub Groups 177

List of figures

Figure 2.1: Number of new dengue fever cases each week from 2012 Jan to 2015 Dec 38
Figure 2.2: Geographic and temporal variation in dengue clusters 39
Figure 2.3: Estimated effect of being near a dengue cluster using different cut-off distances..... 40
Figure 3.1: Proportion of Muslims and non-Muslims who were in-utero during Ramadan, by age 84
Figure 4.1: Trends for life satisfaction, for samples trimmed based on increasingly narrower ranges of propensity score 127
Figure 4.2: Trends for life, social, and daily activities satisfaction 128
Figure 4.3: Trends for household income, economic, health satisfaction as well as health condition 129
Figure 4.4: Density plots / histograms of key demographics before (left panel) and after (right panel) trimming.. 130

Abstract

This dissertation consists of four papers in applied microeconomics / the economics of health and ageing that analyse the causal effect of public policies, or of events / issues amenable to policy intervention. Chapter 1 provides an overview of the papers in this dissertation.

Chapter 2 investigates whether a major and growing environmental disamenity – dengue fever – leads to protective behavior that increases residential electricity consumption. Being near a dengue cluster leads to a persistent increase in electricity consumption in 4-room and 5-room/bigger flats (by 1.7% and 1.1% respectively). In addition, electricity consumption rises discontinuously when a dengue cluster's risk classification is upgraded from yellow to red. This increased electricity consumption cost $11.9 to 16.3 million per annum (in 2015 Singapore dollars), or 7% – 12% of the overall costs of dengue in Singapore.

Chapter 3 studies the effect of in-utero exposure to mild nutritional shocks during Ramadan on an individual's later-life outcomes. In-utero exposure to Ramadan leads to poorer subjective well-being across a broad range of domains (overall life, social and family life, daily activities, economic, and health satisfaction), self-rated health condition, and poorer mental well-being. In addition, exposed individuals report higher rates of diagnosed cardiovascular conditions and higher body mass index (among women). We find no evidence that these results are driven by selective timing of pregnancies, differing survey participation rates, or seasonal effects.

Chapter 4 examines the effect of an exogenous permanent income shock on subjective well-being. This permanent income shock is the introduction of Singapore's first national non-contributory pension, the Silver Support Scheme. The pension improved the life satisfaction of recipients; this effect appears to be driven by social, household income, and economic satisfaction. Consistent with the predictions of the permanent income hypothesis, well-being improved at

announcement, but did not improve significantly further at disbursement of the pension. Lastly, we find evidence that the marginal utility of income varies – recipients who reported being less financially prepared for retirement exhibited larger increases in well-being.

Lastly, Chapter 5 reports results from a pragmatic, randomized controlled trial of CareHub, a new transitional care program (TCP) in Singapore's National University Hospital that aims to contain costs, reduce re-hospitalizations, and improve patient quality of life. CareHub reduced unplanned cardiac-related readmissions by 39% and unplanned cardiac-related days in hospital by 56%. In addition, we found suggestive evidence that CareHub reduced patient anxiety and depression, and improved the quality of transitional care. In all, CareHub achieved net cost savings of about S$1,300 per patient over six months, suggesting that a carefully designed TCP can reduce resource utilization while improving quality of life.

Acknowledgements

I would like to thank Jing Li, my advisor, for her advice and support over the years. Jing's style of advising – encouraging me to develop my own ideas and providing the space to pursue a variety of projects, while providing much-appreciated advice and support – has been invaluable in helping me develop as an economist. In addition, I would like to thank my other committee members: Tomoki Fujii, Christine Ho, and Rhema Vaithianathan, all of whom have given very helpful advice and comments over the course of past few years.

I would also like to thank the following people and teams: Anthony Tay, who put in a substantial amount of effort in accommodating special circumstances during my candidature; the SMU SOE faculty, who have provided many helpful comments on my research and from whom I have learnt much; the friendly and extremely efficient administrative staff at SOE and CREA; Phil Phan and the dedicated NUHS team responsible for implementing CareHub, without whom the evaluation of CareHub would not have been possible; as well as my coursemates, for interesting conversations on economics over the years.

Lastly, I'd also like to thank my family for their support. In particular, I'd like to express my appreciation of Yanying, my co-author, best friend, wife, and fellow minion in thrall to our new Evil Overlord. I will miss working so closely with such a creative, sharp, and careful collaborator when we head off to new jobs. Finally, I dedicate this dissertation to our new-born Evil Overlord, who has done absolutely nothing to contribute to it, except deigning to stay quiet for sufficiently long periods of time to allow his newfound minions to finish putting our dissertations together.

1 Introduction

This dissertation consists of four papers in applied microeconomics – in particular, the economics of health and ageing. The topics covered contribute to a number of different fields: health, labour/pensions, and energy economics. The wide range of topics reflects my aim in embarking upon the Ph.D.: to gain the skills necessary to carry out rigorous analysis of public policies. These topics were chosen to allow me to gain experience in analysing Singaporean policy-related issues in different fields and in using different datasets as well as applied microeconomic techniques. The underlying theme uniting these chapters, therefore, is that all chapters analyse the causal effect of public policies, or of events / issues amenable to policy intervention.

Chapter 2 studies whether a major and growing environmental disamenity – dengue fever – leads to protective behavior that increases residential electricity consumption. Singaporean households near clusters of dengue cases are aware that they are exposed to dengue risk as the government puts up prominently displayed alert banners in dengue clusters. These households may then engage in protective behavior to reduce their risk of contracting dengue. We focus on a type of behaviour (which is common enough to be mentioned in news articles in major Singaporean newspapers) that is likely to lead to a rise in electricity consumption: the closure of doors/windows and turning on the air-conditioner to keep mosquitoes out.

We construct a unique panel dataset with nationwide coverage from administrative sources, and use spatial and temporal variation in proximity to dengue clusters to estimate fixed-effects models and identify the causal effect of proximity to dengue clusters on residential electricity consumption. Being near a cluster leads to a rise in electricity consumption for 4-room and 5-room/bigger flats (by 1.7% and 1.1% respectively), but not 3-room flats. This rise persists in the months after exposure to dengue risk. Households' responses also increase discontinuously when dengue risk classification on alert banners is upgraded from yellow (≤9 cases) to red (≥10 cases), which is consistent with our hypothesis that an observed increase in electricity consumption is driven by the protective behaviour of closing doors and windows, as other factors related to electricity consumption are unlikely to vary with dengue risk intensity in this discontinuous manner. After accounting for the social cost of carbon, we find that increased electricity consumption due to dengue outbreaks costs around $11.9 to 16.3 million per annum (in 2015 Singapore dollars), constituting 7% – 12% of the overall costs of dengue in Singapore.

These results mainly contribute to the literature on residential electricity consumption, by providing evidence that a previously unstudied environmental disamenity can lead to a persistent rise in electricity consumption that is economically significant and comparable to falls induced by some interventions that try to reduce electricity consumption (e.g., the social comparisons studied in, say, Allcott, 2011; Ayres et al., 2013). With the expected increase in the geographical range of dengue fever (Hales et al., 2002; Smith et al., 2014), and the likely increase in prevalence of air-conditioners in currently dengue-endemic countries as they grow richer (Isaac & Van Vuuren, 2009; Akpinar-Ferrand & Singh, 2010), our results also suggest that it may become increasingly important to account for responses to environmental disamenities such as vector-borne diseases (of which dengue fever is one) when simulating the effects of climate change on electricity consumption.

Chapter 3 evaluates the effect of in-utero exposure to a mild nutritional shock – maternal fasting during Ramadan – on subjective well-being measures such as life, economic and health satisfaction, as well as mental and physical well-being in old age. Using the Singapore Life Panel (SLP), a new, high-frequency dataset from Singapore (an environment that is well-suited for studying the long-term effects of in-utero undernutrition), and exploiting the plausible randomness of in-utero exposure to maternal fasting during Ramadan as a natural experiment, we find that the relatively mild nutritional shocks experienced during Ramadan can have lasting effects on the long-term outcomes of the child. Exposed individuals have lower life, social/family, daily activities, economic, and health satisfaction, and they rate their own health condition more poorly. They also seem to have poorer mental well-being. In line with predictions from the foetal origins hypothesis (FOH), these individuals are more likely to be diagnosed with cardiovascular conditions and have a higher body-mass index (for women). In our sample, the effects from exposure are most pronounced for individuals exposed in the second trimester. These results are unlikely to be driven by seasonal effects common to all individuals, selective timing of pregnancies, or differing survey participation behaviour.

These results extend the literature by studying exposure effects on subjective well-being and mental well-being, which may paint a fuller picture of an individual's health and socio-economic outcomes. In particular, life satisfaction is widely used as a utility proxy and can provide information on the overall welfare effects of in-utero Ramadan exposure, compared to past studies that study more specific outcomes. Furthermore, our results with respect to anthropometric measures and diagnosed health conditions conform more closely to those predicted by the FOH, strengthening existing evidence on the link between in-utero Ramadan exposure and health. As we study a milder nutritional shock, our findings also add credence to the idea that the negative effects predicted by the FOH may apply to milder nutritional disruptions in developed countries too.

Chapter 4 examines the effect of a permanent income shock on subjective well-being. The permanent income shock we study is the introduction of the Silver Support Scheme (SSS), a new non-contributory pension targeted at the poorest 20 – 30% of Singaporean elderly. Using data from the SLP and a difference-in-differences strategy, we find that recipients of SSS payouts, who receive an average of around S$500 per quarter in our sample, experienced a statistically significant improvement of 2.5% of the baseline mean (or 0.11 SD) in overall life satisfaction upon SSS announcement; there seems to be no additional improvement after the disbursement of SSS payouts (i.e. the disbursement effect is not statistically different from the announcement effect). These results seem to be driven in part by improvements in recipients' satisfaction with their social contacts and family life, household income and economic situation. In addition, we find that the effect of receiving SSS payouts differed by subjective financial preparedness for retirement, but not by asset levels.

Our results mainly contribute to the literature on income and subjective well-being. While the permanent income hypothesis (PIH) implies that forward-looking individuals' utility should react only to unanticipated, but not anticipated income shocks (see e.g. Cai & Park, 2016), empirical evidence on how subjective well-being reacts to unanticipated versus anticipated income shocks is scarce. We add to this literature by using the high-frequency SLP and a natural experiment from Singapore to separately estimate the announcement and disbursement effects of income increases on subjective well-being, and provide evidence that is consistent with the predictions of the PIH. We also add to the evidence on how marginal utility of income may vary by individual characteristics, when we examine heterogeneous effects of receiving SSS payouts along different dimensions.

Finally, Chapter 5 reports results from a pragmatic, randomized controlled trial of CareHub, a new transitional care program (TCP) in Singapore's National University Hospital (NUH). While TCPs have been implemented and evaluated in many settings, we approach the issue of transitional care from a novel angle by asking if further gains can be by integrating and refining existing TCPs within a hospital. CareHub aims to do so by offering a protocolised 'one-stop shop' for post-discharge patient follow-ups.

Using administrative data from NUH, as well as survey data specifically collected for this evaluation, we examine a comprehensive set of primary outcomes on healthcare services utilisation, and an extensive list of secondary outcomes, including patients' quality of life, quality of transitional care, and selected clinical outcomes. We find that CareHub reduced unplanned cardiac-related readmissions by 39% and unplanned cardiac-related days in hospital by 56%. In addition, we found suggestive evidence that CareHub reduced patient anxiety / depression, and improved the quality of transitional care. In all, CareHub achieved net cost savings of about S$1,300 per patient over six months, suggesting that a carefully designed TCP that integrates and refines existing TCPs can reduce resource utilization while improving quality of life.

2 The effect of environmental disamentities on electricity consumption: Evidence from dengue outbreaks in Singapore

(Co-authored with Yanying Chen)

2.1 Introduction

Rising energy consumption is a pressing environmental and economic concern worldwide, and electricity consumption is an important contributor to this rise (International Energy Agency, 2017). The importance of electricity consumption has motivated a large literature on the determinants of residential electricity demand, which includes a strand that uses field or natural experiments to study how policies and nudges can affect residential electricity demand[1]. Another strand considers electricity consumption as an input in the production of household comfort (Quigley & Rubinfeld, 1989), and studies how it can be affected by temperature or local environmental disameneties (e.g. noise or air pollution), if households increase their electricity usage to mitigate these disameneities. While a large literature on the effects of temperature exists[2], fewer papers use exogenous variation to study the effect of localised environmental disamenities (which may be more amenable than temperature to policy intervention) on residential electricity consumption. Agarwal et al. (2016), who look at the effect of nearby construction on residential electricity usage, is the only other paper on this topic that we are aware of.

[1] Most of these policies and nudges lead to reductions in electricity demand. See, e.g., Reiss and White (2008); Allcott (2011); Aroonruengsawat et al. (2012); Ayres et al. (2013); Jacobsen and Kotchen (2013); Delmas and Lessem (2014); Alberini and Towe (2015); Agarwal et al. (2017); List et al. (2017), which look at how interventions such as peer comparisons / public pressure, the use of smart meters, or building codes can reduce residential electricity consumption. Some policies, however, may lead to increases in electricity demand (e.g. Sexton, 2015).

[2] See e.g. Moral-Carcedo and Vicens-Otero (2005); Ihara et al. (2008); Aroonruengsawat and Auffhammer (2011); Deschênes and Greenstone (2011); Lee and Chiu (2011).

In this paper, we study the effect of a major and growing environmental disamenity – dengue fever outbreaks – on residential electricity consumption in Singapore. Dengue fever, a mosquito-borne viral disease, is a global health threat. Approximately 3.9 billion people in 128 countries are at risk of dengue infection, 390 million are infected annually (World Health Organisation, 2016), and the geographical limits of dengue fever are expected to grow with climate change (Hales et al., 2002; Smith et al., 2014).

Households exposed to dengue[3] may engage in different forms of protective behaviour to mitigate this environmental disamenity. We focus on a type of behaviour that is likely to lead to a rise in electricity consumption: the closure of doors/windows and turning on the air-conditioner to keep mosquitoes out. This type of protective behaviour is sufficiently common in Singapore that articles/infographics in major Singaporean newspapers have made reference to it (e.g. Lee & Kumar, 2016; Singapore Power, 2016), and may grow increasingly common in other dengue-endemic countries, as they grow richer and are increasingly able to afford the use of air conditioners (Isaac & Van Vuuren, 2009; Akpinar-Ferrand & Singh, 2010).

We construct a unique monthly panel dataset with nationwide coverage to study the effects of exposure to localised dengue risk on electricity consumption. It combines data from several administrative sources and contains building-level information on the location of dengue clusters, the number of dengue cases in each cluster and average monthly residential electricity

[3] Households know if they are exposed to localised dengue risk as information is provided by prominently displayed alert banners in dengue clusters. These banners also provide information on government classifications of localised risk intensity via colour codes. Clusters are defined as *"a locality with active transmission where intervention is targeted. It is formed when two or more cases have onset within 14 days and are located within 150m of each other (based on residential and workplace addresses as well as movement history)."* (National Environment Agency, 2016b).

consumption at the building (henceforth referred to as block) and flat-type[4] level. The data cover 2013 and 2014, which span the worst dengue epidemic in Singapore's history.

Causal identification comes from spatial and temporal variation in proximity to dengue clusters. This allows us to use fixed effects to account for unobserved block-level heterogeneity and region-varying time fixed effects, which capture most factors that dengue epidemiologists deem likely to affect the location and duration of a cluster (Pang & Leo, 2015). The remaining variation is largely driven by movement of infected mosquitoes and humans, which is likely to be exogenous with respect to electricity consumption.

We find that households increase their electricity consumption when they are near dengue clusters (i.e. when they are exposed to dengue risk), and the response varies with socio-economic status[5]. Being near a cluster leads to a rise in electricity consumption for 4-room and 5-room/bigger flats (by 1.7% and 1.1% respectively)[6], but not 3-room flats. This rise persists in the months after exposure to dengue risk. Households' responses also increase discontinuously when dengue risk classification on alert banners is upgraded from yellow (≤9 cases) to red (≥10 cases). This discontinuous increase supports our hypothesis that an observed increase in electricity consumption is driven by the protective behaviour of closing doors and windows, as other factors related to electricity consumption are unlikely to vary with dengue risk intensity in this discontinuous manner. Our results are robust to the inclusion of proximity to construction sites[7]. In addition, we consider the effect of other identification issues that we are unable to address in Section 2.6 (e.g. measurement error), and conclude that these issues lead to an underestimation of

[4] Flat-types refer to different types of apartments which can be located within the same residential block. These include (in order of increasing size) 3-, 4-, 5-room/bigger flats.
[5] We proxy socio-economic status by flat-type, which is highly correlated with income in Singapore (Department of Statistics, 2016a).
[6] Percentages are computed off the mean monthly electricity consumption statistics in **Table 1**.
[7] Construction sites may breed *Aedes* mosquitoes. This may lead to endogeneity as electricity consumption increases with construction site proximity (Agarwal et al., 2016).

the true effect. After accounting for the social cost of carbon, we find that increased electricity consumption due to dengue outbreaks costs around $11.9 to 16.3 million per annum (in 2015 Singapore dollars), constituting 7% – 12% of the overall costs of dengue in Singapore[8].

Our results are most related to the literature on residential electricity consumption. Existing papers using experimental or quasi-experimental methods to study the determinants of residential electricity consumption have found that electricity demand can be affected by temperature (Aroonruengsawat & Auffhammer, 2011; Deschênes & Greenstone, 2011; Auffhammer & Mansur, 2014), environmental disamenities (Agarwal et al., 2016), price shocks (e.g. Reiss & White, 2008), or policy-related determinants such as building codes (Aroonruengsawat et al., 2012; Jacobsen & Kotchen, 2013), salience of information (Carroll et al., 2014; Sexton, 2015), social nudges (Allcott, 2011; Ayres et al., 2013; Allcott & Rogers, 2014; Agarwal et al., 2017), or the desire to signal environmental friendliness (Delmas & Lessem, 2014).

Our paper contributes to this literature by providing evidence that a previously unstudied environmental disamenity – dengue fever outbreaks – can lead to a persistent rise in electricity consumption[9]. The magnitude of the rise we observe (1.1 – 1.7%) is comparable to falls induced by some interventions that try to reduce electricity consumption (e.g., the social comparisons studied in, say, Allcott, 2011; Ayres et al., 2013), and about 10 – 15% of the rise in residential

[8] Annual overall cost of dengue, inclusive of direct and indirect cost of illness and control cost, is estimated at S$131 – 176 mil (in 2015 Singapore dollars), over the period of 2000 to 2009 (Carrasco et al., 2011).

[9] In the interest of readability, we relegate our paper's contributions to other literatures in this footnote. Our results also contribute to a nascent literature on protective behaviour in response to health risks. Papers in this literature have studied behavioural responses to lowered air or water quality (Neidell, 2009; Zivin & Neidell, 2009; Zivin et al., 2011; Janke, 2014; Zhang & Mu, 2016), mercury exposure (Shimshack & Ward, 2010), or folic acid advisories (Herrera‐Araujo, 2016). Papers on responses to the risk of contracting vector-borne diseases are rare in this literature; the only other paper we are aware of is by Dammert et al. (2014), who investigate the effect of using mobile phones to encourage preventive measures against dengue. In addition, our results also have implications for the literature estimating the economic cost of dengue. Most papers in this literature do not account for costs incurred due to protective behaviour (e.g. Clark et al., 2005; Suaya et al., 2009; Carrasco et al., 2011; Thalagala et al., 2016). Our paper provides an innovative way of accounting for some costs arising from protective behavior, and our results imply that such costs can form a substantial fraction of the total costs of dengue, supporting Rodriguez et al. (2016)'s argument that such costs should be accounted for as well.

electricity consumption that temperature increase in the "business-as-usual" climatic scenario may lead to by 2099 (Deschênes & Greenstone, 2011).

These results suggest that tackling environmental disamenities can provide benefits (in this case, reduced electricity consumption) beyond those that are typically discussed (such as improved health). In addition, the increased electricity consumption in response to dengue outbreaks we see is likely to become more common worldwide, due to the expected increase in the geographical range of dengue fever (Hales et al., 2002; Smith et al., 2014), as well as the likely increase in prevalence of air-conditioners in currently dengue-endemic countries as they grow richer (Isaac & Van Vuuren, 2009; Akpinar-Ferrand & Singh, 2010). Our results thus also suggest that it may become increasingly important to account for responses to environmental disamenities such as vector-borne diseases (of which dengue fever is one) when simulating the effects of climate change on electricity consumption.

2.2 Background

Dengue is a viral disease transmitted by the *Aedes* mosquito. Mild symptoms include high fever, intense headaches, muscle/joint pains, and vomiting; severe cases can lead to death due to plasma leaking, respiratory distress, severe bleeding or organ impairment (World Health Organisation, 2016).

In 2013, Singapore experienced her worst dengue epidemic to date, which claimed 8 fatalities (Ministry of Health, 2014). A total of 22,170 cases were reported, 380% more than the number of cases reported in 2012. The number of dengue cases remained high in 2014, before falling substantially in 2015 (**Figure 2.1**).

A spike in dengue risk is likely to lead to protective behaviour only if households understand the consequences of contracting dengue, and have information about dengue exposure risk. Both conditions are likely to hold in Singapore.

First, most Singapore residents understand the consequences of dengue. Dengue updates and deaths are reported in Singapore's mainstream news, which makes it clear that people of all age groups are susceptible. E.g., the first dengue death in 2013 was that of a 20-year-old male (e.g. Tham (2013); Khalik (2013); Khalik (2014)). Second, information on dengue cluster location (i.e. exposure risk), and risk intensity, is easily accessible. Singapore's National Environment Agency (NEA) displays large banners in active dengue clusters. These banners are displayed in prominent locations within the clusters, and also provide information on risk classifications (colour-coded as red(highest), yellow or green(lowest)[10]) (National Environment Agency, 2016b). NEA also publishes this information online.

NEA advises households to practise protective behaviors (e.g. spraying insecticide) (National Environment Agency, 2016a). In addition, there are media reports that individuals close windows and doors to prevent mosquitoes from flying in, especially since the use of insect screens is uncommon in Singapore (e.g. Lee and Kumar (2016)). In hot, humid Singapore, closing all windows and doors often means that households are likely to "purchase comfort" (Quigley & Rubinfeld, 1989) by consuming more electricity to cool their homes using fans or air-conditioners[11]. This practice so common that it has even appeared in articles and infographics in major Singaporean newspapers (e.g. Lee & Kumar, 2016; Singapore Power, 2016).

---

[10] A red-coded banner indicates that 10 or more cases have been reported in the cluster; a yellow-coded banner indicates fewer than 10 cases; a green-coded banner indicates that the cluster has no new cases but will be monitored for the next 21 days.

[11] 52% and 94% of Singapore resident households in the lowest and highest income quintile own air-conditioners respectively (Department of Statistics, 2014).

2.3 Data and Variables

2.3.1 Data Sources and Construction of the Combined Dataset

We outline how we construct our panel dataset here. Appendix A provides more detail. Our dataset is compiled using data with nationwide coverage from three administrative sources: the Energy Market Authority (EMA), National Environment Agency (NEA), and Housing and Development Board (HDB). We have block-level information on dengue cluster locations, number of dengue cases per cluster, and average monthly electricity consumption at the block and flat-type[12] level for households living in public housing (approximately 80% of resident households live in public housing in Singapore). Postal codes, which uniquely identify blocks, are used to derive postal districts which we use as regional indicators. We also have the number of flats of each flat-type within a public housing block, which is used to weight the averaged block-level electricity data.

We create three balanced panel datasets (one each for 3-room, 4-room and 5-room/bigger flats[13]), and compute the distance between a block and the nearest dengue cluster in each month (this will be used to construct our dengue risk proxy). Each dataset covers monthly observations for the 23-month period from May 2013 to March 2015.

Lastly, we drop newly-built blocks. As households start moving into new blocks, average block-level electricity consumption increases rapidly. This phenomenon cannot be captured by region-specific time fixed effects and could bias our estimates if these new blocks are near dengue clusters. (Section 2.5 provides evidence that this is indeed an issue.)

[12] Flat-types refer to different types of apartments which can be located within the same residential block. These include (in order of increasing size) 3-, 4-, 5-room/bigger flats.

[13] As of 2015, about 74% of Singapore resident households stay in these flat-types (Department of Statistics, 2016b). We will not be analysing data on 1-/2-room flats as they represent a very small percentage of households.

2.3.2 Construction of Key Variables

A key independent variable is our proxy for households' perceived risk of dengue exposure. We construct this dummy variable using the distance between a public housing block and its nearest dengue cluster. The variable "*near cluster*" takes value one if the distance is within a certain cut-off value. We elaborate on the choice of cut-off values in Section 2.5.2.

Our second key independent variable tests if any increase in electricity consumption induced by exposure to dengue risk persists, even when households are no longer near a dengue cluster. We construct a dummy variable "*persist*" that takes value one if a block was near a cluster before, but is not near a cluster in that particular month.

2.3.3 Summary Statistics

Table 2.1 presents summary statistics for our three flat-type-specific datasets. The mean monthly electricity consumption increases noticeably with the size of flats – from 279.80kWh for 3-room flats to 467.97kWh for 5-room flats. The probabilities of being within 500m or 1km of a dengue cluster are almost the same across the different flat-types, ranging from around 0.40 to 0.67 for 500m and 1km respectively. This is not surprising as the geographical spread of public housing blocks is similar across flat-types. Lastly, **Figure 2.2** provides some visual evidence that the location of dengue clusters varied both spatially and temporally over the period we study.

2.4 Identification Strategy and Empirical Model

We use fixed effects models to estimate the effect of exposure to dengue risk (i.e. being near a dengue cluster where alert banners are prominently displayed) on households' electricity consumption. We perform our analyses separately for 3-room, 4-room, and 5-room/bigger flats, to allow estimates to vary by flat-type. We do so as flat-type is strongly correlated with households' socio-economic status[14], which could influence how they perceive and react to dengue risk. E.g., richer households could be more averse to dengue risk; they can also afford to turn on air conditioners more often.

Our models take the general form:

$$Y_{it} = \beta D_{it} + \alpha_i + \gamma_{kt} + \varepsilon_{it} \qquad (1)$$

where Y_{it} denotes mean monthly electricity consumption for a block *i* in year-month *t* while D_{it} is a vector of treatment-related variables (e.g. whether a block is near a dengue cluster). Block fixed effects α_i control for time-invariant unobserved block-level heterogeneity such as socio-economic status, risk aversion, extent of herd immunity, and density of buildings[15]. Postal-district-year-month fixed effects γ_{kt} control for region-varying time trends. In particular, they control for micro-climates (time-varying, regional differences in temperature and rainfall), and also absorb time-varying shocks common to all blocks, such as electricity prices and media coverage of dengue. ε_{it} is the error term. Standard errors are clustered at the block level.

Our rich set of fixed effects control for most factors that dengue epidemiologists deem likely to affect the duration and location of dengue clusters (Pang & Leo, 2015). The remaining variation in any block's proximity to a dengue cluster is likely to be driven by movement of infected humans and mosquitoes, and is likely to be exogenous with respect to electricity consumption. Our models are thus likely to identify causal effects.

[14] See Table 1 on page 10 of the Key Household Income Trends 2015 report (Department of Statistics, 2016a), for information on how mean per-capita income per household member varies according to flat type.

[15] These features can be considered time-invariant given the short duration of our study period and the fact that about 90% of Singaporeans own their public housing flat (Housing Development Board, 2014) and are less likely to move.

2.5 Main Empirical Specifications and Results

We test (i) if proximity to dengue clusters[16] induces increased electricity consumption; (ii) how persistent any increase we observe is; and (iii) if the extent of the effect is influenced by the risk classification that NEA gives each cluster.

2.5.1 Effect of Being Near a Dengue Cluster

Our primary specification is

$$Y_{it} = \beta_1(near\ cluster)_{it} + \beta_2(persist)_{it} + \alpha_i + \gamma_{kt} + \varepsilon_{it} \qquad (2)$$

where $(near\ cluster)_{it}$ is a dummy variable that takes value one if a block is near a dengue cluster in year-month *t*, and $persist_{it}$ is a dummy variable that takes value one if a block has been near a cluster before year-month *t*, but is not near a cluster in year-month *t*.

β_1 captures the effect of being near a cluster in year-month *t*, i.e. the treatment effect. β_2 captures persistence after treatment ended, i.e. the persistence effect. The variable $persist_{it}$ is important as the literature often finds that changes in electricity consumption are persistent (see e.g. Allcott & Rogers, 2014; Agarwal et al., 2016). Excluding $persist_{it}$ would lead to a downward bias in β_1, as blocks exhibiting increased electricity consumption after treatment would be wrongly tagged as "non-treated".

2.5.2 Cut-off Distances

$(Near\ cluster)_{it}$ can be seen as a proxy for households' perception of dengue risk. Ideally, the cut-off distance should be defined so that households beyond the cut-off distance perceive dengue risk to be slight or do not respond to dengue risk. Households within the cut-off will be taken to be "treated", while households outside the cut-off will be taken as "non-treated". Having said this, we are aware that in reality, both groups differ in terms of the intensive, rather than

[16] Individuals know if they are near a cluster, as dengue alert banners are prominently displayed at dengue clusters.

extensive margin of treatment. Hence, our coefficient estimates capture relative treatment and persistence effects.

In the absence of information on households' risk preferences, it is not clear what the cut-off distance should be. If cut-offs are too small, households that respond to dengue risk would be wrongly tagged as "non-treated"; if the cut-offs are too large, households that do not respond to dengue risk would be wrongly tagged as "treated". Both cases bias our coefficients downwards. We use this hypothesised relationship between the magnitude of the effects and the cut-off distance to determine appropriate cut-offs.

Estimating **Eq (2)** using cut-offs in increasing intervals of 100m (from 100m to 1,500m), should result in coefficient estimates that rise, then fall, as cut-off distances increase. If such a pattern exists, the distance that corresponds to the peak of the inverted-U curve of coefficients would be a reasonable cut-off. We estimate **Eq (2)** separately for each flat-type, as socio-economic status could affect the perception and reaction to dengue risk differently, leading to different choices of cut-off distances and coefficient estimates.

Figure 2.3 presents these results graphically. (The full tables of coefficients for these results are in Appendix B.) Our coefficient estimates are consistent with our hypothesised inverted U-shaped curve, and suggest that 3-room households may not increase electricity consumption in response to dengue risk. Residents of these flats typically have lower incomes and may not use more air-conditioning even if they close their doors/windows. Given the lack of a response in 3-room flats, we exclude them from our subsequent regressions, which study how the effects of being near a cluster vary over time and by NEA's cluster risk classification. Last, applying the criterion outlined above to these results, we define the cut-offs as 1100m for 4-room flats, and 1400m for 5-room/bigger flats.

2.5.3 Effect of Proximity to Dengue Cluster on Electricity Consumption

Table 2.2 reports results from estimating **Eq (2)** using the cut-offs defined earlier. First, we observe that excluding newly-built blocks from our sample is important in avoiding upward bias. Estimates in columns 1 and 3, which include new blocks, are larger than those in columns 2 and 4, which exclude new blocks. This is because average electricity consumption (at the block-level) in new blocks starts low, and increases rapidly as households move in. Including new blocks in our sample would confound our estimates when these new blocks are near dengue clusters. The sample used for our regressions from this point on will thus exclude these newly-built blocks.

Columns 2 and 4 report coefficient estimates from **Eq (2)**, using our preferred sample (without new blocks), for 4-room and 5-room/bigger flats respectively. Proximity to a dengue cluster increases electricity consumption by 6.64kWh (or 1.7% of the mean monthly electricity consumption) for 4-room flats, and 5.28kWh (or 1.1% of the mean) for 5-room/bigger flats, likely due to protective behaviour (closing doors and windows and turning on the fan or air-conditioner) households engage in when they face increased dengue risk.

This rise is persistent: electricity consumption in the months after exposure to dengue risk stay elevated (by 8.89 kWh and 6.09 kWh for 4-room and 5-room/bigger flats respectively)[17]. The persistence of changes in electricity consumption in response to exogenous shocks is often interpreted as evidence of habit formation (see e.g. Allcott and Rogers (2014)). In our case, the presence of persistence effects may also be due to increased risk aversion after households were near a cluster, though we are unable to separate the relative contribution of each effect with the data we possess.

[17] The coefficients for our persistence variable may be larger than those of our treatment variables as (i) our treatment estimates are attenuated more than our persistence estimates (as discussed in Section 6); and (ii) households that first started using air-conditioning in response to increased dengue risk may become habituated to using air-conditioning and start using air-conditioning more often even after the threat of dengue has diminished.

2.5.4 Dynamic Response to Being Near a Dengue Cluster

Next, we examine the dynamics of the treatment and persistence effects with using **Eq (3)**:

$$\begin{aligned} Y_{it} = \quad & \beta_1(near\ cluster)_{it} \\ & + \beta_{11}(near\ cluster \times\ no.of\ months\ after\ first\ treated)_{it} \\ & +\beta_2(persist)_{it} + \beta_{21}(persist \times\ no.of\ months\ post\ cluster)_{it} \\ & +\alpha_i + \gamma_{kt} + \varepsilon_{it} \end{aligned} \tag{3}$$

where (*no. of months after first treated*) is a variable representing the number of months a block spends near a cluster. (*no. of months post cluster*) is a representing the number of months since "treatment" ended.

β_1 and β_2 estimate the overall treatment and persistence effects respectively. β_{11} and β_{21} test if the treatment and persistence effects vary over time.

Table 2.3 shows that β_{11} is not statistically significant for both flat-types, suggesting that treatment effects do not vary over time. β_{21} is statistically significant for 5-room/bigger flats (the persistence effect increases at a small amount of 0.75kWh every month) but not 4-room flats, suggesting that persistence effects do not decay over the period studied.

2.5.5 Effect of Risk Classification Within a Cluster

Lastly, we examine how the treatment and persistence effects vary with NEA's risk classification of dengue risk within a cluster. To estimate the effect of being in a cluster with a higher risk classification, we interact (*red banner*), a dummy variable for clusters that NEA deems high risk (i.e. those with at least 10 cases), with the *(near cluster)*$_{it}$ and *(persist)*$_{it}$ variables. In addition, we interact the *(near cluster)*$_{it}$ and *(persist)*$_{it}$ variables with different functional forms (linear, quadratic) of the number of cases (*f(cases)*) within each cluster, to verify that households are indeed responding to an upgrade in risk classification, rather than an increased number of cases within a cluster. In these specifications, identification of the effect of increased dengue risk on electricity consumption can be seen as coming from the discontinuous change in risk classifications (from yellow to red) when the number of dengue cases rises from 9 to 10.

Eq (4) shows our general specification:

$$
\begin{aligned}
Y_{it} = \quad & \beta_1(near\ cluster)_{it} + \beta_{11}(near\ cluster \times red\ banner)_{it} \\
& +\beta_{12}(near\ cluster \times\ f(cases))_{it} \\
& +\beta_2(persist)_{it} + \beta_{21}(persist \times\ red\ banner)_{it} \\
& +\beta_{22}(persist \times\ f(cases))_{it} \\
& +\alpha_i + \gamma_{kt} + \varepsilon_{it}
\end{aligned}
\tag{4}
$$

β_1 and β_2 estimate the overall treatment and persistence effects, which we call baseline effects. β_{11} and β_{21} estimate the effects of being near clusters with a higher NEA risk classification (i.e. a red banner) beyond the baseline effects. In specifications which include the linear or quadratic polynomial of the number of dengue cases (*f(cases)*), β_{11} and β_{21} continue to capture the additional effect of being near a high risk cluster, but are now identified off the discontinuity in NEA's risk classification at the 10-case threshold.

Table 2.4 shows that households respond more when the risk classification is upgraded. The estimated values of β_{11} and β_{21} in columns (1) and (4) are positive and statistically significant, implying that being in a high risk, red-coded cluster leads to higher treatment and persistence effects in both 4-room and 5-room/bigger flats.

Columns (2), (3), (5), and (6) of **Table 2.4** report results from specifications that include interactions between *f(cases)* and $(near\ cluster)_{it}$ or $(persist)_{it}$. The coefficients β_{11} and β_{21} remain statistically significant, and are similar to / slightly higher than estimates from columns (1) and (4). Furthermore, the coefficients for the interaction terms between *f(cases)* and $(near\ cluster)_{it}$ or $(persist)_{it}$ are small in magnitude and generally statistically insignificant. These results suggest that electricity consumption increases discontinuously when NEA upgrades a cluster's dengue risk classification from yellow to red, and that the increased electricity consumption in columns (1) and (4) arises from a worsening of the risk classification, rather than an increase in the number of cases within the cluster. In addition, this discontinuous increase provides further support for our hypothesis that the observed increase in electricity consumption is driven by the protective behaviour of closing doors and windows, as other factors related to electricity consumption are unlikely to vary with NEA's dengue risk classification in this discontinuous manner.

2.6 Threats to Identification and Robustness Checks

In Section 2.4, we argue that our rich set of fixed effects address most endogeneity problems. However, some potential threats to identification remain. In this section, we find that our results either remain robust in most checks after we account for these threats, or can be seen as a lower bound of the true effect in cases where we are unable to account for the threats.

First, households' proximity to construction sites (which may contain mosquito breeding grounds and increase the likelihood of nearby areas becoming dengue clusters) could bias our results upwards, as Agarwal et al. (2016) find that households living near construction sites increase their electricity consumption. We control for proximity to public and private housing construction sites[18] in columns 1, 2, 5 and 6 of **Table 2.5**. Our results remain unchanged.

Second, we may not have sufficiently accounted for spatial autocorrelation, which could occur if e.g. electricity consumption habits come up in conversation between residents of neighbouring blocks. We cluster our errors at the level of residential committee (RC) zones[19] in

---

[18] We use public housing construction data retrieved from http://www.teoalida.com/singapore/btolist/ on 19 September 2016 and private housing construction data retrieved from REALIS, an online database maintained by Singapore's Urban Redevelopment Authority.

[19] Residential committees (RC) are organisations set up by the People's Association, a government statutory board, to promote cohesiveness among residents within their respective RC zones in public housing estates. We use the closest

columns 3 and 7 of **Table 2.5**. Our results for 4-room flats remain robust but results for 5-room/bigger flats are weaker.

Third, we use the log of electricity consumption as the dependent variable in columns 4 and 8 of **Table 2.5**. Again, our results for 4-room flats remain robust but results for 5-room/bigger flats are weaker.

Fourth, there is a possibility of reverse causality. Households who close their doors and windows more often may reduce the spread of dengue. Electricity consumption of these households would be negatively correlated with treatment probability, and attenuate our estimates.

Fifth, residential sorting in response to dengue risk could bias our results. Given the extent of the dengue outbreak, the speed at which it occurred, and the relatively short timeframe of our sample, residential sorting in response to dengue risk is unlikely. This is especially so given that most resident households own their home in Singapore. Furthermore, sorting is likely to result in a downward bias, as households that are more concerned about dengue risk are more likely to move away from a dengue prone area and use more electricity in response to dengue risk.

Finally, measurement error could bias our results. However, all known sources of measurement error for our dataset introduce downwards bias, which means that our estimates can be seen as a lower bound of the true effect.

First, our dengue cluster data is updated two to five times a month, while electricity consumption data is reported monthly. To merge these two datasets, we tag a block as "treated" if the block is near a cluster at least once in a month. This means that blocks we identify as "treated" could actually be "untreated" for part of the month, biasing our results downwards.

RC office to each block as a proxy for actual residential committee zone. Data for RC offices is retrieved from https://data.gov.sg/ on 27 September 2016.

Second, electricity consumption is reported by the month of billing. As the date of billing varies by when households registered with the utilities provider, electricity consumption for the month of billing, say November 2013, could reflect electricity consumption for part of October 2013 as well. If a block is tagged as treated in November 2013, the electricity consumption associated with it would partly reflect a household's electricity consumption when the household was not near a dengue cluster. Since our models already account for persistence, such measurement error biases the treatment coefficients downwards.

Last, the left-truncated nature of our dataset means that some blocks could be tagged as "untreated" in May 2013, even if they were near a dengue cluster in, say April 2013. These blocks could still be using more electricity due to the persistent effects of exposure to a dengue cluster. If so, this measurement error would introduce downwards bias into our results, as these wrongly tagged blocks contribute to the baseline off which treatment-related effects are estimated. However, the extent of this downward bias is likely to be mitigated by the fact that May 2013 was near the peak of the 2013 dengue outbreak, and there were relatively few dengue cases in 2012.

2.7 Costs of Protective Behaviour

To round off our analyses, we carry out back-of-the-envelope calculations of the costs of protective behaviour in response to dengue outbreaks in **Table 2.6a** and **Table 2.6b**. Dollar amounts are in 2015 Singapore dollars.

To estimate the absolute increase in electricity consumption due to proximity to a dengue cluster from May 2013 to March 2015, we apply the results in columns 2 and 4 of **Table 2.2** to affected block-year-month observations, accounting for the number of flats per block.

We convert the estimated increase in electricity consumption into economic and carbon costs. Applying the appropriate electrical tariffs and considering that 96% of Singapore's power is generated from natural gas in 2015 (Energy Market Authority, 2015), we find that protective behaviour costs each 4-room and 5-room/bigger household an average of S$41 and S$31 respectively, and resulted in an average release of approximately 89kg and 67kg of carbon dioxide respectively, over the 23-month period from May 2013 to March 2015. Overall, this protective behaviour cost all households S$22.8 to 31.2 million (including social cost of carbon), or S$11.9 to 16.3 million per annum, about 7%–12% of the overall costs of dengue in Singapore[20].

2.8 Conclusion

In this paper, we find that a major and growing environmental disamenity – dengue fever outbreaks – can induce protective behaviour and lead to economically significant increases in electricity consumption. Households near dengue clusters consume more electricity (likely due to closing doors/windows and turning on the air-conditioner to keep mosquitoes out), with the greatest increases being observed for households near clusters with the highest NEA risk classification. This increase is unlikely to be driven by unobserved heterogeneity, as we include block-specific and region-varying time fixed effects, and observe a discontinuous rise in electricity consumption when the risk classification on dengue banners is upgraded.

The magnitude of the increase in electricity consumption we observe (1.1 – 1.7%) is not negligible – it is comparable to falls induced by some interventions that try to reduce electricity consumption (e.g., the social comparisons studied in, say, Allcott, 2011; Ayres et al., 2013), and about 10 – 15% of the rise in residential electricity consumption that temperature increase in the "business-as-usual" climatic scenario may lead to by 2099 (Deschênes & Greenstone, 2011). After including the social cost of carbon, we find that this increase in electricity consumption cost around

[20] The annual overall cost of dengue from 2000 - 2009, inclusive of direct and indirect cost of illness and control cost, but excluding costs due to protective behaviour, is estimated at S$131 – 176 mil (in 2015 Singapore dollars) (Carrasco et al., 2011).

$11.9 to 16.3 million per annum (in 2015 Singapore dollars), constituting 7% – 12% of the overall costs of dengue in Singapore[21].

Our results suggest that tackling environmental disamenities can provide benefits (in this case, reduced electricity consumption) beyond those that are typically discussed (an example of which is improved health). In addition, the increased electricity consumption in response to dengue outbreaks we see is likely to become more common worldwide, due to the expected increase in the geographical range of dengue fever (Hales et al., 2002; Smith et al., 2014), as well as the likely increase in prevalence of air-conditioners in currently dengue-endemic countries as they grow richer (Isaac & Van Vuuren, 2009; Akpinar-Ferrand & Singh, 2010). Our results thus also suggest that it may become increasingly important to account for responses to environmental disamenities such as vector-borne diseases (of which dengue fever is one) when simulating the effects of climate change on electricity consumption.

[21] Annual overall cost of dengue, inclusive of direct and indirect cost of illness and control cost, is estimated at S$131 – 176 mil (in 2015 Singapore dollars), over the period of 2000 to 2009 (Carrasco et al., 2011).

References

Agarwal, S., Rengarajan, S., Sing, T. F., & Yang, Y. (2017). Nudges from School Children and Electricity Conservation: Evidence from the "Project Carbon Zero" Campaign in Singapore. *Energy Economics, 61*, 29-41.

Agarwal, S., Satyanarain, R., Sing, T.-F., & Vollmer, D. (2016). Effects of Construction Activities on Residential Electricity Consumption: Evidence from Singapore's Public Housing Estates. *Energy Economics, 55*, 101 - 111.

Akpinar-Ferrand, E., & Singh, A. (2010). Modeling Increased Demand of Energy for Air Conditioners and Consequent Co 2 Emissions to Minimize Health Risks Due to Climate Change in India. *Environmental science & policy, 13*(8), 702-712.

Alberini, A., & Towe, C. (2015). Information V. Energy Efficiency Incentives: Evidence from Residential Electricity Consumption in Maryland. *Energy Economics, 52*, S30-S40.

Allcott, H. (2011). Social Norms and Energy Conservation. *Journal of Public Economics, 95*, 1082-1095.

Allcott, H., & Rogers, T. (2014). The Short-Run and Long-Run Effects of Behavioural Interventions: Experimental Evidence from Energy Conservation. *American Economic Review, 104*(10), 3003-3037.

Aroonruengsawat, A., & Auffhammer, M. (2011). Impacts of Climate Change on Residential Electricity Consumption: Evidence from Billing Data *The Economics of Climate Change: Adaptations Past and Present* (pp. 311-342): University of Chicago Press.

Aroonruengsawat, A., Auffhammer, M., & Sanstad, A. H. (2012). The Impact of State Level Building Codes on Residential Electricity Consumption. *The Energy Journal, 33*(1), 31.

Auffhammer, M., & Mansur, E. T. (2014). Measuring Climatic Impacts on Energy Consumption: A Review of the Empirical Literature. *Energy Economics, 46*, 522-530.

Ayres, I., Raseman, S., & Shih, A. (2013). Evidence from Two Large Field Experiments That Peer Comparison Feedback Can Reduce Residential Energy Usage. *Journal of Law, Economics, and Organization, 29*(5), 992 - 1022.

Carrasco, L. R., Lee, L. K., Lee, V. J., Ooi, E. E., Shepard, D. S., Thein, T. L., . . . Leo, Y. S. (2011). Economic Impact of Dengue Illness and the Cost-Effectiveness of Future Vaccination Programs in Singapore *PLOS Neglected Tropical Diseases, 5*(12), e1426.

Carroll, J., Lyons, S., & Denny, E. (2014). Reducing Household Electricity Demand through Smart Metering: The Role of Improved Information About Energy Saving. *Energy Economics, 45*, 234-243.

Clark, D. V., Mammen, M. P. J., Nisalak, A., Puthimethee, V., & Endy, T. P. (2005). Economic Impact of Dengue Fever / Dengue Hemorrhagic Fever in Thailand at the Family and Population Levels. *American Journal of Tropical Medicine and Hygiene, 72*(6), 786 - 791.

Dammert, A. C., Galdo, J. C., & Galdo, V. (2014). Preventing Dengue through Mobile Phones: Evidence from a Fieldexperiment in Peru. *Journal of Health Economics, 35*, 147-161.

Delmas, M. A., & Lessem, N. (2014). Saving Power to Conserve Your Reputation? The Effectiveness of Private Versus Public Information. *Journal of Environmental Economics and Management, 67*(3), 353-370.

Department of Statistics. (2014). *Report on the Household Expenditure Survey 2012 / 2013*. Retrieved from Singapore:

Department of Statistics. (2016a). *Key Household Income Trends 2015* Retrieved from Singapore:

Department of Statistics. (2016b). Latest Data. Retrieved from http://www.singstat.gov.sg/statistics/latest-data#22

Deschênes, O., & Greenstone, M. (2011). Climate Change, Mortality, and Adaptation: Evidence from Annual Fluctuations in Weather in the Us. *American Economic Journal: Applied Economics, 3*(4), 152-185.

Energy Market Authority. (2015). *Singapore Energy Statistics 2015*. Retrieved from Singapore: https://www.ema.gov.sg/cmsmedia/Publications_and_Statistics/Publications/SES2015_Final_website_2mb.pdf

Hales, S., De Wet, N., Maindonald, J., & Woodward, A. (2002). Potential Effect of Population and Climate Changes on Global Distribution of Dengue Fever: An Empirical Model. *The Lancet, 360*(9336), 830-834.

Herrera‐Araujo, D. (2016). Folic Acid Advisories: A Public Health Challenge? *Health economics, 25*(9), 1104-1122.

Housing Development Board. (2014). *Hdb Annual Report Fy 2013/2014*. Retrieved from

Ihara, T., Genchi, Y., Sato, T., Yamaguchi, K., & Endo, Y. (2008). City-Block-Scale Sensitivity of Electricity Consumption to Air Temperature and Air Humidity in Business Districts of Tokyo, Japan. *Energy, 33*(11), 1634-1645.

International Energy Agency. (2017). *World Energy Outlook 2017*. Retrieved from

Isaac, M., & Van Vuuren, D. P. (2009). Modeling Global Residential Sector Energy Demand for Heating and Air Conditioning in the Context of Climate Change. *Energy Policy, 37*(2), 507-521.

Jacobsen, G. D., & Kotchen, M. J. (2013). Are Building Codes Effective at Saving Energy? Evidence from Residential Billing Data in Florida. *Review of Economics and Statistics, 95*(1), 34-49.

Janke, K. (2014). Air Pollution, Avoidance Behaviour and Children's Respiratory Health: Evidence from England. *Journal of Health Economics, 38*, 23 - 42.

Khalik, S. (2013). 66-Year-Old Becomes 4th Person to Die from Dengue in Current Epidemic. *The Straits Times*. Retrieved from http://www.straitstimes.com/singapore/66-year-old-becomes-4th-person-to-die-from-dengue-in-current-epidemic

Khalik, S. (2014). Woman in Her Late 50s Becomes First Dengue Death This Year. *The Straits Times*. Retrieved from http://www.straitstimes.com/singapore/woman-in-her-late-50s-becomes-first-dengue-death-this-year

Lee, C.-C., & Chiu, Y.-B. (2011). Electricity Demand Elasticities and Temperature: Evidence from Panel Smooth Transition Regression with Instrumental Variable Approach. *Energy Economics, 33*(5), 896-902.

Lee, S., & Kumar, C. (2016). Residents Take Action in Worst Dengue Clusters. *The Straits Times*. Retrieved from http://www.straitstimes.com/singapore/health/residents-take-action-in-worst-dengue-clusters

List, J. A., Metcalfe, R. D., Price, M. K., & Rundhammer, F. (2017). *Harnessing Policy Complementarities to Conserve Energy: Evidence from a Natural Field Experiment*. Retrieved from

Ministry of Health. (2014). *Vector-Borne Diseases*. Retrieved from Singapore: https://www.moh.gov.sg/content/dam/moh_web/Publications/Reports/2014/Communicable%20Diseases%20Surveillance%20in%20Singapore%202013/Vector%20Borne%20-Zoonotic%20Diseases.pdf

Ministry of Health. (2016). Weekly Infectious Diseases Bulletin Retrieved from https://www.moh.gov.sg/content/moh_web/home/statistics/infectiousDiseasesStatistics/weekly_infectiousdiseasesbulletin.html?year=2012

Moral-Carcedo, J., & Vicens-Otero, J. (2005). Modelling the Non-Linear Response of Spanish Electricity Demand to Temperature Variations. *Energy Economics, 27*(3), 477-494.

National Environment Agency. (2016a). Colour Coded Alert System. Retrieved from http://www.dengue.gov.sg/subject.asp?id=76

National Environment Agency. (2016b). Location of Active Clusters. Retrieved from http://www.dengue.gov.sg/subject.asp?id=74

Neidell, M. (2009). Information, Avoidance Behaviour, and Health. *The Journal of Human Resources, 44*(2), 450 - 478.

Pang, J., & Leo, Y.-S. (2015). Clustering, Climate and Dengue Transmission. *Expert Review of Anti-infective Therapy*, 1 - 10.

Quigley, J. M., & Rubinfeld, D. L. (1989). Unobservables in Consumer Choice: Residential Energy and the Demand for Comfort. *The Review of Economics and Statistics*, 416-425.

Reiss, P. C., & White, M. W. (2008). What Changes Energy Consumption? Prices and Public Pressures. *RAND Journal of Economics, 39*(3), 636 - 663.

Rodriguez, R. C., Carrasquilla, G., Porras, A., Galera-Gelvez, K., Yescas, J. G. L., & Rueda-Gallardo, J. A. (2016). The Burden of Dengue and the Financial Cost to Colombia, 2010 - 2012. *American Journal of Tropical Medicine and Hygiene, 94*(5), 1065 - 1072.

Sexton, S. (2015). Automatic Bill Payment and Salience Effects: Evidence from Electricity Consumption. *The Review of Economics and Statistics, 97*(2), 229 - 241.

Shimshack, J. P., & Ward, M. B. (2010). Mercury Advisories and Household Health Trade-Offs. *Journal of Health Economics, 29*, 674 - 685.

Singapore Power. (2016). The Electricity Times. *Today*, p. 11.

Smith, K. R., Woodward, A., Campbell-Lendrum, D., Chadee, D. D., Honda, Y., Liu, Q., . . . Sauerborn, R. (2014). Human Health: Impacts, Adaptation, and Co-Benefits. *Climate change*, 709-754.

Suaya, J. A., Shepard, D. S., Siqueira, J. B., Martelli, C. T., Lum, L. C. S., Tan, L. H., . . . Halstead, S. B. (2009). Cost of Dengue Cases in Eight Countries in the Americas and Asia: A Prospective Study. *American Journal of Tropical Medicine and Hygiene, 80*(5), 846 - 855.

Thalagala, N., Tissera, H., Palihawadana, P., Amarasinghe, A., Ambagahawita, A., Wilder-Smith, A., . . . Tozan, Y. (2016). Costs of Dengue Control Activities and Hospitalisations in the Public Health Sector During an Epidemic Year in Urban Sri Lanka. *PLOS Neglected Tropical Diseases, 10*(2), e0004466.

Tham, T.-C. (2013). 20-Year-Old Ite Graduate First to Die from Dengue During This Epidemic. *The Straits Times*. Retrieved from http://www.straitstimes.com/singapore/20-year-old-ite-graduate-first-to-die-from-dengue-during-this-epidemic

World Health Organisation. (2016). Dengue and Severe Dengue. Retrieved from http://www.who.int/mediacentre/factsheets/fs117/en/

Zhang, J., & Mu, Q. (2016). Air Pollution and Defensive Expenditures: Evidence from Particulate-Filtering Facemasks. *working paper*.

Zivin, J. G., & Neidell, M. (2009). Days of Haze: Environmental Information Disclosure and Intertemporal Avoidance Behaviour. *Journal of Environmental Economics and Management, 58*, 119 - 128.

Zivin, J. G., Neidell, M., & Schlenker, W. (2011). Water Quality Violations and Avoidance Behaviour: Evidence from Bottled Water Consumption. *American Economic Review: Papers and Proceedings, 101*(3), 448 - 453.

Table 2.1: Summary statistics[1]

Variable	N[2]	Mean	SD	Min	Max
Monthly electricity consumption (kwh)					
3-room flats	48,369	279.8	44.85	56.09	771.53
4-room flats	125,603	384.74	61.95	110.75	1031.02
5-room and bigger flats	115,276	467.97	85.06	97.4	1350.55
Within 500m of a dengue cluster					
3-room flats	48,369	0.40	0.49	0	1
4-room flats	125,603	0.40	0.49	0	1
5-room and bigger flats	115,276	0.38	0.48	0	1
Within 1km of a dengue cluster					
3-room flats	48,369	0.67	0.47	0	1
4-room flats	125,603	0.67	0.47	0	1
5-room and bigger flats	115,276	0.65	0.48	0	1

[1] Statistics are weighted by the number of flats of each type in each block.

[2] Observations are at the block/year-month level. (N/23 months) gives the number of unique blocks.

Table 2.2: Effect of being near a dengue cluster on electricity consumption

Variable	4-room flats		5-room flats	
	With new blocks	Without new blocks	With new blocks	Without new blocks
	(1)	(2)	(3)	(4)
Near cluster	7.98***	6.64***	5.71***	5.28***
	(1.382)	(1.175)	(1.727)	(1.702)
Persist	10.25***	8.89***	6.57***	6.09***
	(1.403)	(1.187)	(1.787)	(1.758)
Block Fixed Effects	YES	YES	YES	YES
Postal District × Year × Month Fixed Effects	YES	YES	YES	YES
Observations	126,086	125,603	115,368	115,276
R-squared	0.706	0.705	0.761	0.761

Notes:
[1] The dependent variable is electricity consumption (kWh)
[2] "Near cluster" is defined as within 1100m of a cluster for 4-room flats and within 1400m of a cluster for 5-room flats.
[3] Regression estimates are weighted by the number of units of each type of flats in each block.
[4] Standard errors are reported in parentheses using clustered standard errors at the block level.

Table 2.3: Decay of persistence

Variable	4-room flats (1)	5-room flats (2)
Near cluster	6.55***	5.09***
	(1.177)	(1.709)
Near cluster × no. of mths after first treated	0.07	-0.02
	(0.047)	(0.056)
Persist	8.74***	4.55**
	(1.262)	(1.895)
Persist × no. of mths after post-cluster	0.16	0.75***
	(0.170)	(0.277)
Block Fixed Effects	YES	YES
Postal District × Year × Month Fixed Effects	YES	YES
Observations	125,603	115,276
R-squared	0.705	0.761

Notes:
[1] The dependent variable is electricity consumption (kWh)
[2] "Near cluster" is defined as within 1100m of a cluster for 4-room flats and within 1400m of a cluster for 5-room flats.
[3] Regression estimates are weighted by the number of units of each type of flats in each block.
[4] Standard errors are reported in parentheses using clustered standard errors at the block level.
[5] ***, **, * represent statistical significance at 1, 5, and 10 percent level respectively.

Table 2.4: Effect of risk classification on electricity consumption

Variable	4-room flats			5-room flats		
	(1)	(2)	(3)	(4)	(5)	(6)
Near cluster	6.11***	6.11***	6.14***	4.76***	4.85***	5.09***
	(1.189)	(1.186)	(1.182)	(1.713)	(1.714)	(1.715)
Near cluster × red banner	**1.97*** **	**2.01*** **	**2.12*** **	**1.74*** **	**3.29*** **	**4.58*** **
	(0.486)	(0.565)	(0.657)	(0.609)	(0.745)	(0.858)
Near cluster × cases	-	-0.00	-0.00	-	-0.03***	-0.08***
	-	(0.007)	(0.014)	-	(0.009)	(0.018)
Near cluster × cases^2	-	-	0.00	-	-	0.00***
	-	-	(0.000)	-	-	(0.000)
Persist	8.21***	8.35***	8.20***	5.47***	5.46***	6.07***
	(1.196)	(1.198)	(1.230)	(1.769)	(1.777)	(1.821)
Persist × red banner	**5.97*** **	**7.58*** **	**6.72** **	**6.21*** **	**6.26** **	**9.13*** **
	(1.383)	(1.952)	(2.629)	(1.853)	(2.672)	(3.504)
Persist × cases	-	-0.05	0.01	-	-0.00	-0.19
	-	(0.032)	(0.106)	-	(0.052)	(0.159)
Persist × cases^2	-	-	-0.00	-	-	0.00
	-	-	(0.001)	-	-	(0.001)
Block Fixed Effects	YES	YES	YES	YES	YES	YES
Postal District × Year × Month Fixed Effects	YES	YES	YES	YES	YES	YES
Observations	125,603	125,603	125,603	115,276	115,276	115,276
R-squared	0.705	0.705	0.705	0.761	0.761	0.761

Notes:
[1] The dependent variable is electricity consumption (kWh)
[2] "Near cluster" is defined as within 1100m of a cluster for 4-room flats and within 1400m of a cluster for 5-room flats.
[3] Regression estimates are weighted by the number of units of each type of flats in each block.
[4] Standard errors are reported in parentheses using clustered standard errors at the block level.
[5] ***, **, * represent statistical significance at 1, 5, and 10 percent level respectively.

Table 2.5: Robustness Checks

	4-room flats				5-room flats			
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Near cluster	6.76***	6.62***	6.64***	0.01***	5.74***	5.29***	5.28	0.005
	(1.174)	(1.173)	(2.501)	(0.003)	(1.700)	(1.702)	(3.693)	(0.003)
Persist	9.02***	8.87***	8.89***	0.02***	6.61***	6.10***	6.09	0.005
	(1.188)	(1.186)	(2.563)	(0.003)	(1.756)	(1.758)	(3.860)	(0.003)
Block Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES
Postal district × Month × Year Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES
Near construction sites and persistence variables	**YES**^	**YES**^^	**NO**	**NO**	**YES**^	**YES**^^	**NO**	**NO**
Level of clustered standard errors	**Block**	**Block**	**RC zone**	**Block**	**Block**	**Block**	**RC zone**	**Block**
Observations	125,603	125,603	125,603	125,603	115,276	115,276	115,276	115,276
R-squared	0.705	0.705	0.537	0.710	0.761	0.761	0.540	0.759

Notes:

[1] The dependent variable is electricity consumption (kWh) in columns (1) – (3) and (5) – (7), log(electricity consumption) in columns (4) and (8).

[2] "Near cluster" is defined as within 1100m of a cluster for 4-room flats and within 1400m of a cluster for 5-room flats.

[3] Regression estimates are weighted by the number of units of each type of flats in each block.

[4] ***, **, * represent statistical significance at 1, 5, and 10 percent level respectively.

^ Distance cut-offs for proximity to a construction site (in specifications (1) and (5)) are defined using the same cut-offs as proximity to a dengue cluster: within 1100m of a construction site for 4-room flats and within 1400m of a construction site for 5-room flats.

^^ Distance cut-offs for proximity to a construction site (in specifications (2) and (6)) are defined using cut-offs which show the highest relative treatment effects of construction sites for specifications (2) and (6): within 300m of a construction site for 4-room flats and within 100m of a construction site for 5-room flats. Results for the construction-related regressions are available upon request.

Table 2.6a: Total additional electricity consumption and carbon dioxide release[1] over the period May 2013 to Mar 2015

Flat type	No. of flats[2]	Average additional electricity consumption / flat (kWh)	Total additional electricity consumption ('000,000 kWh)	Average additional CO_2 release[3]/ flat (kg)	Total additional CO_2 release ('000,000 kg)
4-room flats	329,367	161.18	53.09	88.65	29.20
5-room flats	269,888	122.61	33.09	67.43	18.20
Overall	599,255	-	86.18	-	47.40

[1] Based on estimated coefficients of columns (2) and (4) of **Table 2.2**.
[2] Based on the balanced panel datasets we use for our analyses.
[3] Based on amount of carbon dioxide produced from electricity generation using natural gas, obtained from U.S. Energy Information Administration. Retrieved from https://www.eia.gov/tools/faqs/faq.cfm?id=74&t=11 on 9 May 2016.

Table 2.6b: Total cost of electricity consumption and carbon dioxide release[1] over the period May 2013 to Mar 2015, in 2015 Singapore Dollars

Flat type	No. of flats[2]	Average additional cost / flat of electricity consumption	Total additional cost of electricity consumption[3] ('000,000)	Total additional social cost of carbon[4] ('000,000)	Total cost of protective behaviour studied ('000,000)
4-room flats	329,367	40.83	13.45	0.60 – 5.74	14.05 – 19.19
5-room flats	269,888	31.08	8.39	0.37 – 3.58	8.76 – 11.97
Overall	599,255	-	21.84	0.98 – 9.32	22.82 – 31.16

[1] Based on estimated coefficients of columns (2) and (4) of **Table 2.2** and mean electricity consumption from January-April 2013.
[2] Based on the balanced panel datasets we use for our analyses.
[3] Based on historical electricity tariff rates obtained from Singapore Power. Historical tariff rates are retrieved from http://www.singaporepower.com.sg/irj/go/km/docs/wpccontent/Sites/SP%20Services/Site%20Content/Tariffs/documents/Historical%20Electricity%20Tariff.pdf on 9 May 2016.
[4] Based on the 2015 range of social cost of carbon reported in the Technical Support Document:-Technical Update of the Social Cost of Carbon for Regulatory Impact Analysis-Under Executive Order 12866 by the Interagency Working Group on Social Cost of Carbon, United States Government. We use an exchange rate of 1 2007 USD to 1.5 2007 SGD.

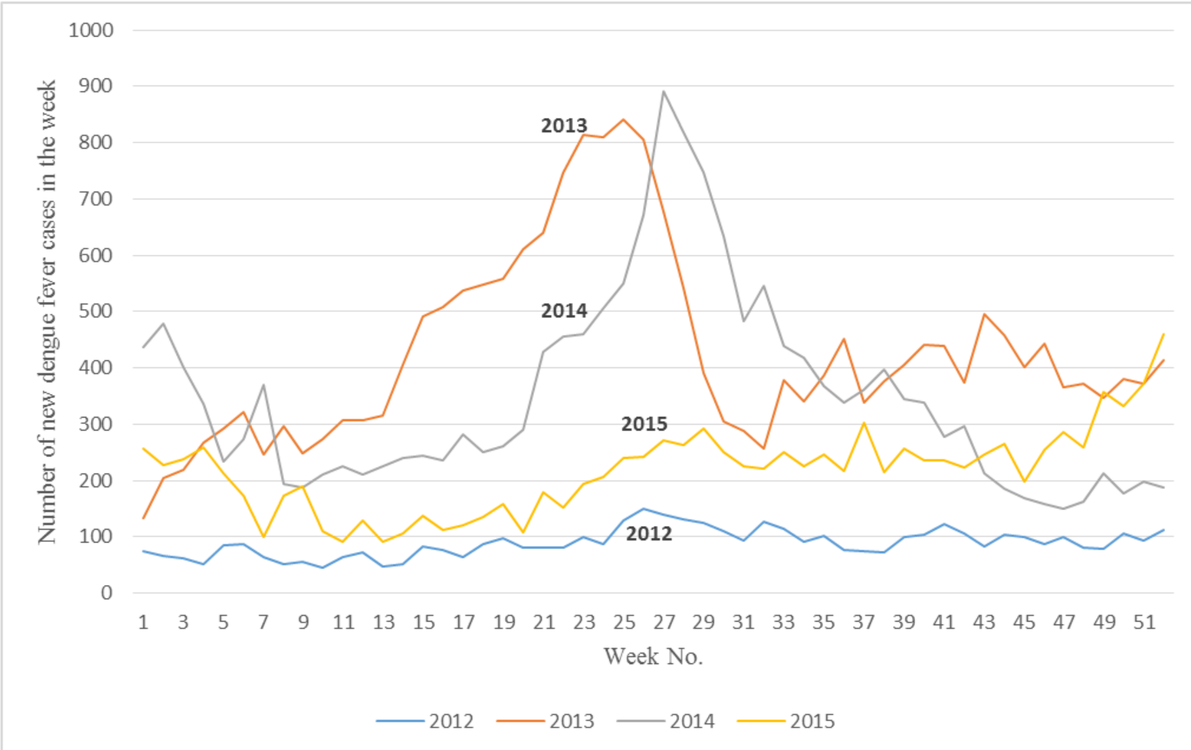
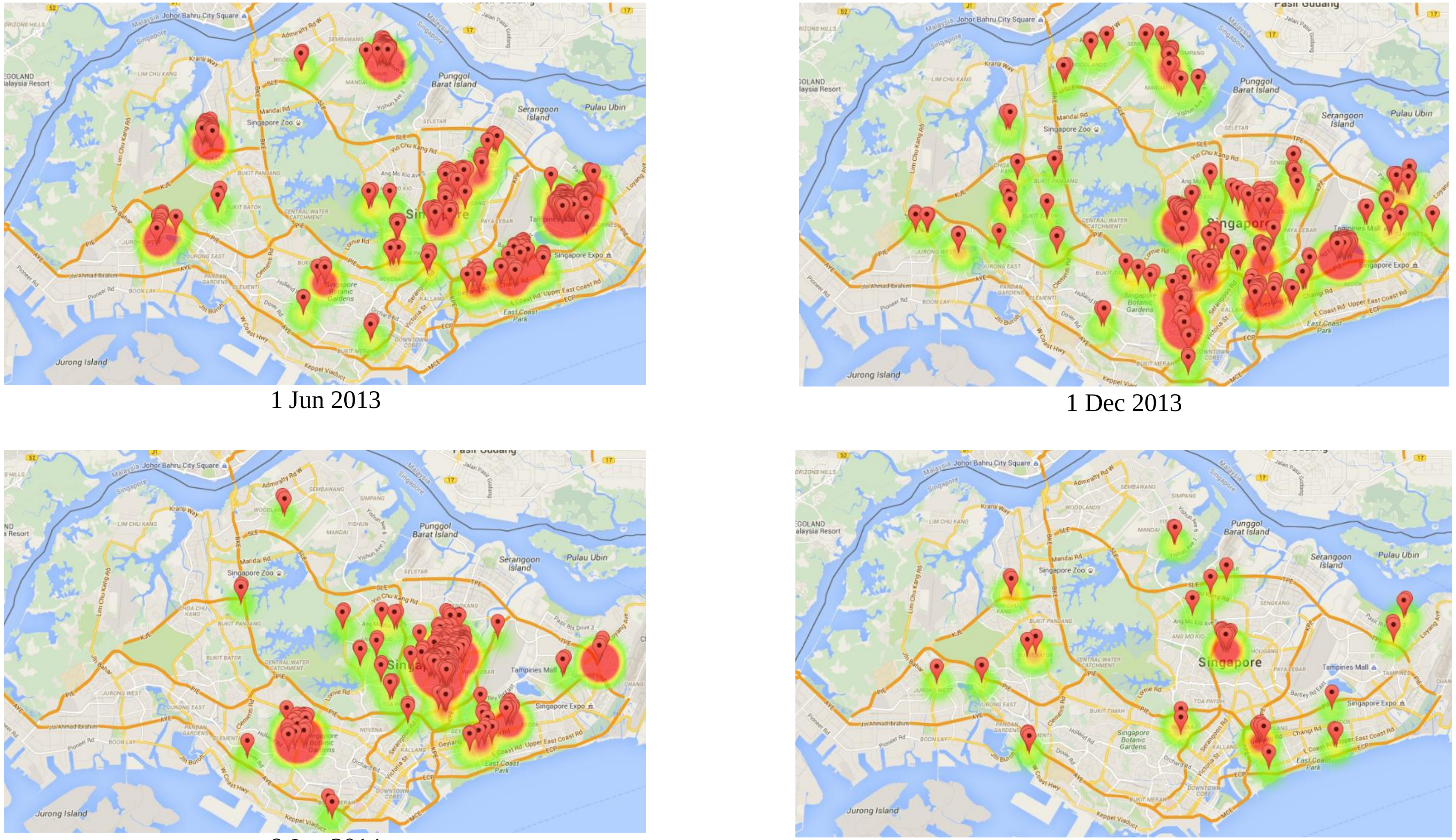


Figure 2.1: Number of new dengue fever cases each week from 2012 Jan to 2015 Dec[22]

[22] Source: Ministry of Health (2016)



1 Jun 2013

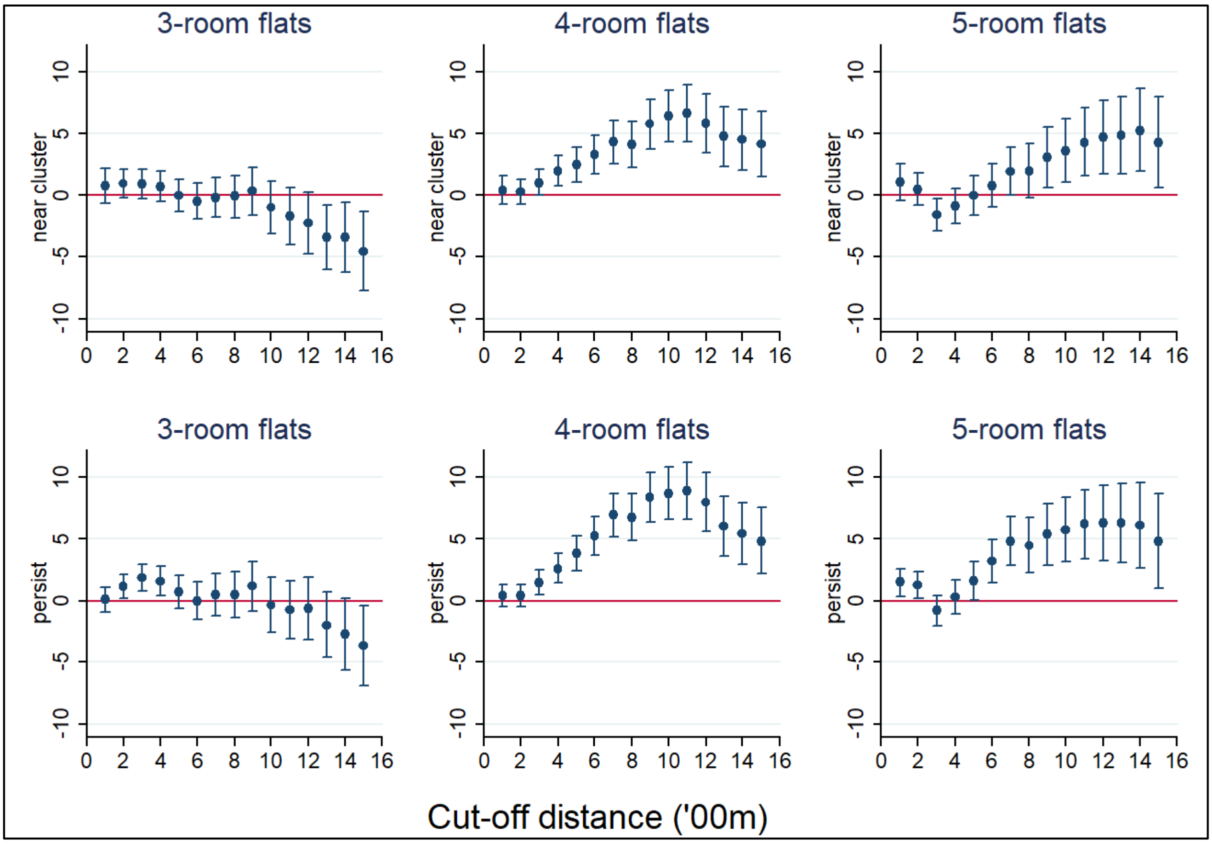
1 Dec 2013

8 Jun 2014

1 Dec 2014

Figure 2.2: Geographic and temporal variation in dengue clusters[23]

[23] Snapshots retrieved from http://outbreak.sgcharts.com/ on 3 May 2016. Map data: Google



Notes: The upper and lower rows respectively plot coefficients for the "near cluster" and "persist" variables from separate regressions using different cut-off distances to define whether a block is near a cluster. Each of these regressions includes block-specific as well as region-specific time fixed effects. Error bars represent 95% confidence intervals.

Figure 2.3: Estimated effect of being near a dengue cluster using different cut-off distances

Appendix A: Detailed information on dataset construction

We compile data with nationwide coverage from three administrative sources. From the Singapore Energy Market Authority (EMA)'s website, we download administrative data on electricity consumption (in kilowatt-hours, or kWh) for the period January 2013 to March 2015, for households living in public housing. This provides data with good nationwide coverage, as about 80% of the resident households stay in public housing in Singapore[24]. EMA provides average monthly consumption data[25] for each flat-type within a block of flats. Flat-types are categorised into (i) 1- or 2- room; (ii) 3-room; (iii) 4-room and (iv) 5-room/bigger flats. Electricity consumption will be our outcome variable.

Locational data on dengue clusters for the period May 2013 to October 2015 comes from the website outbreaks.sgcharts.com. Their data is compiled from administrative data published by NEA. Data is available for red- and yellow-coded clusters but not green-coded clusters. This means that we cannot differentiate between locations that are green-coded dengue clusters and locations that are not dengue clusters. We discuss the effect of this measurement error in Section 2.6.

We match locational data from the abovementioned datasets[26] to obtain the distance between each public housing block and its nearest dengue cluster for each month. We use this distance to construct a proxy for households' perceived dengue exposure risk.

---

[24] The statistic comes from Department of Statistics (2016b). Public housing in Singapore is meant to cater to the majority of the population, unlike in, say, the United States where public housing is meant for low-income households.

[25] The average monthly electricity consumption data is constructed using data at the household level, which is collected based on individual households' billing cycle. Different households within the same block of apartments can be billed based on different cycles. E.g. one household can be billed in March for consumption during 2 February to 1 March, while another can be billed for consumption during 15 February to 14 March.

[26] The dengue data is available at a higher frequency than the monthly electricity consumption data. The number of times dengue data is updated within a month varies by month. To match dengue data with electricity data, we restructure the dengue data as follows. In any month, we treat a location as a part of a dengue cluster if it has been identified in at least one of the updates within that month. To capture the intensity of dengue in any cluster (i.e. number of reported dengue cases) in a particular month, we take the maximum number of cases that are reported for that cluster in that month. We discuss the effect of this measurement error in Section 2.6.

These monthly datasets are combined to form three balanced panel datasets, each representing electricity consumption for 3-room, 4-room, and 5-room/bigger flats[27]. Each dataset tracks electricity consumption at the block level over 23 months from May 2013 to March 2015.

We merge the panel datasets with another dataset that records the number of flats of each flat-type within a block of flats. We obtain this dataset the website http://www.teoalida.com/singapore/hdbdatabase/ (who compiles the data from the Housing and Development Board's[28] website). This dataset allows us to weight the aggregated observations in the panel datasets by the number of relevant flats.

Lastly, we filter out newly built blocks of flats from our datasets. As households start moving into the new blocks, the blocks' average electricity consumption increases substantially over a short period. This phenomenon cannot be captured by region-specific time fixed effects and could bias our estimates if these new blocks are near dengue clusters.

[27] As of 2015, about 74% of Singapore resident households stay in these flat-types (Department of Statistics, 2016b). We will not be analysing data on 1-/2-room flats as they represent a very small percentage of households.
[28] This is the statutory board in charge of public housing in Singapore.

Appendix B: Tables of coefficients associated with Figure 2.3

Table B1: Impact of dengue cluster on electricity consumption (3-room flats)

Variable	Cut-off distances for "near cluster" variable														
	100m	200m	300m	400m	500m	600m	700m	800m	900m	1000m	1100m	1200m	1300m	1400m	1500m
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Near cluster	0.78	0.94	0.91	0.69	0.02	-0.49	-0.18	-0.09	0.36	-0.96	-1.71	-2.25*	-3.40**	-3.40**	-4.55***
	(0.717)	(0.592)	(0.600)	(0.619)	(0.675)	(0.741)	(0.824)	(0.870)	(0.994)	(1.084)	(1.173)	(1.261)	(1.324)	(1.441)	(1.629)
Persist	0.09	1.15**	1.87***	1.58***	0.69	-0.02	0.50	0.50	1.19	-0.35	-0.73	-0.64	-1.98	-2.72*	-3.65**
	(0.508)	(0.492)	(0.549)	(0.607)	(0.677)	(0.777)	(0.878)	(0.927)	(1.027)	(1.123)	(1.198)	(1.272)	(1.357)	(1.483)	(1.656)
Block Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Postal District × Year × Month Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	48,369	48,369	48,369	48,369	48,369	48,369	48,369	48,369	48,369	48,369	48,369	48,369	48,369	48,369	48,369
R-squared	0.725	0.725	0.725	0.725	0.725	0.725	0.725	0.725	0.725	0.725	0.725	0.725	0.725	0.725	0.725

Notes:
[1] The dependent variable is electricity consumption (kWh)
[2] "Near cluster" is defined as within 1100m of a cluster for 4-room flats and within 1400m of a cluster for 5-room flats.
[3] Regression estimates are weighted by the number of units of each type of flats in each block.
[4] Standard errors are reported in parentheses using clustered standard errors at the block level.
[5] ***, **, * represent statistical significance at 1, 5, and 10 percent level respectively.

Table B2: Impact of dengue cluster on electricity consumption (4-room flats)

Variable	Cut-off distances for "near cluster" variable														
	100m	200m	300m	400m	500m	600m	700m	800m	900m	1000m	1100m	1200m	1300m	1400m	1500m
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Near cluster	0.41	0.29	1.04*	1.99***	2.47***	3.34***	4.33***	4.11***	5.77***	6.42***	6.64***	5.83***	4.77***	4.49***	4.16***
	(0.575)	(0.508)	(0.531)	(0.609)	(0.719)	(0.801)	(0.874)	(0.947)	(1.035)	(1.065)	(1.175)	(1.214)	(1.219)	(1.246)	(1.347)
Persist	0.41	0.44	1.46***	2.61***	3.83***	5.26***	6.93***	6.77***	8.37***	8.70***	8.89***	8.00***	6.05***	5.44***	4.84***
	(0.452)	(0.462)	(0.505)	(0.604)	(0.715)	(0.808)	(0.885)	(0.968)	(1.045)	(1.076)	(1.187)	(1.222)	(1.239)	(1.278)	(1.376)
Block Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Postal District × Year × Month Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	125,603	125,603	125,603	125,603	125,603	125,603	125,603	125,603	125,603	125,603	125,603	125,603	125,603	125,603	125,603
R-squared	0.705	0.705	0.705	0.705	0.705	0.705	0.705	0.705	0.705	0.705	0.705	0.705	0.705	0.705	0.705

Notes:
[1] The dependent variable is electricity consumption (kWh)
[2] "Near cluster" is defined as within 1100m of a cluster for 4-room flats and within 1400m of a cluster for 5-room flats.
[3] Regression estimates are weighted by the number of units of each type of flats in each block.
[4] Standard errors are reported in parentheses using clustered standard errors at the block level.
[5] ***, **, * represent statistical significance at 1, 5, and 10 percent level respectively.

Table B3: Impact of dengue cluster on electricity consumption (5-room flats)

Variable	Cut-off distances for "near cluster" variable														
	100m	200m	300m	400m	500m	600m	700m	800m	900m	1000m	1100m	1200m	1300m	1400m	1500m
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Near cluster	1.05	0.50	-1.56**	-0.88	-0.033	0.78	1.94*	1.98*	3.09**	3.64***	4.30***	4.71***	4.86***	5.28***	4.27**
	(0.761)	(0.662)	(0.668)	(0.731)	(0.807)	(0.906)	(1.009)	(1.120)	(1.253)	(1.317)	(1.396)	(1.517)	(1.591)	(1.702)	(1.885)
Persist	1.48***	1.25**	-0.83	0.30	1.63**	3.19***	4.83***	4.48***	5.37***	5.74***	6.20***	6.30***	6.27***	6.09***	4.84**
	(0.565)	(0.567)	(0.632)	(0.715)	(0.800)	(0.904)	(1.009)	(1.132)	(1.275)	(1.339)	(1.428)	(1.553)	(1.642)	(1.758)	(1.937)
Block Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Postal District × Year × Month Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observa-tions	115,276	115,276	115,276	115,276	115,276	115,276	115,276	115,276	115,276	115,276	115,276	115,276	115,276	115,276	115,276
R-squared	0.761	0.761	0.761	0.761	0.761	0.761	0.761	0.761	0.761	0.761	0.761	0.761	0.761	0.761	0.761

Notes:
[1] The dependent variable is electricity consumption (kWh)
[2] "Near cluster" is defined as within 1100m of a cluster for 4-room flats and within 1400m of a cluster for 5-room flats.
[3] Regression estimates are weighted by the number of units of each type of flats in each block.
[4] Standard errors are reported in parentheses using clustered standard errors at the block level.
[5] ***, **, * represent statistical significance at 1, 5, and 10 percent level respectively.

3 Effects of In-utero Exposure to Maternal Fasting during Ramadan on Subjective Well-being and Health in Old Age

(Co-authored with Yanying Chen)

3.1 Introduction

A large number of studies have shown that exposure to adverse conditions in-utero, such as malnutrition, has lasting effects on the health and human capital of the child. Most work that looks at the causal effects of in-utero malnutrition, however, has utilised unusual events such as famines as their source of exogenous variation (Almond & Currie, 2011). As Almond and Mazumder (2011) note, results from studies of such severe events may be less applicable to milder and more common nutritional shocks (e.g. in-utero undernutrition due to morning sickness). A key policy question, then, is whether these milder nutritional shocks – for which successful intervention is more likely – also negatively affect a child's long-term outcomes. A recent strand of work[29] looks at precisely this question. These studies use maternal fasting during the Islamic holy month of Ramadan as a natural experiment, and find that even this relatively mild in-utero nutritional shock may have long-term negative effects on the child's health and socio-economic outcomes.

We use a new, high-frequency panel survey of elderly Singaporeans to add to this literature in two main ways. First, we extend the literature by evaluating the effect of in-utero exposure to Ramadan on currently unexamined evaluative subjective well-being measures such as life, economic and health satisfaction, as well as individuals' mental well-being. These measures may paint a fuller picture of an individual's socio-economic and health outcomes compared to more typically used objective measures such as income and employment. In

[29] Almond and Mazumder (2011); Van Ewijk (2011); Van Ewijk et al. (2013); Almond et al. (2015); Greve et al. (2015); Majid (2015); Schultz-Nielsen et al. (2016). See Section 3.2 for a review.

particular, life satisfaction, which is widely used as a utility proxy[30], can provide information on the overall welfare effects of in-utero Ramadan exposure.

Second, our study is set in an environment well suited to detect direct evidence of later-life medical problems predicted by the foetal origins hypothesis (FOH) *and* provide results that are more generalizable to non-immigrant populations. Compared to developing countries (where most of the long-term health-related studies in this literature have come from so far), the postnatal environment in Singapore is likely to conform more closely to environments that the FOH predicts will lead to health issues in old age[31]. We are thus able to carry out a more severe test of the FOH and detect more direct evidence (compared to past studies) of health effects predicted by the FOH in our analyses (see Section 3.5). Furthermore, as we study a largely non-immigrant population, our results are likely to be more generalisable to non-immigrant populations, compared to studies looking at immigrant populations.

As with other papers in this literature, our identification strategy compares outcomes between individuals who are exposed to Ramadan in-utero and those who are not. This avoids endogeneity associated with the decision to fast, and yields causal estimates of intent-to-treat effects of in-utero exposure to fasting during Ramadan[32]. A key identifying assumption is the lack of selective timing of pregnancy with respect to Ramadan. We show, in Section 3.6, that we find no evidence of selective timing in our sample.

Our data comes from the Singapore Life Panel, a longitudinal, population-representative survey of Singapore citizens and permanent residents aged between 50 and 70 (and their

---

[30] See e.g. Frey and Stutzer (2002) and Kahneman and Krueger (2006) for evidence on the suitability of using subjective well-being as a utility proxy.

[31] Singapore's postnatal environment is likelier to be food abundant and present a mismatch between the in-utero and postnatal environments. The FOH predicts that this may cause health problems in later life. E.g., adverse in-utero conditions may lead the foetus to develop physiological adaptations that would be evolutionarily advantageous in a food-scarce postnatal environment, but counterproductive in a food-abundant one. This could lead to higher risks of obesity, cardiovascular disease, and diabetes in later life (see e.g. Bateson et al., 2004; Gluckman et al., 2008). See Section 2 for more details.

[32] While we argue that the effects we observe are likely due to fasting, we are aware that other behavioural changes during Ramadan related to sleep patterns or fluid intake could contribute to our results.

spouses)[33], which is carried out monthly. We highlight two advantages of using this dataset. First, we can average the subjective well-being measures across waves, reducing the influence that transient events in any wave have on respondents' self-reported subjective well-being[34]. Second, the Centre for Research on the Economics of Aging (CREA), which runs this survey, made a strong effort to verify the accuracy of respondents' birthdates, which reduces error in our key treatment variables, the Ramadan exposure indicators (derived based on birthdates).

We find that individuals who were exposed to Ramadan in-utero have lower life satisfaction and poorer mental well-being, consistent with their lower levels of satisfaction within narrower domains such as social contacts/family life and daily activities. These results appear to be partially driven by lower economic satisfaction, as well as poorer self-rated health condition and satisfaction[35], which corresponds with what we observe for diagnosed medical conditions. Consistent with the predictions of the foetal origins hypothesis (FOH), individuals exposed in the second trimester are more likely to suffer from at least one cardiovascular condition. However, contrary to the predictions of the FOH, in-utero Ramadan exposure does not appear to lead to increased rates of diabetes in our sample. Lastly, we find that in-utero exposure to Ramadan increases the body mass index of women, and decreases the height and weight of men. In Sections 3.5 and 3.6, we show that our results are unlikely to be driven by seasonal effects, selective timing of pregnancies, or differing survey participation behaviour, and that they are robust to alternate specifications.

In general, our results on diagnosed medical conditions and anthropometric measures are consistent with predictions of the FOH, as well as findings from studies that use famines as exogenous nutritional shocks. We are able to find direct evidence that in-utero exposure to Ramadan increases rates of cardiovascular conditions and body mass index, strengthening

[33] Permanent residents make up only 2% of the Muslim sample we used in the study.
[34] For evidence that self-reported subjective well-being is partially influenced by transient events, see e.g. Schwarz and Strack (1999) (as cited in Krueger and Schkade (2008))
[35] Results for self-rated health satisfaction and condition hold mainly for those exposed in the second trimester.

existing evidence on the link between in-utero exposure to maternal fasting and adult health. As hypothesised earlier, our ability to detect this direct evidence may be due to the more food-abundant postnatal environment experienced by our sample.

The rest of this paper proceeds as follows. Section 3.2 provides background information and summarises findings from related papers. Section 3.3 describes our data. Section 3.4 explains our identification strategy and empirical specifications. Section 3.5 presents our results. Section 3.6 lists our robustness checks, and Section 3.7 concludes.

3.2 Background and Literature Review

3.2.1 Ramadan and Fasting

Ramadan is an Islamic holy month that lasts 29 – 30 days each year, and practicing Muslims fast during Ramadan by abstaining from food and water from dawn to dusk. Exemptions from fasting due to reasons such as illness or pregnancy are allowed, but exempted individuals may need to make up for the missed fasting and/or donate to the poor[36].

The start and end dates of Ramadan are determined by the Islamic calendar. This is a lunar calendar, which results in Ramadan shifting forward by about 11 days on the Gregorian calendar each year. Ramadan thus cycles through the entire Gregorian calendar every 33 years – a feature that a number of economists (e.g. Almond and Mazumder (2011), Van Ewijk (2011)) use to control for seasonality[37] when studying the in-utero effects of maternal fasting during Ramadan on various child outcomes.

While pregnant women may be excused from fasting during Ramadan, surveys summarised by Almond and Mazumder (2011) found 70 – 90% of pregnant women in different countries fast during Ramadan. In particular, a survey of pregnant Singaporean Muslim women

[36] See, e.g., the section on the Islamic ruling on pregnancy and fasting in page 5 of this FAQ by Majlis Ugama Islam Singapura (2016). Majlis Ugama Islam Singapura, or the Islamic Religious Council of Singapore, is a government statutory board overseeing Muslim affairs in Singapore.
[37] Controlling for seasonality is important as a large literature suggests that the season in which a child is born may have effects on its birth and adult outcomes such as birthweight or lifespan. See e.g. Doblhammer and Vaupel (2001) and Chodick et al. (2009).

by Joosoph et al. (2004) finds that 87% of pregnant Muslim mothers surveyed fasted for at least a day, while 74% fasted for more than 10 days. Results from the survey suggest that this could be because most (79%) mothers surveyed did not feel that fasting would harm themselves or their child. This perception could be driven by the lack of good evidence in 2004 that Ramadan fasting could have harmful effects on the foetus.

Fasting can lead to reduced caloric intake and other physiological changes which may affect the foetus. For example, more than 90% of pregnant women in Iran who fasted during Ramadan had a calorie deficit of at least 500 kcal in the 24 hours before breaking the fast (Arab, 2004). Even if total caloric intake is sufficient, fasting may lead to biochemical changes in the mother's body that affect the foetus. E.g., the metabolic profile of pregnant women who fast (especially those who fast in the day) can rapidly reach levels usually seen only in starving individuals, as pregnant women require energy for both themselves and their foetus[38].

3.2.2 Mechanisms linking in-utero Ramadan exposure to adult outcomes

The biomedical / epidemiological literature on the foetal origins hypothesis (FOH) suggests mechanisms by which adverse in-utero conditions described above may affect the child's long-term health negatively[39]. Almond and Mazumder (2011)'s online appendix and Van Ewijk (2011) summarise how the FOH may link in-utero Ramadan exposure to poorer long-run outcomes, while Godfrey and Barker (2001), Bateson et al. (2004), Gluckman et al. (2008), Gluckman et al. (2009), and Almond and Currie (2011) provide more general reviews of the FOH literature. This section summarises material from these reviews, and additional papers cited below, to outline how these mechanisms may lead to poorer health and economic outcomes in old age.

[38] See Almond and Mazumder (2011) and Van Ewijk (2011) for a summary of the biomedical literature on this.
[39] We acknowledge that other aspects of Ramadan, such as sleep deprivation due to having to wake up before sunrise, or rapid rises in blood sugar due to the heavy consumption of sweet drinks and snacks in the evening, may also contribute to the observed effects of Ramadan.

The foetal origins hypothesis (FOH) suggests that a foetus' response to adverse conditions in-utero, such as a lack of nutrition, can result in lasting physiological changes that may not lead to immediate health problems, but may predispose a child to poorer health (especially heart disease, hypertension, and type II diabetes) in adulthood[40]. Animal experimental studies suggest that even relatively short periods (48 hours or less) of adverse conditions such as undernutrition appear sufficient to lead to these changes (see e.g. Barker, 1997; Kett & Bertram, 2004). These changes are thought to occur as they enable the foetus to adapt to the adverse prenatal environment or potentially poorer postnatal environment signalled by the adverse in-utero conditions. While enabling the foetus to survive in the short run and better adapt to a food-scarce postnatal environment, these adaptations can compromise long-run health, especially if the postnatal environment is food-abundant.

For example, the foetus may divert nutrients to the brain at the expense of muscle and organ development in the limbs and trunk (e.g. the kidney) to improve short-term survivability. This could result in fewer kidney nephrons, increasing adult hypertension risk (Mackenzie & Brenner, 1995; Godfrey & Barker, 2001). In addition, the foetus may develop a "thrifty phenotype". This involves reduced skeletal muscle development and increased fat deposition, which would be evolutionarily advantageous in a food-scarce postnatal environment. If, however, the postnatal environment turns out to be food-abundant, the child will be at higher risk of obesity in later life. In turn, obesity results in higher risks of hypertension, heart disease, and diabetes (see e.g. Bateson et al., 2004; Gluckman et al., 2008).

Not all physiological changes arise due to the adaptations to improve short-term or post-natal survivability. Impaired in-utero nutrition can also slow foetal growth directly, leading to

[40] The general term for how stimuli during crucial periods of development (e.g. during foetal development) lead to lasting changes in physiology is known as "programming". Recent research indicates that this "programming" is likely to occur through epigenetic modification – environmentally induced changes to gene expression. See Vo and Hardy (2012) for a review of the epigenetic mechanisms by which adverse prenatal conditions lead to poorer adult health (especially heart disease, hypertension, and diabetes).

other changes that affect long-term health. E.g., impaired foetal growth may reprogram the foetus' hypothalamic-pituitary-adrenal axis, leading to hypertension or diabetes (Godfrey & Barker, 2001; Seckl & Holmes, 2007).

Beyond health, other channels can link in-utero nutritional shocks to poorer well-being in later life. Foetal malnutrition could impair physical and cognitive development, reducing initial human capital endowment and the ability to accumulate human capital. Alternatively, poorer adult health may impair one's ability to participate in the labour market (Heckman, 2007; Almond et al., 2015; Majid, 2015). Both channels suggest ways by which in-utero shocks could affect earnings, wealth accumulation, and eventually well-being in old age.

3.2.3 Famines as a natural experiment for studying the effects of foetal malnutrition

The mechanisms above give us some reason to believe that the epidemiological studies (see e.g. Godfrey and Barker (2001)) showing a correlation between in-utero malnutrition (reflected in birthweight) and poorer adult health may reflect a causal relationship.

The evidence for a causal relationship is strengthened by papers using famines as natural experiments to assess the causal impact of foetal malnutrition on the child's later-life outcomes. For example, Roseboom et al. (2003) report that people exposed the Dutch famine of 1944-45 in early gestation are more likely to rate their health as poor. Ravelli et al. (1998) and Stein et al. (2006) find evidence suggesting that prenatal exposure to famine is linked to lowered glucose tolerance and higher hypertension prevalence in adults respectively. In addition, Ravelli et al. (1999) and Stein et al. (2007) find that prenatal exposure to the Dutch Hunger Winter led to an increased body mass index (BMI) in middle-aged women.

3.2.4 Long-term impact of in-utero exposure to Ramadan fasting

A growing literature on a milder nutritional shock – prenatal exposure to maternal Ramadan fasting – found similar results on adult health. Almond and Mazumder (2011) find that Muslims in Uganda and Iraq who are exposed to Ramadan fasting in-utero during the early pregnancy are likelier to experience disabilities as adults. Using Indonesian data, Van Ewijk (2011) shows that exposed individuals have poorer objective (but not subjective) general health. Consistent with the foetal origins hypothesis (FOH), these individuals are also likelier to report symptoms that suggest coronary heart problems and diabetes. Somewhat surprisingly (given FOH predictions), exposed individuals were not more likely to suffer from hypertension. Van Ewijk et al. (2013) also find that adult Muslims who are exposed in-utero are slightly thinner[41] and have smaller stature.

The shadow of in-utero exposure to Ramadan stretches beyond health, and extends into cognitive and labour market outcomes. Majid (2015) finds that exposed children in Indonesia score lower on cognitive and math tests. Similarly, Almond et al. (2015) find that exposed Pakistani and Bangladeshi students in England have poorer academic outcomes at age seven. Using Danish data, Greve et al. (2015) document that exposed Muslim students score lower on national exams. Beyond cognition, Majid (2015) shows that exposed Indonesian adults work fewer hours per week and are more likely to be self-employed, while Schultz-Nielsen et al. (2016) find that exposed Muslim adults in Denmark are less likely to be employed.

In all, past research outlined in Section 3.2 suggests that in-utero Ramadan exposure is likely to affect the long-term health and economic outcomes of a child. In the rest of this paper, we extend this literature by examining the effects of in-utero Ramadan exposure on previously unexamined subjective measures of well-being, which may provide a more complete picture of the overall effects of Ramadan exposure. We also present complementary results on objective health outcomes that conform more closely to predictions from medical theory, strengthening existing evidence on the effects of in-utero Ramadan exposure on adult health.

[41] As Susser and Ananth (2013) note, this finding is different from that of the Dutch famine studies. E.g., Ravelli et al. (1999) and Stein et al. (2007) find that prenatal exposure to the Dutch Hunger Winter led to increased BMI in middle-aged women. We speculate that this may be because the postnatal environment faced by individuals in Van Ewijk et al. (2013)'s study is not sufficiently nutrient-rich to induce the maladaptive outcomes (i.e. higher BMI) of their thrifty phenotype.

3.3 Data and Variables

3.3.1 Data source and sample definition

We use data from waves 0 – 17 of the Singapore Life Panel (SLP), a population-representative survey of about 15,000 Singaporean citizens and permanent residents aged 50 to 70 (and their spouses) that is repeated monthly. The Centre for Research on the Economics of Aging (CREA) in Singapore Management University runs this survey, and has gone to some length to ensure that responses are population-representative, and that attrition from the panel remains low. CREA has also verified that the data collected by the SLP is indeed population-representative – distributions of variables such as income, educational level, and disease prevalence match Singapore government statistics.

In our analysis, we split our sample into Muslims (for the main analysis) and non-Muslims (for falsification checks). As the SLP does not collect data on religion, we use the ethnic group of Malays as a proxy for Muslims, as 99% of Singaporean Malays aged 15 and over are Muslims (Singapore Department of Statistics, 2011). Compared to earlier studies examining in-utero Ramadan exposure, our Muslim sample is relatively small, as only 12% of Singaporeans aged 50-70 are Malays (Singapore Department of Statistics, 2016). This disadvantage is negated somewhat as the postnatal environment in Singapore is more likely to be food abundant. Given the potential biological mechanisms outlined in Section 3.2, we expect the effects of in-utero Ramadan exposure to be stronger in our sample, counteracting the disadvantage of our smaller sample somewhat. To boost our sample size slightly, we include in our Muslim sample a very small group of individuals whose reported ethnicities meant that they are likely to be Muslims, e.g. Arab, Javanese.

3.3.2 Variables

Our key treatment variables are indicators for whether an individual was in-utero during Ramadan. Following earlier studies (e.g. Almond & Mazumder, 2011; Van Ewijk, 2011;

Majid, 2015), we determine if an individual is exposed to Ramadan in-utero by checking if Ramadan[42] overlaps with the period in which an individual was likely to be in-utero. We assume that an individual was in-utero in the 266 days (average length of human pregnancies) before his birth date[43]. We operationalise in-utero Ramadan exposure in two ways. The first is a single dummy variable that takes value one if Ramadan overlaps with pregnancy. The second is a set of three dummy variables that are trimester-specific. Each trimester dummy takes value one if Ramadan overlaps with that trimester. If Ramadan straddles two adjacent trimesters, both relevant dummies take value one.

Misclassification error can arise when pregnancies last longer than 266 days. In such cases, those who are exposed may be tagged as non-exposed, attenuating our estimates. To avoid this, we use a separate "buffer" indicator to tag those conceived less than 3 weeks after the end of Ramadan[44]. Our reference group thus consists of individuals who are definitely not exposed to Ramadan in-utero, based on their self-reported birth dates.

The second main source of misclassification error comes from preterm births. This may mean that some individuals classified as being exposed in their first trimester may actually not have been exposed. This attenuates the estimates of the overall and first trimester effects of being exposed to Ramadan. This also complicates the interpretation of the trimester results somewhat, as some individuals classified as exposed in the second / third trimester may have actually been exposed during the first / second trimester.

The main outcome variables we study capture evaluative subjective well-being (e.g., life satisfaction, economic satisfaction[45]), mental well-being (e.g. how much of the time the

[42]Ramadan dates are from http://www.timeanddate.com/holidays/saudi-arabia/ramadan-begins (accessed 15 Jun 2016). These dates match those used in Almond and Mazumder (2011) for our period of study.
[43] Birth dates are self-reported by panel members. This may lead to misclassification due to misreporting by individuals, but this concern is assuaged somewhat by the lack of heaping around prominent dates such as 1 Jan.
[44] We follow Van Ewijk (2011) and choose 3 weeks as the buffer period, as Kieler et al. (1995) find that few pregnancies last beyond 3 weeks past term.
[45] While the SLP also contains data on income and employment, these variables are less suitable for analysis as individuals in our sample are close to retirement. Individuals exposed to Ramadan in-utero could have lower

respondent was happy), self-reports of whether individuals have been told by doctors that they have medical conditions such as hypertension, as well as self-reports of height and weight. These variables capture individuals' status as at the time of survey and therefore their later-life outcomes. In cases where individuals respond to the same question over several waves (i.e. questions on subjective well-being and mental well-being), their responses are averaged across the number of waves in which they responded. Section 3.3.3 provides more details and reports descriptive statistics for key outcome variables.

3.3.3 Summary Statistics

Table 3.1 reports summary statistics for some demographic variables, exposure indicators and selected outcome variables. Slightly less than half (47 – 48%) of our Muslim and non-Muslim samples are male, and our respondents are close to retirement age in general: the mean age for Muslims is 58; that for non-Muslims is 59. Respondents in our sample are not highly educated, which is not surprising for this birth cohort. About 70% of the sample has some secondary education, while only 18% of Muslims and 32% of non-Muslims have some form of postsecondary education. The proportions of Muslim and non-Muslim individuals exposed to Ramadan are similar across our exposure-related variables.

In addition to basic demographics, we report our main outcome variables, partly to provide some context for interpreting our regression results in Section 3.5. One such variable of interest is self-reported body mass index (BMI). The mean BMI for Muslims and non-Muslims is about 27 and 24 respectively, indicating that the average person in this cohort is somewhat overweight. This suggests that the postnatal environment in Singapore for these individuals was likely to be food-abundant, and thus likely to provide a suitable environment for testing predictions from the foetal origins hypothesis[46].

incomes because of lower health and human capital, or higher incomes because they were less able to accumulate wealth and are not yet able to retire.

[46] See footnote 31 and Section 3.2 for more details on why this is the case.

Turning to self-reports of diagnosed chronic conditions, we find that self-reported disease diagnosis rates are similar to rates reported in government surveys. Our samples' reported diabetes rates were 26% and 16% for Muslims and non-Muslims respectively, and hypertension rates were 36% and 35% for Muslims and non-Muslims respectively. In comparison, the 2010 National Health Survey finds that diabetes prevalence is 19% among those aged 50 – 59 and 29% among those aged 60 – 69, while hypertension prevalence rates were 32% and 53% for those aged 50 – 59 and 60 – 69 respectively (Ministry of Health, 2011). Given that SLP sample means are not age-adjusted and are diagnosis, rather than prevalence rates, our self-reported diagnosis rates are not too far from government-reported prevalence rates, suggesting that under-reporting of diagnosis should not be a major issue.

To assess an individual's evaluative subjective well-being, we look at a broad indicator of overall life satisfaction, as well as satisfaction within narrower domains such as economic situation and health. These variables are rated on a scale of 1 (worst) to 5 (best). These data are collected via the following questions: (i) "Taking all things together, how satisfied are you with your life as a whole these days?"; (ii) "How satisfied are you with your social contacts and family life?; (iii) "How satisfied are you with your daily activities, and if you are working, your job?"; (iv) "How satisfied are you with your overall economic situation?"; (v) "How satisfied are you with your health"; and (vi) "Would you say your health is excellent, very good, good, fair, or poor?". For the first five questions, respondents choose from "very dissatisfied", "dissatisfied", "neither satisfied nor dissatisfied", "satisfied", and "very satisfied", while the options to the last question are "poor", "fair", "good", "very good" and "excellent". These questions are asked monthly, so each respondent's answers are averaged across the number of waves in which they participated. This reduces the influence that transient events in any one interview[47] could have on respondents' subjective well-being. On average, respondents report being somewhat satisfied with the various domains of their life (3 corresponds to "neither satisfied nor dissatisfied"; 4 corresponds to "satisfied"), and they rate their health condition between fair (2) and good (3).

There is evidence to suggest that these single-item subjective well-being measures can contain useful information about individuals' actual utility / well-being. Life satisfaction questions correlate well with demographic variables such as education and income (see e.g. Kahneman and Krueger (2006) or Diener et al. (2013) for reviews on the validity of life satisfaction data). Self-assessed health has also been found to be consistent with objective health (e.g. Wu et al., 2013) and is an independent predictor of mortality, even after conditioning on objective health (Idler & Benyamini, 1997).

Lastly, four questions provide information on respondents' mental well-being: (i) "During the past 30 days, how much of the time have you felt worn out?"; (ii) "During the past 30 days, how much of the time have you been a happy person?; (iii) "Overall in the last 30 days, how much difficulty did you have sleeping, such as falling asleep, waking up frequently during the night or waking up too early in the morning?"; and (iv) "Overall in the last 30 days, how much of a problem did you have with feeling sad, low, or depressed?". For the first two questions, responses are given on a scale of 1 to 6 (1 – All of the time, 2 – Most of the time, 3 – A good bit of the time, 4 – Some of the time, 5 – A little of the time, 6 -- None of the time). For the last two questions, responses are given on a scale of 1 to 5 (1 – None, 2 – Some, 3 – Moderate, 4 – Severe, 5 – Extreme). For analyses and reporting, we flip the scale where relevant, so that higher values always represent better mental well-being. As with the evaluative subjective well-being measures, respondents' answers to each question are averaged across the

[47] See e.g. Schwarz and Strack (1999) (as cited in Krueger and Schkade (2008)) that transient events can affect self-reported satisfaction measures.

number of waves in which they participated[48]. We summarise the data from these four questions into a mental well-being index, by first normalising[49] the averaged responses to each question, and then summing the four normalised measures.

In all, the basic demographics, exposure indicators and outcome variables are generally similar for our Muslim and non-Muslim respondents, providing some reassurance that non-Muslims can serve as a reasonable, albeit imperfect, comparison group for falsification checks.

3.4 Identification Strategy and Empirical Specifications

Our basic specification, estimated by OLS, is

$$Y_i = \alpha + \beta Exposure_i + \gamma X_i + \varepsilon_i \qquad (1)$$

where Y_i is the later life outcome of individual i and $Exposure_i$ captures in-utero Ramadan exposure either as a single dummy or a set of three dummies for different trimesters. X_i covers individual-specific covariates which include age (defined as the difference between 2016 and the birthyear), age-squared, gender, a buffer indicator[50], and calendar month of birth dummies. ε_i is the error term. Standard errors are clustered at the household level.

Our identification strategy follows earlier papers (e.g. Almond & Mazumder, 2011; Van Ewijk, 2011) studying the effect of in-utero exposure to Ramadan, and compares individuals who were exposed in-utero against those who were not. We thus estimate intent-to-treat effects since we do not know if individuals were indeed exposed in-utero. While this underestimates the exposure effect, it avoids endogeneity associated with parents' decision to fast. The identifying assumption, that allows us to interpret the estimated effect as causal, is that parents do not practice selective timing of pregnancies with respect to Ramadan. If there is selective timing, our results could be biased since unobserved characteristics of parents can influence

[48] These questions are posed quarterly instead of monthly, resulting in a smaller available sample.
[49] We normalise the averaged responses in the Muslim and non-Muslim samples by using the mean and standard deviation statistics from the respective samples.
[50] As mentioned in Section 2, misclassification error can arise when pregnancies last more than 266 days. In such cases, those who are exposed may be tagged as non-exposed, attenuating our estimates. To avoid this, we use a separate "buffer" indicator to tag those conceived less than 3 weeks after the end of Ramadan.

whether an individual was exposed to Ramadan in-utero as well as his/her later-life outcomes. In Section 3.6, we argue that the identifying assumption is generally valid in our study.

For each outcome, we estimate **Eq (1)** on two samples: Muslims only; and non-Muslims only. The estimation on non-Muslims provides a falsification check. Since the ethnic composition of Singapore is such that the non-Muslim sample is much larger than the Muslim sample, we avoid the possibility that a non-significant result in the falsification test may be due to lower statistical power rather than a null result (Susser & Ananth, 2013).

We also perform a difference-in-differences (DiD) analysis on the combined sample of Muslims and non-Muslims, to control for possible seasonal effects that are not already captured by the month-of-birth fixed effects. We do so as most of our Muslims-only sample was born in years where Ramadan fell in the months of January to July, and month-of-birth fixed effects may not be sufficient to control for seasonality[51]. To **Eq (1)**, we add a dummy variable for Muslim, as well as the interaction terms between (i) the Muslim indicator and individual-specific covariates (X_i) and (ii) the Muslim indicator and exposure indicators. The coefficient(s) for the interaction term(s) between the Muslim and exposure indicator(s), β_1, is the coefficient of interest. The full specification is shown in **Eq (2)** below.

$$\begin{aligned} Y_i \quad = \quad & \alpha + \beta_1 (Exposure \times Muslim)_i + \beta_2 Exposure_i + \beta_3 Muslim_i \\ & + \gamma_1 X_i + \gamma_2 (Muslim \times X_i) + \varepsilon_i \end{aligned} \tag{2}$$

Having said this, we argue that season-of-birth effects are less likely to confound the effects of in-utero exposure to Ramadan in Singapore. Singapore is located on the equator, and has much lower seasonal variation in day length and temperature as compared to temperate countries. This removes a key source of variation that is likely to drive the observed effect of

[51] We acknowledge that non-Muslims in Singapore are not a perfect control due to differences such as dietary habits between the two groups. Nonetheless, the falsification checks on the non-Muslim sample and the DiD reduce the chance that seasonal effects are driving our results.

seasonality on child outcomes in temperate countries (Chodick et al., 2009). Furthermore, Singapore – even in the 1940s to 1960s – was an urban centre, rather than an agricultural economy. This means that rainfall and the consequent changes in food availability, which is a major driver of seasonal effects in less developed / tropical countries (Yamauchi, 2012; Rocha & Soares, 2015), is less likely to lead to seasonality effects in child outcomes in Singapore.

3.5 Results

For all results tables, we show (i) the trimester-specific effects and the overall exposure effect of in-utero exposure to Ramadan, (ii) results from our falsification tests, and (iii) our DiD results. In the interest of space, we describe only results from the Muslim-only sample in the main text. These are similar to the DiD results, as in-utero exposure to Ramadan generally has no effect on outcomes in our non-Muslim sample.

3.5.1 Life satisfaction and mental well-being

We start by looking at individuals' overall life satisfaction, which may also be seen as a utility proxy. Life satisfaction is measured on a scale of 1 (worst) to 5 (best). Responses are averaged over the number of waves an individual participated. In-utero Ramadan exposure may affect overall life satisfaction in several ways discussed in Section 3.2, e.g. through its effect on health and/or economic outcomes. Column (1) of **Table 3.2** shows that individuals exposed to Ramadan in-utero have lower overall life satisfaction -- a marginally significant overall effect of -0.10 (3% worse than non-exposed individuals relative to the mean or 0.18SD). This effect is strongest in trimester two, where in-utero exposure has a significant effect of -0.078 (2% of the mean / 0.13SD).

An alternative broad measure of subjective well-being is an individual's mental well-being, captured using a mental well-being index constructed from four questions (details in Section 3.3.3). A higher index value corresponds to better mental well-being. Column (4) in **Table 3.2** shows that exposure leads to poorer mental well-being in old age – a statistically significant effect of -0.80 (0.26 SD[52]). Trimester-wise, in-utero exposure during the second trimesters leads to a statistically significant effect of -0.65 (0.21SD[52]). Results from analyses using the individual questions are in **Tables A1 and A2** in the appendix.

3.5.2 Social/Family and Daily activities satisfaction

To investigate potential drivers of the above results, we delve into narrower domains of evaluative subjective well-being. We start with social/family life and daily activities satisfaction. A higher value of the dependent variable indicates a higher level of satisfaction. **Table 3.3** (Columns (1) and (4)) show that the overall exposure effect is not statistically significant for both measures, but individuals exposed to Ramadan during their second trimester feel less satisfied in both domains. Their exposure leads to a statistically significant effect of -0.079 (2% relative to the mean or 0.15SD) and -0.097 (3% relative to the mean or 0.17SD) for social/family and daily activities satisfaction respectively. It is possible that poorer satisfaction in these aspects are driven by health and economic status.

3.5.3 Economic satisfaction

We look next at individuals' overall economic satisfaction[53]. Column (1) in **Table 3.4** shows that in-utero exposure to Ramadan results in individuals reporting lower economic satisfaction (their economic satisfaction rating is a statistically significant 0.15 lower than those who were not exposed). Relative to the mean, those exposed were about 4% less satisfied with their overall economic conditions than non-exposed individuals (or 0.21SD). Trimester-wise, exposure during the second trimester leads to a marginally significant decrease of 0.084 in economic satisfaction rating (2% of the mean or 0.12 SD). These results are consistent with

[52] Coefficient is not expressed in terms of the mean, since the index is the sum of normalised measures.
[53] As we argued earlier, while the SLP also contains data on income and employment, these variables are less suitable targets for analysis as individuals in our sample are close to retirement. Individuals exposed to Ramadan could have lower incomes because of lower health and human capital, or higher incomes because they were less able to accumulate wealth and are not yet able to retire.

those on life satisfaction and mental well-being, suggesting that the exposure effect on economic outcomes may contribute to effects on broader outcomes.

3.5.4 **Self-**rated **health**

We explore the other potential driver – respondents' self-rated health satisfaction and condition[54] (rated on a scale of 1(worst) to 5(best)). **Table 3.5** suggests that those exposed to Ramadan in-utero may have worse self-rated health for both outcomes (though the overall effects are not statistically significant). In terms of trimester effects, Columns (1) and (4) show that exposure during the second trimester leads to a statistically significant 0.12 and 0.10 fall in health condition and health satisfaction respectively. Relative to the mean, individuals exposed in the second trimester rate their health 3-4% (or 0.15SD) worse than those non-exposed. Again, these results are consistent with those on life satisfaction.

Separately, we note that our result on self-rated health differs from that of Van Ewijk (2011), the only other paper which looks at the effect of in-utero Ramadan exposure on self-rated health. While Van Ewijk (2011) finds that in-utero exposure to Ramadan has negative effects on general objective health[55], he finds positive effects for self-rated health. The difference between our papers may be due to the trickier nature of the question used to measure self-rated health in Van Ewijk (2011)'s study[56].

3.5.5 Diagnosed medical conditions

To understand why exposed individuals reported worse self-rated health, we turn to objective measures of respondents' health: a question that asks respondents if they have ever been told by a doctor that they have medical conditions such as hypertension, stroke, heart problems, diabetes, cancer, psychiatric conditions, or arthritis.

[54] There is evidence suggesting that such measures are useful predictors of mortality (see e.g. Idler & Benyamini, 1997).
[55] General objective health is rated by professional health workers on a 9-point scale.
[56] The question asked whether an individual's own health was better or worse than that of another person of the same age and sex.

Given the bulk of the predictions by the foetal origins hypothesis (FOH), we would expect results for cardiovascular problems and/or diabetes, but not other conditions such as cancer, to correspond with the self-rated health results. This is indeed what we see in general. **Table 3.6** shows the effect of in-utero exposure to Ramadan on the self-reported rates of cardiovascular conditions (hypertension, heart problems and stroke) and diabetes. Column (1) shows that exposure during the second trimester results in a statistically significant 9.3 percentage point increase in the probability of having at least one cardiovascular condition. This translates to a 22% increase relative to the mean rate in the Muslim sample, which may explain why individuals exposed in the second trimester reported poorer health satisfaction and condition. **Table A3** in the appendix shows that much of this result is driven by hypertension.

We do not, however, find an effect on diabetes in our sample (see Column (4) of **Table 3.6**)[57]. In addition, we find that in-utero exposure to Ramadan has no effect on psychiatric problems or cancer (see **Table A4**)[58], or the ability to carry out activities of daily living (results available on request).

In all, our results conform closely to the main predictions of the FOH, and strengthen existing evidence (e.g. Van Ewijk (2011)) on the link between in-utero exposure to maternal fasting and adult health. While Van Ewijk (2011) find that exposure leads to an increase in reports of chest pains among older Muslims (an indirect indication of cardiovascular conditions), he does not find effects on hypertension. As argued earlier, this may be because the postnatal environment faced by our Singaporean sample is likelier to be food abundant,

[57] We note that exposure to Ramadan leads to a statistically significant lower probability that non-Muslims will suffer from diabetes. In light of the multiple outcomes analysed, this significant effect is likely to be just due to chance.

[58] Exposure in the second trimester increases the probability of being diagnosed with arthritis by a marginally significant 0.036, but this effect may be spurious. To the best of our knowledge, arthritis has not been specifically predicted by medical theory. As this result is only marginally significant, there is a higher chance that the result may simply be due to chance.

which means that the "thrifty phenotype" developed by individuals exposed to Ramadan in-utero is likelier to be a maladaptation which can cause health issues in later life.

3.5.6 Anthropometric measures59

Other than medical conditions such as cardiovascular problems, the maladaptation of the "thrifty phenotype" to the postnatal environment could also manifest in the form of faster weight gain (see e.g. Bateson et al., 2004; Gluckman et al., 2008) . **Table 3.7** shows results on body mass index (BMI) based on the full sample of Muslims as well as separate samples of men and women, following studies which use the Dutch famine as a natural experiment (Ravelli et al., 1999; Stein et al., 2007). In-utero Ramadan exposure leads to a marginally significant increase in BMI for women (with the strongest effects present in those exposed during trimester 2) but not for men[60]. While this differs from Van Ewijk et al. (2013) who also look at the effect of in-utero Ramadan exposure on BMI, our result is consistent with predictions from the foetal origins hypothesis (FOH), as well as empirical work from the Dutch famine studies by Ravelli et al. (1999) and Stein et al. (2007). Again, we believe that our results differ from Van Ewijk et al. (2013) (who use Indonesian data) because Singapore may be a more suitable location to observe these particular effects predicted by the FOH.

Consistent with the BMI effects we find for women, **Table A6** shows that in-utero exposure leads to a statistically insignificant increase in weight for women. In contrast, overall in-utero exposure leads to a marginally significant decrease of 5.31kg for men, with the effect being most pronounced in the first trimester (column (7)). In-utero exposure to Ramadan also leads to a statistically significant decrease in height of about 2.2cm in our overall sample (see **Table A7**). This effect is strongest when the foetus is exposed in trimester one, and is driven mostly by men.

[59] The sample size here is much smaller than others as weight / height data was collected only in one wave.
[60] We also study if individuals are overweight (BMI>=23) or obese (BMI>=27.5). Exposed women have a marginally significant higher probability of being overweight (see **Table A5** in appendix), but we do not find statistically significant results for obesity (results available on request).

3.6 Validity of identifying assumption and robustness checks

3.6.1 Selective timing of pregnancy

The causal interpretation of our results depends on the validity of the identifying assumption that parents do not practice selective timing of pregnancies with respect to Ramadan. Our checks in this section show no evidence of such selective timing.

We start by checking if observed characteristics of individuals' parents vary significantly by individuals' in-utero exposure to Ramadan. The lack of a statistically significant difference would provide reassurance that other unobserved characteristics of their parents do not differ across individuals who were and were not exposed to Ramadan in-utero.

We do so by estimating a variant of **Eq (1)** on the Muslim-only sample, using characteristics of respondents' parents as the dependent variables. While we do not have information on parental socioeconomic characteristics at the point of pregnancy, our dataset does contain three variables that may act as summary measures of the health and socio-economic status of the individuals' parents. They are (i) whether the parent is alive; (ii) the current age of the parent if he/she is alive; and (iii) the age at which the parent passed away if the parent is not alive. These three variables are available separately for the father and mother.

From **Table 3.8**, we see that only 3 out of 24 coefficients are marginally significant (see column (6)), which is close to what one would expect from chance, under the null hypothesis. The marginally significant coefficient in Column (3) has a sign opposite from what we would expect, i.e. if selective timing of pregnancies is driving our results, we would expect a negative correlation between the age at which the father passes away and whether the individual was exposed in-utero, as age is a proxy for better health and socio-economic status. Both marginally significant coefficients in Column (4) are opposite in sign, suggesting that there is no clear relationship between the age at which the mother passes away and in-utero Ramadan exposure.

The check above, however, may not be able to detect time-varying selection that Ahsan (2015) argues that we should be concerned about as well. Such time-varying selection could arise from increased availability of contraceptives / family planning programmes, which might lead to educated mothers timing their pregnancies to avoid Ramadan.

In our view, such time-varying selection is unlikely to be an issue in our sample. Most of our respondents were born before 1966, the year in which a statutory board was established to oversee family planning in Singapore (National Library Board, 2016). Before this, only a small percentage of the population received family planning services. Moreover, unlike the intensive Bangladeshi family planning programme studied by Ahsan (2015) (which gave free contraceptives to all women of reproductive ages, door to door every fortnight), free or subsidised supplies were provided only to the needy (Lim, 2010). Nonetheless, to rule out time-varying selection that could have arisen due to the setting up of the family planning statutory board, we re-estimate the results on later-life outcomes on a sub-sample of Muslims born before 1966. Our results are generally robust (see **Table 3.9**).

Beyond this, we would expect that if time-varying selection did occur, the proportion of Muslims exposed to Ramadan in-utero would vary by age, relative to the proportion of non-Muslims in-utero during the same period. We do not observe such a pattern.

Figure 3.1 plots the proportion of Muslims and non-Muslims who were in-utero during Ramadan, by age. Visually, there is little evidence of a fall in Ramadan exposure for younger Muslims relative to non-Muslims. When we test this more formally by regressing the exposure indicators on age and calendar-month-of-birth fixed effects (see **Table 3.10**), the DiD results in columns (3), (6), (9), and (12) show little evidence that in-utero exposure to Ramadan varies significantly by age, once we take population-wide trends into account.

3.6.2 Differing survey participation behaviours between treatment and control

Many outcomes we use are computed as averages across the number of waves in which individuals responded. Our results could hence be driven by differing participation behaviour across exposed and non-exposed individuals. We compare the number of times individuals responded to the surveys, across the exposed and non-exposed groups of the Muslim-only sample[61]. We find no evidence of such differing participation between these two groups (t-statistic = -0.0935, p-value = 0.9255).

3.6.3 Other regression specifications

We test if our results are sensitive to different regression specifications and sample definitions. Instead of including age as a scale variable, we replace it with birth year fixed effects. Where appropriate, we estimate **Eq (1)** with a logistic regression. We also exclude permanent residents from our sample, keeping only Singapore citizens, with the aim of defining samples with an even greater degree of homogeneity in experiences. Our results are generally robust to these alternative specifications (results available on request).

3.7 Conclusion

Using a new, high-frequency dataset from Singapore (an environment that is well-suited for studying the long-term effects of in-utero undernutrition), and exploiting the plausible randomness of in-utero exposure to maternal fasting during Ramadan as a natural experiment, we find that the relatively mild nutritional shocks experienced during Ramadan can have lasting effects on the long-term outcomes of the child. Exposed individuals have lower life, social/family, daily activities, economic, and health satisfaction, and they rate their own health condition more poorly. They also seem to have poorer mental well-being. In line with predictions from the foetal origins hypothesis (FOH), these individuals are more likely to be diagnosed with cardiovascular conditions and have a higher body-mass index (for women). In

[61] Individuals estimated to be conceived less than 21 days after the end of Ramadan are excluded in this check.

our sample, the effects from exposure are most pronounced for individuals exposed in the second trimester. These results are unlikely to be driven by seasonal effects common to all individuals, selective timing of pregnancies, or differing survey participation behaviour.

These results extend the literature by studying exposure effects on evaluative subjective well-being and mental well-being, which may paint a fuller picture of an individual's health and socio-economic outcomes. In particular, life satisfaction is widely used as a utility proxy and can provide information on the overall welfare effects of in-utero Ramadan exposure, compared to past studies that study more specific outcomes. Furthermore, our results with respect to anthropometric measures and diagnosed health conditions conform more closely to those predicted by the FOH, strengthening existing evidence on the link between in-utero Ramadan exposure and health.

As we study a milder nutritional shock, our findings add credence to the idea that the negative effects predicted by the FOH may apply to milder nutritional disruptions in developed countries too. Our findings also have implications for advice on fasting while pregnant during Ramadan. While Islamic law exempts Muslim women from fasting during Ramadan if they are pregnant, many still do so. Findings from our paper (and others in this literature) provide new information that parents and healthcare workers could use when considering how best to manage fasting during pregnancy. For policymakers who wish to bolster existing country-related findings, replication studies using other samples from the same country could be conducted. Further research on whether managing the fasting process more carefully could mitigate the effects of Ramadan fasting, or on elucidating the biological and epigenetic mechanisms that link fasting to poorer health outcomes, would also be helpful in advancing this line of research and in recommending more appropriate interventions.

References

Ahsan, M. N. (2015). Do Parents Selectively Time Birth Relative to Ramadan? Evidence from Matlab, Bangladesh Job Market Paper.

Almond, D., & Currie, J. (2011). Killing Me Softly: The Fetal Origins Hypothesis. *The Journal of Economic Perspectives, 25*(3), 153-172.

Almond, D., Mazumder, B., & van Ewijk, R. (2015). In Utero Ramadan Exposure and Children's Academic Performance. *The Economic Journal, 125*(589), 1501-1533.

Almond, D., & Mazumder, B. A. (2011). Health Capital and the Prenatal Environment: The Effect of Ramadan Observance During Pregnancy. *American Economic Journal: Applied Economics, 3*(4), 56-85.

Arab, M. (2004). Ketonuria and Serum Glucose of Fasting Pregnant Women at the End of a Day in Ramadan. *Acta Medica Iranica, 42*(3), 209-212.

Barker, D. J. (1997). Fetal Nutrition and Cardiovascular Disease in Later Life. *British Medical Bulletin, 53*(1), 96-108.

Bateson, P., Barker, D., Clutton-Brock, T., Deb, D., D'Udine, B., Foley, R. A., . . . Lahr, M. M. (2004). Developmental Plasticity and Human Health. *Nature, 430*(6998), 419-421.

Chodick, G., Flash, S., Deoitch, Y., & Shalev, V. (2009). Seasonality in Birth Weight: Review of Global Patterns and Potential Causes. *Human Biology, 81*(4), 463-477.

Diener, E., Inglehart, R., & Tay, L. (2013). Theory and Validity of Life Satisfaction Scales. *Social Indicators Research, 112*(3), 497-527.

Doblhammer, G., & Vaupel, J. W. (2001). Lifespan Depends on Month of Birth. *Proceedings of the National Academy of Sciences, 98*(5), 2934-2939.

Frey, B. S., & Stutzer, A. (2002). What Can Economists Learn from Happiness Research? *Journal of Economic literature, 40*(2), 402-435.

Gluckman, P. D., Hanson, M. A., Buklijas, T., Low, F. M., & Beedle, A. S. (2009). Epigenetic Mechanisms That Underpin Metabolic and Cardiovascular Diseases. *Nature Reviews Endocrinology, 5*(7), 401-408.

Gluckman, P. D., Hanson, M. A., Cooper, C., & Thornburg, K. L. (2008). Effect of in Utero and Early-Life Conditions on Adult Health and Disease. *New England Journal of Medicine, 359*(1), 61-73.

Godfrey, K. M., & Barker, D. J. (2001). Fetal Programming and Adult Health. *Public Health Nutrition, 4*(2b), 611-624.

Greve, J., Schultz-Nielsen, M. L., & Tekin, E. (2015). *Fetal Malnutrition and Academic Success: Evidence from Muslim Immigrants in Denmark*. Retrieved from

Heckman, J. J. (2007). The Economics, Technology, and Neuroscience of Human Capability Formation. *Proceedings of the National Academy of Sciences, 104*(33), 13250-13255.

Idler, E. L., & Benyamini, Y. (1997). Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies. *Journal of Health and Social Behavior*, 21-37.

Joosoph, J., Abu, J., & Yu, S. (2004). A Survey of Fasting During Pregnancy. *Singapore Med J, 45*(12), 583-586.

Kahneman, D., & Krueger, A. B. (2006). Developments in the Measurement of Subjective Well-Being. *The Journal of Economic Perspectives, 20*(1), 3-24.

Kett, M. M., & Bertram, J. F. (2004). Nephron Endowment and Blood Pressure: What Do We Really Know? *Current Hypertension Reports, 6*(2), 133-139.

Kieler, H., Axelsson, O., Nilsson, S., & Waldenströ, U. (1995). The Length of Human Pregnancy as Calculated by Ultrasonographic Measurement of the Fetal Biparietal Diameter. *Ultrasound in Obstetrics and Gynecology, 6*(5), 353-357.

Krueger, A. B., & Schkade, D. A. (2008). The Reliability of Subjective Well-Being Measures. *Journal of Public Economics, 92*(8), 1833-1845.

Lim, I. (2010). Singapore's First Family Planning Services. Retrieved from http://eresources.nlb.gov.sg/infopedia/articles/SIP_1650_2010-02-26.html
Mackenzie, H. S., & Brenner, B. M. (1995). Fewer Nephrons at Birth: A Missing Link in the Etiology of Essential Hypertension? *American Journal of Kidney Diseases, 26*(1), 91-98.
Majid, M. F. (2015). The Persistent Effects of in Utero Nutrition Shocks over the Life Cycle: Evidence from Ramadan Fasting. *Journal of Development Economics, 117*, 48-57.
Majlis Ugama Islam Singapura. (2016). Women and Ramadan. In M. U. I. Singapura (Ed.). Singapore.
Ministry of Health. (2011). *National Health Survey 2010*. Singapore: Ministry of Health Retrieved from https://www.moh.gov.sg/content/moh_web/home/Publications/Reports/2011/national_health_survey2010.html.
National Library Board. (2016). Singapore Family Planning & Population Board Is Established. Retrieved from http://eresources.nlb.gov.sg/history/events/0b5c578a-f160-4958-97b0-d3e78ba38a05
Ravelli, A. C., van der Meulen, J. H., Michels, R., Osmond, C., Barker, D. J., Hales, C., & Bleker, O. P. (1998). Glucose Tolerance in Adults after Prenatal Exposure to Famine. *The Lancet, 351*(9097), 173-177.
Ravelli, A. C., van der Meulen, J. H., Osmond, C., Barker, D. J., & Bleker, O. P. (1999). Obesity at the Age of 50 Y in Men and Women Exposed to Famine Prenatally. *The American Journal of Clinical Nutrition, 70*(5), 811-816.
Rocha, R., & Soares, R. R. (2015). Water Scarcity and Birth Outcomes in the Brazilian Semiarid. *Journal of Development Economics, 112*, 72-91.
Roseboom, T. J., Van Der Meulen, J. H., Ravelli, A. C., Osmond, C., Barker, D. J., & Bleker, O. P. (2003). Perceived Health of Adults after Prenatal Exposure to the Dutch Famine. *Paediatric and Perinatal Epidemiology, 17*(4), 391-397.
Schultz-Nielsen, M. L., Tekin, E., & Greve, J. (2016). Labor Market Effects of Intrauterine Exposure to Nutritional Deficiency: Evidence from Administrative Data on Muslim Immigrants in Denmark. *Economics & Human Biology, 21*, 196-209.
Schwarz, N., & Strack, F. (1999). Reports of Subjective Well-Being: Judgmental Processes and Their Methodological Implications. *Well-being: The foundations of hedonic psychology, 7*, 61-84.
Seckl, J. R., & Holmes, M. C. (2007). Mechanisms of Disease: Glucocorticoids, Their Placental Metabolism and Fetal'programming'of Adult Pathophysiology. *Nature Clinical Practice Endocrinology & Metabolism, 3*(6), 479-488.
Singapore Department of Statistics. (2011). *Census of Population 2010, Statistical Release 1, Demographic Characteristics, Education, Language and Religion*. Retrieved from
Singapore Department of Statistics. (2016). Singapore Residents by Age Group, Ethnic Group and Gender, End June, Annual. Retrieved from https://data.gov.sg/dataset/resident-population-by-ethnicity-gender-and-age-group?view_id=2f273a73-de27-4242-a602-8165789f96d3&resource_id=d5d35678-cd15-4f43-b347-d671d1870655
Stein, A. D., Kahn, H. S., Rundle, A., Zybert, P. A., van der Pal–de Bruin, K., & Lumey, L. (2007). Anthropometric Measures in Middle Age after Exposure to Famine During Gestation: Evidence from the Dutch Famine. *The American Journal of Clinical Nutrition, 85*(3), 869-876.
Stein, A. D., Zybert, P. A., Van der Pal-de Bruin, K., & Lumey, L. (2006). Exposure to Famine During Gestation, Size at Birth, and Blood Pressure at Age 59 Y: Evidence from the Dutch Famine. *European Journal of Epidemiology, 21*(10), 759-765.

Susser, E., & Ananth, C. V. (2013). Invited Commentary: Is Prenatal Fasting During Ramadan Related to Adult Health Outcomes? A Novel and Important Question for Epidemiology. *American Journal of Epidemiology, 177*(8), 737-740.

Van Ewijk, R. (2011). Long-Term Health Effects on the Next Generation of Ramadan Fasting During Pregnancy. *Journal of Health Economics*, *30*(6), 1246-1260.

Van Ewijk, R., Painter, R. C., & Roseboom, T. J. (2013). Associations of Prenatal Exposure to Ramadan with Small Stature and Thinness in Adulthood: Results from a Large Indonesian Population–Based Study. *American Journal of Epidemiology, 177*(8), 729-736.

Vo, T., & Hardy, D. B. (2012). Molecular Mechanisms Underlying the Fetal Programming of Adult Disease. *Journal of Cell Communication and Signaling*, *6*(3), 139-153.

Wu, S., Wang, R., Zhao, Y., Ma, X., Wu, M., Yan, X., & He, J. (2013). The Relationship between Self-Rated Health and Objective Health Status: A Population-Based Study. *BMC Public Health*, *13*(1), 1.

Yamauchi, F. (2012). Prenatal Seasonality, Child Growth, and Schooling Investments: Evidence from Rural Indonesia. *Journal of Development Studies, 48*(9), 1323-1341.

Table 3.1: Summary statistics for basic demographics and major outcome variables

Variable	Muslims					Non-Muslims				
	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max
Basic Demographics										
Male	1,480	0.47	0.50	0	1	14,046	0.48	0.50	0	1
Age	1,480	58.13	5.92	40	76	14,046	59.09	6.04	40	83
At least some sec educ	1,447	0.69	0.46	0	1	13,886	0.73	0.44	0	1
At least some post-sec educ	1,447	0.18	0.38	0	1	13,886	0.32	0.47	0	1
Dummy variables for exposure to Ramadan in-utero										
Ever exposed	1,480	0.84	0.37	0	1	14,046	0.85	0.36	0	1
Exposed 1st trimester	1,480	0.36	0.48	0	1	14,046	0.34	0.47	0	1
Exposed 2nd trimester	1,480	0.31	0.46	0	1	14,046	0.34	0.47	0	1
Exposed 3rd trimester	1,480	0.32	0.47	0	1	14,046	0.33	0.47	0	1
Subjective well-being variables										
Life satisfaction	1,442	3.82	0.58	1	5	13,852	3.63	0.65	1	5
Feeling happy (s011)	867	4.20	0.85	1	6	10,093	3.81	0.89	1	6
Feeling sad (s013)	865	3.92	0.79	1	5	10,092	3.92	0.69	1	5
Feeling worn out (s010)	866	3.65	0.97	1	6	10,096	3.80	0.93	1	6
Diff with sleep (s012)	867	3.64	0.82	1	5	10,092	3.69	0.77	1	5
Mental well-being index	864	0.0010	3.13	-12.39	7.58	10,092	0.00011	3.17	-13.40	8.11
Social/Family satisfaction	1,442	3.98	0.52	1.41	5	13,847	3.74	0.58	1	5
Daily activities satisfaction	1,047	3.72	0.58	1.25	5	11,120	3.53	0.63	1	5
Economic satisfaction	1,440	3.42	0.70	1	5	13,833	3.31	0.73	1	5
Health satisfaction	1,441	3.61	0.68	1	5	13,839	3.44	0.72	1	5
Health condition	1,441	2.88	0.77	1	5	13,843	2.75	0.78	1	5

Table 1: Summary statistics for basic demographics and major outcome variables (continued)

Variable	Muslims					Non-Muslims				
	N	Mean	SD	Min	Max	N	Mean	SD	Min	Max
Health-related variables										
Cardiovascular conditions	1,480	0.42	0.49	0	1	14,046	0.39	0.49	0	1
Hypertension	1,480	0.36	0.48	0	1	14,046	0.35	0.48	0	1
Diabetes	1,480	0.26	0.44	0	1	14,046	0.16	0.37	0	1
Height (m)	481	1.60	0.10	1.10	2.00	7,415	1.62	0.09	1.05	2.00
Weight (kg)	481	68.86	13.54	39	136	7,415	63.23	13.39	32	200
BMI	481	26.91	5.29	16.26	55.56	7,415	23.94	4.55	12.89	75.65

Notes:

[1] Life, social/family life, daily activities, economic and health satisfaction, as well as health condition are measured on a scale of 1 – 5. 1 indicates the worst option and 5 indicating the best. Since individuals can respond to these questions in more than one wave, their responses for each outcome are averaged across the number of waves in which they responded.

[2]Affect index is the sum of normalised affect measures, with each measure being averaged across the number of waves in which respondents participated. A higher value reflects a better state of well-being.

[3]Cardiovascular conditions include hypertension, heart problems, and stroke.

Table 3.2: Effect of in-utero exposure to Ramadan on overall life satisfaction and mental well-being index

	Life satisfaction			Mental well-being index		
	Muslims	Non-Muslims	DiD	Muslims	Non-Muslims	DiD
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Panel A						
Trimester 1	-0.0680	0.000174	-0.0681	-0.273	-0.0501	-0.223
	(0.0451)	(0.0156)	(0.0475)	(0.328)	(0.0898)	(0.338)
Trimester 2	-0.0779**	0.00776	-0.0857**	-0.649**	0.0485	-0.698**
	(0.0392)	(0.0137)	(0.0413)	(0.273)	(0.0777)	(0.281)
Trimester 3	-0.0407	-0.00302	-0.0377	-0.254	-0.00249	-0.251
	(0.0442)	(0.0155)	(0.0465)	(0.298)	(0.0880)	(0.308)
Panel B						
Exposed	-0.103*	0.0109	-0.114*	-0.803**	-0.131	-0.673*
	(0.0566)	(0.0190)	(0.0594)	(0.384)	(0.104)	(0.395)
Observations	1,442	13,852	15,294	864	10,092	10,956
R-squared	0.013	0.003	0.011	0.039	0.006	0.009
Mean	3.82	3.63		0.0010	0.00011	
S.D.	0.58	0.65		3.13	3.17	

Notes:

[1] Standard errors clustered at the household level in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.

[2] The first dependent variable is the self-reported life satisfaction of each respondent, averaged across the number of waves in which they responded. It is measured on a scale of 1 – 5, with 1 indicating "very dissatisfied" and 5 indicating "very satisfied". The second dependent variable is the sum of normalised measures of mental well-being, with each measure being averaged across the number of waves in which respondents participated. A higher value reflects a better state of well-being.

[3] Panels A and B show results from regressions using different Ramadan exposure dummies.

[4] Results shown in the Muslim and non-Muslim columns are coefficients of Ramadan exposure variables, in a regression that also controls for birth month dummies, gender, age, and age-squared, as well as a buffer dummy that takes value 1 if the predicted conception date occurs 21 days or less after the end of Ramadan.

[5] Results shown in the DiD column are coefficients of the Ramadan exposure variable interacted with a Muslim indicator, in a regression similar to the one described in footnote (3), but adjusted for the DiD specification.

Table 3.3: Effect of in-utero exposure to Ramadan on social/family life satisfaction and satisfaction with daily activities

	Social/Family satisfaction			Daily activities satisfaction		
	Muslims	Non-Muslims	DiD	Muslims	Non-Muslims	DiD
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Panel A						
Trimester 1	-0.0289	-0.0137	-0.0152	-0.0396	-0.00159	-0.0380
	(0.0400)	(0.0140)	(0.0422)	(0.0518)	(0.0169)	(0.0542)
Trimester 2	-0.0785**	-0.00839	-0.0701*	-0.0971**	0.0168	-0.114**
	(0.0367)	(0.0124)	(0.0386)	(0.0483)	(0.0149)	(0.0502)
Trimester 3	-0.0459	-0.0222	-0.0237	-0.0465	0.000366	-0.0469
	(0.0416)	(0.0139)	(0.0436)	(0.0523)	(0.0166)	(0.0545)
Panel B						
Exposed	-0.0705	-0.0141	-0.0564	-0.0780	-0.0158	-0.0623
	(0.0484)	(0.0168)	(0.0510)	(0.0581)	(0.0200)	(0.0611)
Observations	1,442	13,847	15,289	1,047	11,120	12,167
R-squared	0.017	0.002	0.018	0.016	0.006	0.014
Mean	3.98	3.74		3.72	3.53	
S.D.	0.52	0.58		0.58	0.63	

Notes:
[1] Standard errors clustered at the household level in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.
[2] The dependent variables are respondents' self-reported satisfaction with social contacts and family life and daily activities respectively, averaged across the number of waves in which they responded. They are measured on a scale of 1 – 5, with 1 indicating "very dissatisfied" and 5 indicating "very satisfied".
[3] Panels A and B show results from regressions using different Ramadan exposure dummies.
[4] Results shown in the Muslim and non-Muslim columns are coefficients of Ramadan exposure variables, in a regression that also controls for birth month dummies, gender, age, and age-squared, as well as a buffer dummy that takes value 1 if the predicted conception date occurs 21 days or less after the end of Ramadan.
[5] Results shown in the DiD column are coefficients of the Ramadan exposure variable interacted with a Muslim indicator, in a regression similar to the one described in footnote (3), but adjusted for the DiD specification.

Table 3.4: Effect of in-utero exposure to Ramadan on economic satisfaction

	Economic satisfaction		
	Muslims	Non-Muslims	DiD
Variables	(1)	(2)	(3)
Panel A			
Trimester 1	-0.0913	0.00336	-0.0947
	(0.0556)	(0.0178)	(0.0581)
Trimester 2	-0.0839*	0.0208	-0.105**
	(0.0476)	(0.0154)	(0.0497)
Trimester 3	-0.0356	0.00796	-0.0436
	(0.0528)	(0.0174)	(0.0553)
Panel B			
Exposed	-0.150**	0.0166	-0.167**
	(0.0662)	(0.0210)	(0.0692)
Observations	1,440	13,833	15,273
R-squared	0.011	0.005	0.007
Mean	3.42	3.31	
S.D.	0.70	0.73	

Notes:
[1] Standard errors clustered at the household level in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.
[2] Dependent variable is the self-reported economic satisfaction of each respondent, averaged across the number of waves in which they responded. It is measured on a scale of 1 – 5, with 1 indicating "very dissatisfied" and 5 indicating "very satisfied"
[3] Panels A and B show results from regressions using different Ramadan exposure dummies.
[4] Results shown in the Muslim and non-Muslim columns are coefficients of Ramadan exposure variables, in a regression that also controls for birth month dummies, gender, age, and age-squared, as well as a buffer dummy that takes value 1 if the predicted conception date occurs 21 days or less after the end of Ramadan.
[5] Results shown in the DiD column are coefficients of the Ramadan exposure variable interacted with a Muslim indicator, in a regression similar to the one described in footnote (3), but adjusted for the DiD specification.

Table 3.5: Effect of in-utero exposure to Ramadan on general self-rated health

	Health condition			Health satisfaction		
	Muslims	Non-Muslims	DiD	Muslims	Non-Muslims	DiD
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Panel A						
Trimester 1	-0.00132	-0.00906	0.00774	-0.0466	-0.00528	-0.0413
	(0.0614)	(0.0190)	(0.0640)	(0.0522)	(0.0174)	(0.0547)
Trimester 2	-0.115**	0.0171	-0.132**	-0.100**	0.0172	-0.117**
	(0.0526)	(0.0161)	(0.0547)	(0.0448)	(0.0149)	(0.0470)
Trimester 3	0.0158	-0.00261	0.0184	0.0138	-0.0174	0.0311
	(0.0590)	(0.0184)	(0.0616)	(0.0494)	(0.0172)	(0.0521)
Panel B						
Exposed	-0.0518	-0.00274	-0.0491	-0.0628	-0.0165	-0.0462
	(0.0772)	(0.0227)	(0.0802)	(0.0601)	(0.0208)	(0.0634)
Observations	1,441	13,843	15,284	1,441	13,839	15,280
R-squared	0.032	0.012	0.016	0.034	0.006	0.013
Mean	2.88	2.75		3.61	3.44	
S.D.	0.77	0.78		0.68	0.72	

Notes:
[1] Standard errors clustered at the household level in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.
[2] Dependent variables are the self-reported health satisfaction and health condition of each respondent, averaged across the number of waves in which they responded. Both are measured on a scale of 1 – 5, with 1 indicating the worst option (e.g. poor for health condition) and 5 indicating the best (e.g. excellent for health condition)
[3] Panels A and B show results from regressions using different Ramadan exposure dummies.
[4] Results shown in the Muslim and non-Muslim columns are coefficients of Ramadan exposure variables, in a regression that also controls for birth month dummies, gender, age, and age-squared, as well as a buffer dummy that takes value 1 if the predicted conception date occurs 21 days or less after the end of Ramadan.
[5] Results shown in the DiD column are coefficients of the Ramadan exposure variable interacted with a Muslim indicator, in a regression similar to the one described in footnote (3), but adjusted for the DiD specification.

Table 3.6: Effect of in-utero exposure to Ramadan on cardiovascular conditions and diabetes

	Cardiovascular conditions			Diabetes		
	Muslims	Non-Muslims	DiD	Muslims	Non-Muslims	DiD
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Panel A						
Trimester 1	-0.00281	0.00585	-0.00866	-0.0343	-0.0113	-0.0230
	(0.0376)	(0.0114)	(0.0390)	(0.0349)	(0.00888)	(0.0358)
Trimester 2	0.0925***	-0.00779	0.100***	0.00566	-0.0183**	0.0240
	(0.0328)	(0.00980)	(0.0341)	(0.0292)	(0.00761)	(0.0301)
Trimester 3	-0.0332	0.00708	-0.0403	-0.00653	-0.0136	0.00709
	(0.0365)	(0.0112)	(0.0380)	(0.0334)	(0.00880)	(0.0344)
Panel B						
Exposed	0.0394	0.00511	0.0343	-0.00665	-0.0195*	0.0128
	(0.0460)	(0.0139)	(0.0478)	(0.0423)	(0.0111)	(0.0435)
Observations	1,480	14,046	15,526	1,480	14,046	15,526
R-squared	0.056	0.054	0.054	0.040	0.025	0.033
Mean	0.42	0.39		0.26	0.16	
S.D.	0.49	0.49		0.44	0.37	

Notes:
[1] Standard errors clustered at the household level in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.
[2] Dependent variables are dummy variables for whether respondent indicated that he/she had been diagnosed with the medical condition by a doctor. Cardiovascular diseases include hypertension, heart problems, and stroke.
[3] Panels A and B show results from regressions using different Ramadan exposure dummies.
[4] Results shown in the Muslim and non-Muslim columns are coefficients of Ramadan exposure variables, in a regression that also controls for birth month dummies, gender, age, and age-squared, as well as a buffer dummy that takes value 1 if the predicted conception date occurs 21 days or less after the end of Ramadan.
[5] Results shown in the DiD column are coefficients of the Ramadan exposure variable interacted with a Muslim indicator, in a regression similar to the one described in footnote (3), but adjusted for the DiD specification.

Table 3.7: Effect of in-utero exposure to Ramadan on body mass index

	Sample: all			Sample: women only			Sample: men only		
	Muslims	Non-Muslims	DiD	Muslims	Non-Muslims	DiD	Muslims	Non-Muslims	DiD
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A									
Trimester 1	0.338	-0.0743	0.412	0.829	-0.0719	0.900	-0.396	-0.0868	-0.309
	(0.643)	(0.150)	(0.649)	(1.092)	(0.204)	(1.078)	(0.907)	(0.225)	(0.902)
Trimester 2	0.526	0.104	0.421	2.086*	0.149	1.937*	-0.942	0.0157	-0.958
	(0.648)	(0.134)	(0.650)	(1.119)	(0.189)	(1.102)	(0.760)	(0.185)	(0.754)
Trimester 3	0.125	-0.228	0.353	1.096	-0.333	1.429	-0.943	-0.147	-0.795
	(0.722)	(0.150)	(0.723)	(1.092)	(0.215)	(1.080)	(0.869)	(0.205)	(0.861)
Panel B									
Exposed	0.511	-0.264	0.776	1.405	-0.391	1.796	-0.442	-0.155	-0.287
	(0.756)	(0.189)	(0.767)	(1.178)	(0.255)	(1.174)	(0.944)	(0.277)	(0.954)
Observations	481	7,415	7,896	263	3,940	4,203	218	3,475	3,693
R-squared	0.044	0.016	0.041	0.076	0.005	0.055	0.084	0.006	0.019
Mean	26.91	23.94		27.51	23.45		26.17	24.5	
S.D.	5.29	4.55		5.59	4.56		4.82	4.49	

Notes:
[1] Standard errors clustered at the household level in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.
[2] Dependent variable is body mass index (weight in kg/(height in m)^2). Headings at the top of each column indicate the sub-sample used.
[3] Panels A and B show results from regressions using different Ramadan exposure dummies.
[4] Results shown in the Muslim and non-Muslim columns are coefficients of Ramadan exposure variables, in a regression that also controls for birth month dummies, gender, age, and age-squared, as well as a buffer dummy that takes value 1 if the predicted conception date occurs 21 days or less after the end of Ramadan.
[5] Results shown in the DiD column are coefficients of the Ramadan exposure variable interacted with a Muslim indicator, in a regression similar to the one described in footnote (3), but adjusted for the DiD specification.

Table 3.8: Parental characteristics and in-utero exposure to Ramadan

(Sample contains only Muslims)

	Father			Mother		
	Whether alive	Current age if alive	Age when passed away	Whether alive	Current age if alive	Age when passed away
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Panel A						
Trimester 1	-0.00251	0.0615	0.730	0.0232	1.058	-2.515*
	(0.0244)	(1.700)	(1.275)	(0.0361)	(0.817)	(1.393)
Trimester 2	-0.000523	1.236	1.046	-0.0235	0.0502	-0.289
	(0.0213)	(1.362)	(1.089)	(0.0309)	(0.657)	(1.188)
Trimester 3	-0.00970	-0.518	1.425	0.00727	0.688	2.340*
	(0.0229)	(1.709)	(1.136)	(0.0346)	(0.773)	(1.301)
Panel B						
Exposed	-0.0325	0.0153	2.649*	-0.0203	0.459	1.043
	(0.0294)	(2.363)	(1.483)	(0.0450)	(0.928)	(1.644)
Observations	1,431	160	1,191	1,434	479	884
R-squared	0.074	0.388	0.013	0.081	0.290	0.042

Notes:

[1] Standard errors clustered at the household level in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.

[2] Dependent variables are indicated at the top of each column. Variables in the left and right panels refer to data related to the father and the mother of the respondent respectively.

[3] Panels A and B show results from regressions using different Ramadan exposure dummies.

[4] Results shown are coefficients of Ramadan exposure variables, in a regression that also controls for birth month dummies, gender, age, and age-squared, as well as an buffer dummy that takes value 1 if the predicted conception date occurs 21 days or less after the end of Ramadan.

Table 3.9: Robustness check for Family Planning Policy (born before 1966)
(Sample contains only Muslims)

Variables	Life satisfaction (1)	Affect index (2)	Social/ Family satisfaction 3)	Daily activities satisfaction (4)	Economic satisfaction (5)	Health condition (6)	Health satisfaction (7)	Cardiovas-cular conditions (8)	BMI (Women) (9)	Weight (Men) (10)	Height (Men) (11)
Panel A											
Trimester 1	-0.0628	-0.330	-0.0200	-0.0347	-0.0857	-0.00187	-0.0405	-0.0126	1.037	-4.716**	-0.0370**
	(0.0464)	(0.342)	(0.0410)	(0.0533)	(0.0567)	(0.0637)	(0.0537)	(0.0383)	(1.194)	(2.362)	(0.0183)
Trimester 2	-0.0784*	-0.671**	-0.0742**	-0.108**	-0.0892*	-0.123**	-0.106**	0.0950***	1.946	-3.145	-0.00571
	(0.0402)	(0.287)	(0.0377)	(0.0502)	(0.0489)	(0.0543)	(0.0460)	(0.0334)	(1.217)	(2.281)	(0.0155)
Trimester 3	-0.0360	-0.225	-0.0444	-0.0339	-0.0111	0.0213	0.0181	-0.0370	1.146	-3.865	-0.0118
	(0.0449)	(0.313)	(0.0425)	(0.0537)	(0.0537)	(0.0606)	(0.0508)	(0.0377)	(1.215)	(2.558)	(0.0174)
Panel B											
Exposed	-0.0955	-0.867**	-0.0597	-0.0666	-0.128*	-0.0533	-0.0604	0.0321	0.883	-5.778*	-0.0394
	(0.0594)	(0.422)	(0.0505)	(0.0616)	(0.0688)	(0.0819)	(0.0631)	(0.0482)	(1.418)	(3.272)	(0.0276)
Observation	1,404	832	1,404	1,012	1,402	1,403	1,403	1,442	241	215	215
R-squared	0.012	0.040	0.016	0.016	0.014	0.033	0.035	0.057	0.080	0.096	0.108

Notes:
[1] Standard errors clustered at the household level in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.
[2] Dependent variables are indicated at the top of each column.
[3] Panels A and B show results from regressions using different Ramadan exposure dummies.
[4] Results shown are coefficients of Ramadan exposure variables, in a regression that also controls for birth month dummies, gender, age, age-squared, as well as a buffer dummy that takes value 1 if the predicted conception date occurs 21 days or less after the end of Ramadan.

Table 3.10: Relationship between probability of being in-utero during Ramadan and age

	Exposed in Trimester 1			Exposed in Trimester 2			Exposed in Trimester 3			Exposed		
	Muslims	Non-Muslims	DiD	Muslims	Non-Muslims	DiD	Muslims	Non-Muslims	DiD	Muslims	Non-Muslims	DiD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Coefficient (SE)	-0.00531**	-0.00314***	-0.00217	0.00483**	0.00230***	0.00253	0.00541**	0.00194***	0.00348	0.00228	-0.000317	0.00260
	(0.00226)	(0.000725)	(0.00237)	(0.00217)	(0.000721)	(0.00228)	(0.00231)	(0.000721)	(0.00241)	(0.00175)	(0.000527)	(0.00182)
Observations	1,480	14,046	15,526	1,480	14,046	15,526	1,480	14,046	15,526	1,480	14,046	15,526

Notes:
[1] Robust standard errors in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.
[2] Dependent variables are the indicator variables for in-utero Ramadan exposure. They take on value 1 if Ramadan was likely to overlap with a particular period of pregnancy, and 0 otherwise.
[3] For each dependent variable, 3 regression models are estimated. In the first and second models, the dependent variable is regressed on age and calendar-month-of-birth fixed effects, for the Muslim and non-Muslim sample respectively. Results shown in the Muslim and non-Muslim columns are coefficients of the age variable.
[4] In the third model (DiD), the dependent variable is regressed on age and calendar-month-of-birth fixed effects, as well as age interacted with a Muslim indicator and a Muslim indicator. Results shown in the DiD column are coefficients of the age variable interacted with a Muslim indicator.

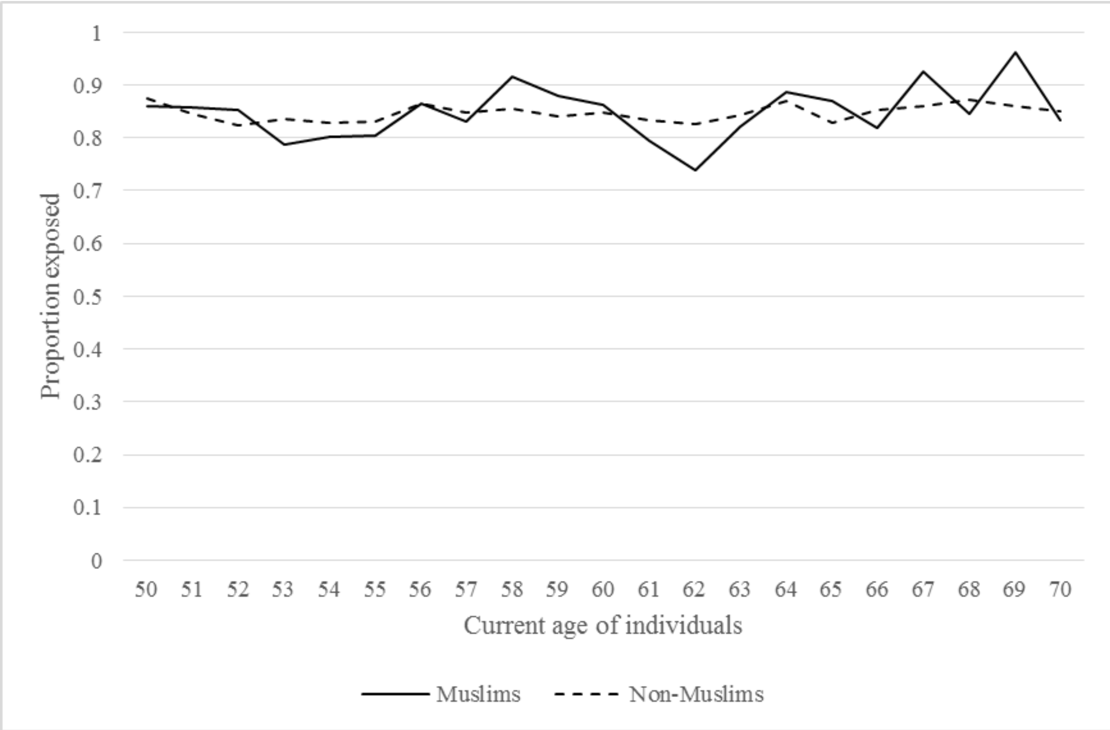


Figure 3.1: Proportion of Muslims and non-Muslims who were in-utero during Ramadan, by age

Appendix: Additional Tables

Table A 1: Effect of in-utero exposure to Ramadan on time spent feeling happy and problem with feeling sad/low/depressed

	Feeling happy			Feeling sad		
	Muslims	Non-Muslims	DiD	Muslims	Non-Muslims	DiD
Variables	(1)	(2)	(3)	(4)	(5)	(6)
Panel A						
Trimester 1	-0.153*	0.00823	-0.162*	-0.0224	-0.0108	-0.0116
	(0.0874)	(0.0252)	(0.0903)	(0.0774)	(0.0194)	(0.0792)
Trimester 2	-0.163**	0.0199	-0.183**	-0.140**	0.00150	-0.141**
	(0.0777)	(0.0218)	(0.0800)	(0.0669)	(0.0171)	(0.0684)
Trimester 3	-0.158*	0.00353	-0.161*	-0.0810	0.00591	-0.0869
	(0.0815)	(0.0246)	(0.0844)	(0.0753)	(0.0192)	(0.0770)
Panel B						
Exposed	-0.184*	-0.0292	-0.154	-0.119	-0.0247	-0.0940
	(0.104)	(0.0294)	(0.107)	(0.0872)	(0.0227)	(0.0894)
Observations	867	10,093	10,960	865	10,092	10,957
R-squared	0.026	0.004	0.019	0.034	0.005	0.008
Mean	4.20	3.81		3.92	3.92	
S.D.	0.85	0.89		0.79	0.69	

Notes:

[1] Standard errors clustered at the household level in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.

[2] The first dependent variable is the amount of time respondent felt happy, averaged across the number of waves in which they responded. It is measured on a scale of 1 – 6, with 1 indicating "none of the time" and 6 indicating "all of the time". The second dependent variable is respondent's problem with feeling sad, low, or depressed, averaged across the number of waves in which they responded. It is measured on a scale of 1 – 5, with 1 indicating "extreme" and 5 indicating "none". For both variables, a higher value indicates a better state of well-being, as we have flipped the scale where relevant.

[3] Panels A and B show results from regressions using different Ramadan exposure dummies.

[4] Results shown in the Muslim and non-Muslim columns are coefficients of Ramadan exposure variables, in a regression that also controls for birth month dummies, gender, age, and age-squared, as well as a buffer dummy that takes value 1 if the predicted conception date occurs 21 days or less after the end of Ramadan.

[5] Results shown in the DiD column are coefficients of the Ramadan exposure variable interacted with a Muslim indicator, in a regression similar to the one described in footnote (3), but adjusted for the DiD specification.

Table A 2: Effect of in-utero exposure to Ramadan on time spent feeling worn out and difficulty with sleeping

	Feeling worn out			Difficulty with sleeping		
	Muslims	Non-Muslims	DiD	Muslims	Non-Muslims	DiD
Variables	(1)	(2)	(3)	(4)	(5)	(6)
<u>Panel A</u>						
Trimester 1	-0.0525	-0.0104	-0.0421	-0.00170	-0.0247	0.0230
	(0.103)	(0.0260)	(0.106)	(0.0880)	(0.0218)	(0.0899)
Trimester 2	-0.0991	0.0258	-0.125	-0.146**	-0.00325	-0.143**
	(0.0872)	(0.0228)	(0.0894)	(0.0710)	(0.0188)	(0.0728)
Trimester 3	0.0667	0.0167	0.0500	-0.0319	-0.0248	-0.00703
	(0.0984)	(0.0253)	(0.101)	(0.0790)	(0.0212)	(0.0811)
<u>Panel B</u>						
Exposed	-0.204	0.00290	-0.207	-0.183*	-0.0502*	-0.133
	(0.134)	(0.0304)	(0.137)	(0.103)	(0.0260)	(0.105)
Observations	866	10,096	10,962	867	10,092	10,959
R-squared	0.042	0.014	0.018	0.027	0.005	0.007
Mean	3.65	3.8		3.64	3.69	
S.D.	0.97	0.93		0.82	0.77	

Notes:
[1] Standard errors clustered at the household level in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.
[2] The first dependent variable is the amount of time respondent felt worn out, averaged across the number of waves in which they responded. It is measured on a scale of 1 – 6, with 1 indicating "all of the time" and 6 indicating "none of the time". The second dependent variable is respondent's difficulty with sleeping, averaged across the number of waves in which they responded. It is measured on a scale of 1 – 5, with 1 indicating "extreme" and 5 indicating "none". For both variables, a higher value indicates a better state of well-being, as we have flipped the scale where relevant.
[3] Panels A and B show results from regressions using different Ramadan exposure dummies.
[4] Results shown in the Muslim and non-Muslim columns are coefficients of Ramadan exposure variables, in a regression that also controls for birth month dummies, gender, age, and age-squared, as well as a buffer dummy that takes value 1 if the predicted conception date occurs 21 days or less after the end of Ramadan.
[5] Results shown in the DiD column are coefficients of the Ramadan exposure variable interacted with a Muslim indicator, in a regression similar to the one described in footnote (3), but adjusted for the DiD specification.

Table A 3: Effect of in-utero exposure to Ramadan on cardiovascular diseases

	Hypertension			Heart Problems			Stroke		
	Muslims	Non-Muslims	DiD	Muslims	Non-Muslims	DiD	Muslims	Non-Muslims	DiD
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A									
Trimester 1	-0.0228	0.00123	-0.0240	-0.0134	-0.00358	-0.00984	0.000744	0.00280	-0.00206
	(0.0375)	(0.0112)	(0.0389)	(0.0288)	(0.00720)	(0.0295)	(0.0137)	(0.00364)	(0.0141)
Trimester 2	0.0701**	-0.0114	0.0815**	0.0400*	0.00671	0.0333	-0.000809	-0.00277	0.00196
	(0.0322)	(0.00971)	(0.0335)	(0.0242)	(0.00628)	(0.0248)	(0.0111)	(0.00306)	(0.0114)
Trimester 3	-0.0276	0.00279	-0.0304	-0.0217	-0.000596	-0.0211	-0.000963	0.000281	-0.00124
	(0.0356)	(0.0110)	(0.0371)	(0.0262)	(0.00710)	(0.0270)	(0.0115)	(0.00342)	(0.0119)
Panel B									
Exposed	0.0255	-0.00693	0.0325	-0.0150	-0.00322	-0.0117	0.0128	0.000402	0.0124
	(0.0449)	(0.0138)	(0.0467)	(0.0343)	(0.00890)	(0.0353)	(0.0140)	(0.00427)	(0.0145)
Observations	1,480	14,046	15,526	1,480	14,046	15,526	1,480	14,046	15,526
R-squared	0.049	0.044	0.044	0.052	0.028	0.032	0.033	0.007	0.011
Mean	0.36	0.35		0.13	0.10		0.028	0.022	
S.D.	0.48	0.48		0.34	0.30		0.17	0.15	

Notes:
[1] Standard errors clustered at the household level in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.
[2] Dependent variables are dummy variables for whether respondent indicated that he/she had been diagnosed with the medical condition by a doctor.
[3] Panels A and B show results from regressions using different Ramadan exposure dummies.
[4] Results shown in the Muslim and non-Muslim columns are coefficients of Ramadan exposure variables, in a regression that also controls for birth month dummies, gender, age, and age-squared, as well as an buffer dummy that takes value 1 if the predicted conception date occurs 21 days or less after the end of Ramadan.
[5] Results shown in the DiD column are coefficients of the Ramadan exposure variable interacted with a Muslim indicator, in a regression similar to the one described in footnote (3), but adjusted for the DiD specification.

Table A 4: Effect of in-utero exposure to Ramadan on other illnesses

	Psychiatric problems			Arthritis			Cancer		
	Muslims	Non-Muslims	DiD	Muslims	Non-Muslims	DiD	Muslims	Non-Muslims	DiD
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A									
Trimester 1	0.00460	0.00315	0.00145	0.00262	-0.00320	0.00582	0.0199	0.00673	0.0132
	(0.00879)	(0.00365)	(0.00948)	(0.0211)	(0.00828)	(0.0226)	(0.0126)	(0.00487)	(0.0135)
Trimester 2	0.0115	0.000878	0.0107	0.0360*	0.00129	0.0347*	0.00530	-0.000319	0.00562
	(0.00755)	(0.00332)	(0.00821)	(0.0193)	(0.00714)	(0.0205)	(0.0123)	(0.00432)	(0.0129)
Trimester 3	0.0107	0.00476	0.00590	0.0178	0.00164	0.0162	0.00245	-0.00123	0.00368
	(0.00902)	(0.00367)	(0.00966)	(0.0221)	(0.00811)	(0.0234)	(0.0132)	(0.00485)	(0.0140)
Panel B									
Exposed	0.0140	0.00340	0.0106	0.0289	0.00799	0.0209	0.0121	-0.000559	0.0127
	(0.00917)	(0.00427)	(0.0101)	(0.0249)	(0.00981)	(0.0267)	(0.0147)	(0.00603)	(0.0158)
Observations	1,480	14,046	15,526	1,480	14,046	15,526	1,480	14,046	15,526
R-squared	0.020	0.002	0.003	0.022	0.023	0.024	0.014	0.006	0.006
Mean	0.014	0.025		0.091	0.137		0.034	0.042	
S.D.	0.12	0.16		0.29	0.34		0.18	0.20	

Notes:

[1] Standard errors clustered at the household level in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.

[2] Dependent variables are dummy variables for whether respondent indicated that he/she had been diagnosed with the medical condition by a doctor.

[3] Panels A and B show results from regressions using different Ramadan exposure dummies.

[4] Results shown in the Muslim and non-Muslim columns are coefficients of Ramadan exposure variables, in a regression that also controls for birth month dummies, gender, age, and age-squared, as well as an buffer dummy that takes value 1 if the predicted conception date occurs 21 days or less after the end of Ramadan.

[5] Results shown in the DiD column are coefficients of the Ramadan exposure variable interacted with a Muslim indicator, in a regression similar to the one described in footnote (3), but adjusted for the DiD specification.

Table A 5: Effect of in-utero exposure to Ramadan on being overweight (BMI>=23)

	Sample: all			Sample: women only			Sample: men only		
	Muslims	Non-Muslims	DiD	Muslims	Non-Muslims	DiD	Muslims	Non-Muslims	DiD
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A									
Trimester 1	0.00110	-0.00373	0.00483	0.0700	0.00181	0.0682	-0.0794	-0.0112	-0.0682
	(0.0558)	(0.0165)	(0.0573)	(0.0815)	(0.0228)	(0.0823)	(0.0866)	(0.0236)	(0.0866)
Trimester 2	0.0217	0.00824	0.0135	0.0916	0.00775	0.0839	-0.0500	0.00485	-0.0549
	(0.0543)	(0.0142)	(0.0552)	(0.0796)	(0.0196)	(0.0797)	(0.0796)	(0.0203)	(0.0793)
Trimester 3	0.0111	-0.0157	0.0269	0.0904	-0.0301	0.121*	-0.0970	-0.00694	-0.0900
	(0.0547)	(0.0160)	(0.0561)	(0.0714)	(0.0222)	(0.0726)	(0.0858)	(0.0231)	(0.0858)
Panel B									
Exposed	0.0496	-0.0163	0.0659	0.181*	-0.0125	0.194*	-0.0918	-0.0278	-0.0640
	(0.0703)	(0.0194)	(0.0718)	(0.102)	(0.0267)	(0.103)	(0.0969)	(0.0282)	(0.0979)
Observations	481	7,415	7,896	263	3,940	4,203	218	3,475	3,693
R-squared	0.019	0.019	0.032	0.072	0.007	0.033	0.047	0.006	0.013
Mean	0.78	0.55		0.80	0.49		0.76	0.62	
S.D.	0.41	0.50		0.40	0.50		0.43	0.49	

Notes:
[1] Standard errors clustered at the household level in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.
[2] Dependent variable is whether the individual is overweight, as defined by a body mass index (weight in kg/(height in m)^2) that is greater than or equal to 23. Headings at the top of each column indicate the sub-sample used.
[3] Panels A and B show results from regressions using different Ramadan exposure dummies.
[4] Results shown in the Muslim and non-Muslim columns are coefficients of Ramadan exposure variables, in a regression that also controls for birth month dummies, gender, age, and age-squared, as well as a buffer dummy that takes value 1 if the predicted conception date occurs 21 days or less after the end of Ramadan.
[5] Results shown in the DiD column are coefficients of the Ramadan exposure variable interacted with a Muslim indicator, in a regression similar to the one described in footnote (3), but adjusted for the DiD specification.

Table A 6: Effect of in-utero exposure to Ramadan on weight

	Sample: all			Sample: women only			Sample: men only		
	Muslims	Non-Muslims	DiD	Muslims	Non-Muslims	DiD	Muslims	Non-Muslims	DiD
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A									
Trimester 1	-1.083	0.338	-1.421	1.942	0.137	1.805	-4.807**	0.537	-5.344**
	(1.743)	(0.405)	(1.760)	(2.559)	(0.511)	(2.533)	(2.375)	(0.634)	(2.373)
Trimester 2	0.0509	0.179	-0.128	3.775	0.356	3.418	-2.766	-0.0981	-2.668
	(1.741)	(0.343)	(1.744)	(2.783)	(0.435)	(2.732)	(2.265)	(0.536)	(2.245)
Trimester 3	0.501	-0.311	0.812	4.163	-0.646	4.809*	-3.889	-0.0960	-3.793
	(1.938)	(0.387)	(1.941)	(2.813)	(0.512)	(2.774)	(2.520)	(0.585)	(2.495)
Panel B									
Exposed	-0.946	-0.301	-0.645	2.827	-0.617	3.444	-5.305*	-0.0497	-5.256*
	(2.199)	(0.495)	(2.219)	(3.167)	(0.623)	(3.144)	(3.195)	(0.773)	(3.185)
Observations	481	7,415	7,896	263	3,940	4,203	218	3,475	3,693
R-squared	0.074	0.182	0.184	0.072	0.008	0.043	0.099	0.017	0.024
Mean	68.86	63.23		66.18	57.97		72.10	69.19	
S.D.	13.54	13.39		13.44	11.11		12.97	13.26	

Notes:
[1] Standard errors clustered at the household level in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.
[2] Dependent variable is weight in kilogrammes. Headings at the top of each column indicate the sub-sample used.
[3] Panels A and B show results from regressions using different Ramadan exposure dummies.
[4] Results shown in the Muslim and non-Muslim columns are coefficients of Ramadan exposure variables, in a regression that also controls for birth month dummies, gender, age, and age-squared, as well as an buffer dummy that takes value 1 if the predicted conception date occurs 21 days or less after the end of Ramadan.
[5] Results shown in the DiD column are coefficients of the Ramadan exposure variable interacted with a Muslim indicator, in a regression similar to the one described in footnote (3), but adjusted for the DiD specification.

Table A 7: Effect of in-utero exposure to Ramadan on height

	Sample: all			Sample: women only			Sample: men only		
	Muslims	Non-Muslims	DiD	Muslims	Non-Muslims	DiD	Muslims	Non-Muslims	DiD
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A									
Trimester 1	-0.0222**	0.00598***	-0.0282***	-0.00730	0.00423	-0.0115	-0.0370**	0.00801**	-0.0450**
	(0.0101)	(0.00227)	(0.0102)	(0.0141)	(0.00295)	(0.0139)	(0.0183)	(0.00350)	(0.0179)
Trimester 2	-0.0131	-5.37e-05	-0.0131	-0.0153	0.000908	-0.0162	-0.00478	-0.000804	-0.00398
	(0.00888)	(0.00196)	(0.00894)	(0.0109)	(0.00265)	(0.0109)	(0.0151)	(0.00287)	(0.0148)
Trimester 3	-0.000975	0.00362	-0.00459	0.0102	0.00380	0.00642	-0.0161	0.00285	-0.0189
	(0.0107)	(0.00223)	(0.0107)	(0.0124)	(0.00305)	(0.0124)	(0.0172)	(0.00337)	(0.0169)
Panel B									
Exposed	-0.0237*	0.00419	-0.0279**	-0.00889	0.00487	-0.0138	-0.0418	0.00337	-0.0452*
	(0.0138)	(0.00277)	(0.0139)	(0.0146)	(0.00361)	(0.0146)	(0.0270)	(0.00421)	(0.0265)
Observations	481	7,415	7,896	263	3,940	4,203	218	3,475	3,693
R-squared	0.377	0.391	0.392	0.148	0.026	0.041	0.099	0.036	0.044
Mean	1.60	1.62		1.55	1.57		1.66	1.68	
S.D.	0.10	0.09		0.07	0.07		0.08	0.07	

Notes:
[1] Standard errors clustered at the household level in parentheses. ***, **, * represent statistical significance at the 1%, 5% and 10% levels respectively.
[2] Dependent variable is height in metres. Headings at the top of each column indicate the sub-sample used.
[3] Panels A and B show results from regressions using different Ramadan exposure dummies.
[4] Results shown in the Muslim and non-Muslim columns are coefficients of Ramadan exposure variables, in a regression that also controls for birth month dummies, gender, age, and age-squared, as well as an buffer dummy that takes value 1 if the predicted conception date occurs 21 days or less after the end of Ramadan.
[5] Results shown in the DiD column are coefficients of the Ramadan exposure variable interacted with a Muslim indicator, in a regression similar to the one described in footnote (3), but adjusted for the DiD specification.

4 Income and subjective well-being: Evidence from Singapore's first national non-contributory pension

(Co-authored with Yanying Chen)

4.1 Introduction

Since Easterlin (1974)'s work on income and happiness, interest in the empirical link between income and subjective well-being has grown substantially amongst economists. In recent years, more convincing evidence on the causal impact of income on happiness (e.g. Frijters et al., 2004; Gardner & Oswald, 2007) has emerged. In addition, researchers are starting to go beyond average effects to look at heterogeneity in the effect of income on happiness arising from differences in individual characteristics (e.g. personality (Boyce & Wood, 2011) and health (Finkelstein et al., 2013)). Two recent papers also show that – consistent with the permanent income hypothesis – permanent income shocks lead to larger subjective well-being responses than transitory shocks (Bayer & Juessen, 2015; Cai & Park, 2016).

We add to this literature by using a new, high-frequency (monthly) panel and a natural experiment in Singapore to separately estimate the announcement and disbursement effects of an exogenous permanent income increase on life satisfaction and its sub-domains. In addition, we examine how these effects may vary by self-assessed financial preparation for retirement versus net assets, and investigate whether an individual's subjective well-being improves with his or her spouse's receipt of income.

The permanent income increase we study comes in the form of Singapore's first national non-contributory pension scheme (the Silver Support Scheme, or SSS). The SSS targets the neediest 20 – 30% of Singaporean citizens aged 65 and above. Details of the SSS (e.g. exact qualifying criteria and payout quantum) were announced in end-March 2016, followed by the disbursement of its first cash payout in end-July 2016. Eligibility for payouts in the period we study is pre-determined, as it is based on the government's administrative data from 2015.

Eligibility is automatically assessed by the government, and payouts are made automatically to all who are eligible via well-established channels.

The design of the SSS allows us to identify its causal impact via a difference-in-differences (DiD) strategy which includes individual[1] and time fixed effects. Our sample consists of those who are age-eligible (i.e. aged 65 and above in 2016), and the treated group is defined as those who received SSS while controls are those who did not. To address concerns that baseline differences in the characteristics of the treatment and control groups could invalidate the underlying DiD assumption of parallel trends, we construct more similar treatment and control groups by trimming our sample based on their propensity to receive SSS. In addition, we carry out a battery of robustness checks that include the addition of different group-specific time fixed effects and the use of different reweighting schemes to better match the treatment and control groups (e.g. using Abadie (2005)'s semi-parametric DiD on the full sample). These checks are described more fully in Section 4.7.

Our data comes from a new high-frequency longitudinal survey of elderly Singaporeans (the Singapore Life Panel, or SLP) which is population-representative. Singapore citizens and permanent residents aged between 50 and 70, and their spouses, are surveyed monthly. The monthly surveys allow us to track the changes in individuals' subjective well-being from the period before the announcement in end-March 2016, through the post-announcement / pre-disbursement period, to the post-disbursement period after July 2016.

We find that recipients of SSS payouts, who receive an average of around S$500 per quarter in our sample, experienced a statistically significant improvement of 2.5% of the baseline mean (or 0.11 SD) in overall life satisfaction upon SSS announcement; there seems to be no additional improvement after the disbursement of SSS payouts (i.e. the disbursement

---

[1] The use of individual fixed effects when studying subjective well-being is important, to ensure that individual unobserved heterogeneity (e.g. personality differences) does not drive our results (Ferrer- i- Carbonell & Frijters, 2004).

effect is not statistically different from the announcement effect). These results seem to be driven in part by improvements in recipients' satisfaction with their social contacts and family life, household income and economic situation[2]. In addition, we find evidence that the marginal utility of income varies across different dimensions. First, SSS recipients who felt less financially prepared for retirement at the baseline survey wave experienced a higher increase in satisfaction. Surprisingly, we find little evidence of such heterogeneity when we examine heterogeneity in responses by differences in baseline asset levels. Second, the lack of statistically significant overall effects in health satisfaction and condition masks considerable heterogeneity. Those who report being less financially prepared for retirement experienced a statistically significant larger improvement in self-rated health condition. Third, an individual's satisfaction did not improve if only his/her spouse received SSS payouts while he/she did not.

Our findings contribute mainly to the literature on income and subjective well-being. While the permanent income hypothesis (PIH) implies that forward-looking individuals' utility should react only to unanticipated, but not anticipated income shocks (see e.g. Cai & Park, 2016), empirical evidence on how subjective well-being reacts to unanticipated versus anticipated income shocks is scarce. We separately estimate the announcement and disbursement effects of income increases on subjective well-being[3], using a unique monthly longitudinal survey[4]. Our results are consistent with the predictions of the PIH: the release of detailed information about SSS eligibility led to immediate increases in subjective well-being even before payouts started, but there was no additional increase in subjective well-being upon

[2] The result for recipients' satisfaction with the social contacts and family life is only marginally significant, while those for household income and economic situation are statistically significant.

[3] Even in other fields, few non-finance papers have succeeded in capturing announcement effects. Blundell et al. (2011) and Agarwal and Qian (2014) are two examples that manage to do so; both papers show that ignoring announcement effects could bias policy effects downwards.

[4] The high-frequency data also addresses a common source of estimation bias arising from individuals' adaptation to changes in their circumstances over time (see e.g. Clark et al. (2008) for evidence that people adapt to changes in conditions in the context of life satisfaction). Such high frequency data on subjective well-being, however, is uncommon. One example of a paper that tracks subjective well-being almost as frequently is by Frijters et al. (2011), who use quarterly data from Australia to study significant life events.

disbursement of SSS payouts (i.e. the disbursement effect is not statistically different from the announcement effect). We also add to the evidence on how marginal utility of income may vary by individual characteristics, when we examine heterogeneous effects along the dimensions of perceived financial preparedness for retirement, net assets, and spouse's receipt of income. Our finding that the marginal utility of income varies by perceived financial preparedness for retirement (i.e. subjective wealth) but not actual net assets suggests that perceived financial preparedness for retirement could be more important that net assets in understanding how the marginal utility of income varies, at least among the age and socioeconomic group we study.

Beyond the income and subjective well-being literature, our results add to the literature studying the effect of non-contributory pensions on subjective well-being (e.g. Bando et al., 2016; Galiani et al., 2016). The heterogeneous responses we document suggests that researchers should move beyond studying the average effects of non-contributory pensions, to provide a fuller picture of the effects of said pensions. Policymakers wishing to maximize welfare gains may also want to take such heterogeneity into consideration when designing new or refining existing non-contributory pensions. In addition, we add to the external validity of the existing non-contributory pension literature by presenting evidence from a country with different institutions and at a different stage of development, compared to those studied earlier. Lastly, our observation that recipients with a lower perceived level of financial preparedness experienced a larger improvement in self-rated health condition strengthens existing evidence on the causal impact of income on health (e.g. Frijters et al., 2005; Lindahl, 2005).

The rest of this paper proceeds as follows. Section 4.2 gives a brief review of the literature on income and subjective well-being. Section 4.3 provides background information on the Silver Support Scheme (SSS). Section 4.4 describes our data, and in Section 4.5 we elaborate on our identification strategy and empirical model. In Section 4.6, we present our results, while Section 4.7 covers our robustness checks. Section 4.8 concludes.

4.2 Literature review on the impact of income on subjective well-being

An attractive feature of self-reported levels of happiness or subjective well-being is that they can serve as alternative measures of utility and complement traditional revealed preference approaches. The suitability of self-reported happiness or well-being for this purpose is supported by a substantial amount of evidence. For example, there is a strong positive nexus between self-reported well-being and actual well-being (see e.g. Frey & Stutzer, 2002; Kahneman & Krueger, 2006). More recently, Fleurbaey and Schwandt (2015) present survey results suggesting that 90% of their sample do attempt to maximise their subjective well-being.

Much research on subjective well-being within the economic literature focuses on the relationship between income and happiness. Early research using country-level data found that there was little, if any, increase in happiness even when real gross domestic product per capita had risen substantially over the years (e.g. Easterlin, 1974; Easterlin, 1995; Oswald, 1997). In contrast, the positive relationship between income and happiness at a particular snapshot in time within a country is well established, i.e. richer people tend to report higher subjective well-being on average (e.g. Frey and Stutzer (2000), Blanchflower and Oswald (2004)). These seemingly conflicting conclusions can be reconciled if happiness were affected not only by absolute income, but also relative income. For example, Ferrer-i-Carbonell (2005) finds that an individual's happiness is equally dependent on one's own income and the income of one's reference group. In addition, people are happier the larger their income gap as compared to their reference group.

The main criticism levelled at research in this area is its lack of causal identification. For example, unobserved individual heterogeneity such as personality differences could affect both income and happiness. Increases in income can likewise be accompanied by a rise in working hours or occupational risk which may not be accounted for but can also affect subjective well-being.

To address this critical weakness, later papers utilise exogenous shocks such as historic institutional upheavals to establish a causal link between income and happiness. Frijters et al. (2004) find that increased household incomes in East Germany, after the German reunification in 1990, contributed to around 12% of the East Germans' improvement in life satisfaction over time. Similarly, by exploiting the income changes during Russia's post-transition years, Frijters et al. (2006) report that changes in real household incomes explain 10-30% of changes in life satisfaction.

Making use of a much more common wealth shock, Gardner and Oswald (2007) and Apouey and Clark (2015) find that lottery winners have a higher level of mental wellbeing (but not self-assessed overall health).

Other authors draw on income variations resulting from changes in public policies. Kronenberg et al. (2017) report that those who benefited from the 1999 introduction of the United Kingdom National Minimum Wage experienced only limited short-run effects on mental health, while Reeves et al. (2016) find that the same policy reduced the probability of mental ill health for those who benefited. Lachowska (2016) shows that the 2008 economic stimulus tax rebates in the United States reduced feelings of stress and worry, with weaker evidence for an improvement in life and health satisfaction. Boyd-Swan et al. (2016) find positive effects of the 1990 earned income tax credit expansion in the United States, among potentially eligible women, on a range of subjective well-being measures covering mental well-being, overall happiness, and self-esteem.

Apart from natural experiments, Haushofer and Shapiro (2016) and Kilburn et al. (2016) also present positive evidence on happiness, life satisfaction, stress, and future outlook from experimental unconditional cash transfers in Kenya and Malawi respectively.

The introduction of non-contributory pensions is another source of income shocks that generally improved well-being. Case (2004) find that beneficiaries of South Africa's state old age pension experienced a lower level of depression and better self-reported health. Similarly, Galiani et al. (2016) and Bando et al. (2016) show that beneficiaries of a new pension experienced improved mental health and self-worth in Mexico and Peru respectively, but not life satisfaction. Cheng et al. (2016) find that the New Rural Pension Scheme in China led to reduced depression and better self-perceived relative economic situation, while Tseng and Petrie (2014) show that Taiwan's permanent cash injection to elderly farmers improved mental health but not self-assessed health or life satisfaction.

In addition to the focus on causal identification, recent research has moved beyond average effects of income on happiness. For example, Boyce and Wood (2011) investigate the heterogeneity of effects based on personality, while Finkelstein et al. (2013) look at how effects could vary by health status. Two recent papers show that the type of income shocks matters. Consistent with the permanent income hypothesis, permanent income shocks lead to larger subjective well-being responses than transitory shocks (Bayer & Juessen, 2015; Cai & Park, 2016).

Our paper contributes to the literature by separately estimating the announcement and disbursement effects of an exogenous permanent income increase on life satisfaction and its sub-domains. We also examine how these effects may vary by self-assessed financial preparation for retirement versus net assets, and investigate whether an individual's subjective well-being improves with his or her spouse's receipt of income.

4.3 Background on the Silver Support Scheme

The Silver Support Scheme (SSS) is Singapore's first national means-tested, non-contributory pension, which permanently supplements the income of the neediest 20 – 30% of Singaporean citizens aged 65 and above. The SSS is an important addition to Singapore's social security system, which has until now been mostly addressed by a defined contribution system known as the Central Provident Fund (CPF). Singaporeans (and their employers) contribute a proportion of income into their CPF accounts. These contributions are split across three accounts – funds from one account can be used for home purchase, funds from another can be used for healthcare expenses, while the last account sets aside money for retirement.

Details of the SSS (e.g. exact qualifying criteria and payout quantum) were announced in end-March 2016 during Singapore's annual Budget speech[5]. Eligible individuals receive quarterly payouts of S$300 – S$750, depending on the type of public housing (HDB) flat they live in[6]. Singaporeans who live in smaller flats will receive a larger payout, as flat-type is used as a proxy for socioeconomic status. These payouts constitute a significant increase in permanent income, corresponding to 7% – 18% of mean monthly household expenditure among those who received SSS payouts in our sample, or 5 – 13% of labour income at the 20th percentile of full-time resident employees (Department of Statistics, 2016). On average, recipients of SSS payouts in our sample receive around S$500 per quarter. In the period we study, the first payout was made in end-July 2016, followed by another in end-September 2016[7]. We focus on the effects of the announcement and disbursement of these payouts in this study.

Eligibility is automatically determined annually based on a combination of lifetime wages, housing type, housing ownership, and per-capita household income. To qualify,

[5] The Government first announced the introduction of the Silver Support Scheme (SSS) in August 2014, but details on qualifying criteria were not announced then. This implies that even if Singaporeans had some expectations about whether they would receive payouts from SSS, these expectations were probably weak.

[6] Most Singaporeans (80% as of 2016 – see Department of Statistics (2017)) live in high-rise public housing apartments (flats) purchased directly from the government, or in the resale market. These flats are often called HDB flats, after the statutory board (the Housing Development Board) that oversees public housing, and are categorised based on the number of rooms within each flat. The government often uses flat-type as a proxy for socio-economic status to target subsidies and transfers. In the case of the SSS, the payout quantum for individuals living in each type of flat is as follows. 1- and 2-room flats: S$750; 3-room flats: S$600; 4-room flats: S$450; 5-room flats: S$300.

[7] Payouts meant for the year 2017 and after will be made in end-December, end-March, end-June and end-September, ahead of the start of each quarter.

individuals must (i) have contributed no more than S$70,000 to their Central Provident Fund (CPF) accounts by age 55[8]; (ii) live in a 1- to 5-room HDB flat; (iii) not personally own or have a spouse who owns 5-room or larger HDB flats, private property, or multiple properties; and (iv) live in a household with a per-capita income of S$1,100 or below. They must also be Singapore citizens. For our study, receipt of Silver Support Scheme (SSS) payouts made in end-July and end-September 2016 can be seen as exogenous. This is because eligibility for the 2016 payouts is based on government data available in 2015 (i.e. before the announcement of eligibility details in March 2016), and hence pre-determined.

Payments of the SSS payouts are credited to the bank accounts that Singaporeans have already registered with the government[9]. Those without a registered bank account will receive a cheque that is mailed to the residential address they registered with the government. If the cheque is not encashed or banked in within six months, the payouts will be credited into the individuals' CPF account, and can be subsequently withdrawn within a year. This disbursement set-up suggests that Singaporeans who are eligible for the SSS payouts will almost certainly receive their payouts.

4.4 Data and variables

4.4.1 Data source

We use monthly data from waves 0 – 17[10] (covering the period May 2015 – Dec 2016) of a new longitudinal survey of elderly respondents in Singapore, the Singapore Life Panel (SLP). The SLP is a population-representative survey that aims to follow about 15,000 Singaporean citizens and permanent residents aged 50 – 70, as well as their spouses, every

[8] Self-employed persons should also have an average annual net trade income of not more than $22,800 when they were between the ages of 45 and 54.
[9] The Singapore government has a long history of giving out ad-hoc or regular cash transfers to Singaporeans, and hence has efficient systems in place that can be used to disburse any new types of cash transfers.
[10] Wave 0 refers to the baseline survey which was carried out during recruitment (between May – July 2015), while wave 1 was conducted in August 2015. Wave 17 was conducted in December 2016.

month[11]. The survey is run by the Centre for Research on the Economics of Aging (CREA), which has put measures in place to ensure that responses are population representative and that attrition remains low[12]. CREA has also verified that data from the SLP is indeed population representative, by checking that the distribution of variables collected in the SLP (e.g. age, ethnicity, marital status, education, household expenditure) match those from government statistics. More details on the SLP can be found in Vaithianathan et al. (2017).

There are several advantages to using this dataset for our study. First, the high frequency at which the survey is carried out allows us to time the effects of this permanent income shock (i.e. the Silver Support Scheme (SSS)) more precisely, compared to most other surveys which are carried out at yearly or quarterly intervals. Second, the survey had been running for a few months before announcement of the SSS details, allowing us to separately identify the announcement and disbursement effects, as well as examine the credibility of our DiD identifying assumption. Third, as we will see later in this section, the SLP questions on subjective well-being go beyond the more commonly asked life satisfaction question, and include questions on narrower domains of subjective well-being such as satisfaction with one's social life or household income, allowing us to study which of these narrower domains might drive changes in overall life satisfaction. Lastly, the richness of data in the SLP will also allow us to explore how responses to the SSS might vary among different segments of the population.

4.4.2 Variables

We focus our analysis on the effect of receiving a permanent income shock on subjective well-being, and restrict the sample for our main analysis to age-eligible respondents (aged 65

[11] While not everyone responds to every survey wave, the number of participants who respond in any particular wave has remained stable at around 8,000.

[12] For example, the surveys are available in all four major languages spoken by Singaporeans. In addition, while the surveys are conducted over the internet, respondents who are unable to understand the survey questions or who do not have access to the internet can answer the survey over the phone, or at centres set up at convenient locations around Singapore, where the survey will be conducted by trained interviewers. CREA also conducts ongoing campaigns to encourage participation.

and above in 2016) who are Singapore citizens and who live in public housing flats (as of 2016, about 80% of resident households live in public housing flats (Department of Statistics, 2017)). These are three out of the list of eligibility criteria that individuals must meet to qualify for SSS payouts in 2016. This sample includes individuals who receive SSS as well as those who do not. SSS recipients are identified using a deliberately timed quarterly question in the SLP which asks whether the respondent received SSS in the previous month; individuals who report receiving at least one out of the two payouts in 2016[13] are coded as SSS recipients.

Apart from questions on treatment status, our dataset includes a rich set of baseline demographics (age, marital status, gender, ethnicity, education, housing type, number of household members, assets, and self-reported financial preparedness for retirement), as well as measures of different domains of evaluative subjective well-being, which will serve as our outcome variables.

The subjective well-being variables include a broad measure of overall life satisfaction, as well as satisfaction within narrower domains such as household income; these variables are rated from 1 (worst) to 5 (best). Data on these are collected via the following questions: (i) "Taking all things together, how satisfied are you with your life as a whole these days?"; (ii) "How satisfied are you with your social contacts and family life?; (iii) "How satisfied are you with your daily activities, and if you are working, your job?"; (iv) "How satisfied are you with the total income of your household"; (v) "How satisfied are you with your overall economic situation?"; (vi) "How satisfied are you with your health"; and (vii) "Would you say your health is excellent, very good, good, fair, or poor?" [14].

To end off this section, we report summary statistics for the baseline demographics and subjective well-being outcomes for those who are age-eligible for the SSS (i.e. aged 65 and

[13] Once in end-July 2016, and another in end-September 2016.
[14] Options for the first six questions are "very dissatisfied", "dissatisfied", "neither satisfied nor dissatisfied", "satisfied", and "very satisfied"; options for the last question are "poor", "fair", "good", "very good" and "excellent".

above in 2016), disaggregated by treatment status, in **Table 4.1** and **Table 4.2** respectively. Focusing on the full age-eligible sample in **Table 4.1**, we can see that SSS recipients and non-recipients differ in terms of several baseline characteristics. E.g., women are more likely to receive SSS, and SSS recipients report being less financially prepared for retirement than non-SSS recipients.

While these differences are relatively small in magnitude (judging by the normalised differences), and causal identification in our study relies on a Difference-in-Differences (DiD) strategy which allows for differences in baseline characteristics, one might be concerned that the treated and control groups are different enough that the DiD identifying assumption – parallel trends in the absence of treatment – might not hold. In Section 4.5, we discuss our methods for addressing this concern and explain **Table 4.1** in greater detail.

4.5 Identification strategy and empirical model

4.5.1 Identification strategy

We use a difference-in-differences (DiD) strategy to identify the average treatment-on-treated (ATT) effect of receiving Silver Support Scheme (SSS) payouts on a broad range of subjective well-being domains. Our main analysis focuses on age-eligible individuals (aged 65 and above in 2016) who are Singapore citizens and who live in public housing flats[15]. These are three out of the list of eligibility criteria that individuals must meet to qualify for SSS payouts in 2016. This sample includes both SSS recipients (the treated group) and non-recipients (the control group). Waves 0 – 8 (May 2015 – Mar 2016)[16] make up the pre-announcement period; waves 9 – 12 (Apr – Jul 2016) are the post-announcement and pre-disbursement period; and waves 13 – 17 (Aug – Dec 2016) are the post-disbursement period.

[15] As of 2016, 80% of resident households live in public housing (Department of Statistics, 2017).
[16] Wave 0 covers May-Jul 2015.

As we note in Section 4.4, the treated and control groups differ from each other in terms of several baseline characteristics. One might be concerned that differences between these groups could invalidate the DiD identifying assumption (i.e. outcomes in both groups follow the same trend in the absence of treatment).

Our main strategy for addressing this concern is to construct a sample where the treatment and control groups are more similar. We start by using logistic regression to estimate the propensity score for receiving SSS payouts. The covariates are selected from a rich pool of key baseline demographic variables that could affect one's eligibility for SSS payouts[17]. Using the algorithm outlined in Imbens (2015), which proposes a data-driven way of selecting a subset of baseline covariates and their interactions, we select the following covariates and some of their interactions: age, marital status, gender, ethnicity, education, public housing flat type, whether respondent's father is still living, number of household members, number of living children, income of self and spouse, baseline self-assessment of financial preparedness for retirement and baseline satisfaction with one's economic situation.

We then trim the sample progressively at both extreme ends of the propensity score till the treated and control groups are more similar in terms of the baseline characteristics as well as pre-announcement time trends for our outcome variables. **Figure 4.1** plots the unconditional mean of the life satisfaction variable across time, for four samples with different ranges of propensity score: (i) 0.00 to 1.00; (ii) 0.10 to 0.90; (iii) 0.15 to 0.85; and (iv) 0.20 to 0.80. As we restrict the sample to narrower ranges of propensity score, the pre-announcement time trends for the treated and control groups start to converge. For the smallest sample with propensity scores of 0.20 to 0.80, we see that the pre-announcement trends are almost parallel. The same can be observed for other outcome variables in **Figure 4.2** and **Figure 4.3**.

[17] E.g. Total CPF contributions of not more than $70,000 by age 55, living in a household with per capita income of not more than $1,100

We also see a noticeable improvement in the comparability of the treated and control groups in terms of baseline characteristics after trimming. The right panel in **Table 4.1**, which summarises statistics for the trimmed sample with propensity scores of 0.20 to 0.80, shows that differences in education, whether one lives in a 2-room flat, whether one owns a home, and perceived financial preparedness for retirement fall to statistical insignificance, and differences in all other variables shrink noticeably. All normalised differences[18] in the trimmed sample also drop to 0.16 or less in absolute magnitude, well below the value of 0.3 which Imbens (2015) deems "modest"[19]. In addition, **Figure 4.4** shows that the distributions of key covariates related to SSS eligibility become more similar after trimming. Based on our above comparisons of baseline characteristics and pre-treatment trends across different samples and between the treated and control groups, we deem it reasonable to use the sample with propensity scores 0.20 to 0.80 for our main DiD analyses.

While trimming the sample based on observables does not necessarily lead to comparability based on unobservables, it is unlikely that the remaining differences in unobservables will lead to a violation of the common trends assumption for DiD. Differences in unobservables could arguably be critical in the case where individuals self-select into the treatment group, but as we explained in Sections 4.1 and 4.3, the eligibility of individuals for the end-July and September payouts are predetermined before the announcement, and payouts are made automatically to all eligible individuals. In addition, the propensity score we use for trimming is also formulated based on variables that are related to individuals' eligibility for the SSS payouts.

[18] Normalised differences are computed as in Imbens (2015), as the difference in means standardised by the square root of the mean variances of both groups.
[19] **Table A1** in Appendix A compares the baseline characteristics for the treated and control groups for samples with propensity scores 0.10 to 0.90 and 0.15 to 0.85. The differences in means fall as we restrict the sample progressively.

In addition to the visual check of pre-announcement trends, we statistically test the DiD identifying assumption by adding pre-announcement leads to our regression specification. In most of the outcomes analysed, the coefficients of the pre-treatment leads are statistically insignificant, increasing the probability that our identifying assumption is valid. (The exceptions are pre-treatment leads for economic satisfaction and daily activities satisfaction which are significant in March -- the month of SSS announcement, though this could reflect early anticipation about the annual Budget announcements in general.)

Beyond this, we carry out a battery of robustness checks (including different reweighting schemes to better match the treatment and control groups, such as Abadie (2005)'s semi-parametric DiD applied on the full sample and matching-DiD using a 1:1 nearest neighbour match) to verify the validity of our approach. Section 4.7 discusses these checks in detail.

Finally, we note that trimming our sample means that we identify the ATT for only a subset of those who are treated, if treatment effects are heterogeneous. However, we show in Section 4.7 that the ATT effects we estimate in our main specifications are close to those we estimate on the full sample (both with and without Abadie (2005)'s reweighting scheme[20]).

4.5.2 Empirical specifications

We start by estimating the following regression to study the overall effects of the Silver Support Scheme (SSS):

$$\begin{aligned} Y_{it} = \quad & \sum_{k=-3}^{-1} \beta_{pre,k} \left(Treat_i \times Preann_{t,k}\right) \\ & + \beta_{ann}\left(Treat_i \times Anntodisb_t\right) + \beta_{disb}\left(Treat_i \times Postdisb_t\right) \\ & + \alpha_i + \gamma_t + \epsilon_{it} \end{aligned} \tag{1}$$

[20] Abadie (2005)'s semi-parametric DiD involves weighting each control observation by their propensity score, and estimating the effects of receiving SSS using the full age-eligible sample.

where Y_{it} is the outcome variable for respondent i at time t; $Treat$ is a dummy variable that takes on value one if the individual ever received SSS payouts; $Preann_{t,-3}$, $Preann_{t,-2}$, $Preann_{t,-1}$ are the pre-announcement leads (dummy variables that take value 1 if time t corresponds to January, February, and March 2016 respectively)[21]; $Anntodisb$ is a dummy for the period between announcement and disbursement (waves 9 – 12, Apr – Jul 2016); $Postdisb$ is a dummy for the period after disbursement (waves 13 – 17, Aug – Dec 2016) ; while α_i and γ_t are individual and time fixed-effects respectively. The suppressed period consists of waves 0 to 5 (May/Jun/Jul – Dec 2015). Standard errors are clustered at the household level.

$\beta_{pre,k}$ test the assumption of common time trends statistically. Our identifying assumption will be more credible if the coefficients in β_{pre} are statistically insignificant. For the age-eligible sample, β_{ann} captures the announcement effect, and β_{disb} captures the disbursement effect.

Next, we modify equation (1) to study how treatment effects may vary by baseline financial preparedness for retirement and wealth:

$$
\begin{aligned}
Y_{it} = \quad & \sum_{k=-3}^{-1} \beta_{pre,k}\,(Treat_i \times Preann_{t,k}) \\
& +\beta_{ann}(Treat_i \times Anntodisb_t) + \beta_{disb}(Treat_i \times Postdisb_t) \\
& +\beta_{ann,fin}(Treat_i \times Anntodisb_t \times Fin_i) \\
& + \beta_{disb,fin}(Treat_i \times Postdisb_t \times Fin_i) \ + \alpha_i + \gamma_t + \epsilon_{it}
\end{aligned}
\qquad (2)
$$

where Fin can either be (i) a scale variable running from 1 (poor) to 5 (excellent) that measures subjective financial preparedness for retirement, or (ii) a scale variable running from 1(poorest)

[21] March 2016 is considered in the pre-announcement period as the announcement is made only towards the end of March. The survey wave in March would have closed before the announcement is made.

to 5(richest) in terms of pre-announcement assets quintile[22]. Financial preparedness for retirement can serve as a proxy for baseline assets levels, but it also contains additional information about an individual's financial burden, as well as consumption and risk preferences. E.g. compared to another individual with the same level of wealth, a person with lower risk tolerance (and who would want to build up more savings for low probability adverse events) or who must support more dependents would be likely to report being less financially prepared for retirement. Taken together, these two sets of regressions will provide a fuller picture. $\beta_{ann,fin}$ and $\beta_{disb,fin}$ capture heterogeneity in effects across either subjective financial preparedness for retirement or assets.

To the extent that financial preparation for retirement / wealth is correlated with flat-type, $\beta_{ann,fin}$ and $\beta_{disb,fin}$ may also reflect the effects of differing payout levels, since payout levels are determined by the type of flat individuals live in. As such, we estimate another set of regressions which consider heterogeneity across *Fin* within each flat-type, to verify that our results from equation (2) indeed reflect heterogeneity in *Fin*, and not differing payout levels:

$$\begin{aligned} Y_{it} = \quad & \sum_{k=-3}^{-1} \beta_{pre,k} \left(Treat_i \times Preann_{t,k}\right) \\ & + \sum_{f \in F} \beta_{ann,f} \left(Treat_i \times Anntodisb_t \times Flat_{i,f}\right) \\ & + \sum_{f \in F} \beta_{ann,f,fin} \left(Treat_i \times Anntodisb_t \times Flat_{i,f} \times Fin_i\right) \\ & + \sum_{f \in F} \beta_{disb,f} \left(Treat_i \times Postdisb_t \times Flat_{i,f}\right) \\ & + \sum_{f \in F} \beta_{disb,f,fin} \left(Treat_i \times Postdisb_t \times Flat_{i,f} \times Fin_i\right) \\ & + \alpha_i + \gamma_t + \epsilon_{it} \end{aligned} \tag{3}$$

[22] These asset quintiles are based on a sample that includes only those who are age-eligible (aged 65 and above in 2016), and who live in public housing.

$Flat_{i,f}$ is a dummy variable that takes value 1 if the individual lives in flat type *f*. F includes 1/2- , 3-, 4-, 5-room and other flats[23]. $\beta_{ann,f,fin}$ and $\beta_{disb,f,fin}$ are our coefficients of interest.

To investigate if responses may vary by the identity of SSS recipients within a couple, we restrict our sample to respondents who are married and estimate:

$$\begin{aligned} Y_{it} = \quad & \sum_{k=-3}^{-1} \beta_{pre,k} \left(At\ least\ one\ treated_i \times Preann_{t,k}\right) \\ & + \sum_{g \in G} \beta_{ann,g} (Who\_treated_{i,g} \times Anntodisb_t) \\ & + \sum_{g \in G} \beta_{disb,g} (Who\_treated_{i,g} \times Postdisb_t) \\ & + \alpha_i + \gamma_t + \epsilon_{it} \end{aligned} \tag{4}$$

where $At\ least\ one\ treated_i$ is a dummy variable that takes value 1 if either respondent *i* or respondent *i*'s spouse received SSS; $Who\_treated_{i,g}$ can be a dummy variable indicating that (i) only the respondent received SSS; (ii) only the respondent's spouse received SSS[24]; or (iii) both the respondent and his/her spouse received SSS. $\beta_{ann,g}$ and $\beta_{disb,g}$ reflect the heterogeneity of effects based on the identity of SSS recipients within a couple. E.g. $\beta_{disb,respondent\ only}$ will give the effect of only the respondent receiving SSS on his/her subjective well-being in the post-disbursement period[25].

[23] Results for 1/2- and 5-room flats should be interpreted with caution, as relatively fewer people stay in these flat-types. We do not report results for the category "other flats" as it is made up of flats of unknown types. As we have already accounted for all flat-type categories in this specification, we exclude the interaction terms $Treat_i \times Anntodisb_t$ and $Treat_i \times Postdisb_t$ to avoid perfect collinearity.
[24] This variable is captured through a question in the SLP on whether a respondent's spouse received SSS.
[25] The dynamics of how subjective well-being may change over time is interesting too. We estimated these regressions but do not observe any obvious patterns in the dynamics. This is not surprising as Frijters et al. (2011) find that the effects of an improvement in financial situation persists even up to 8 quarters after the event. In the interest of space, we will not report these results, though they are available on request.

4.6 Results

4.6.1 Overall effects

Table 4.3 reports the overall effect of receiving Silver Support Scheme (SSS) payouts on different domains of evaluative subjective well-being (specified at the top of each column). Coefficients for pre-announcement leads are generally statistically insignificant, providing evidence in favour of the DiD identifying assumption. The exceptions are leads for daily activities and economic satisfaction in March (the month of SSS announcement), which may reflect early anticipation about the annual Budget announcements in general, especially in the case of economic satisfaction.

We find that the SSS, which amounts to around S$500 per quarter among recipients in our sample, leads to a statistically significant improvement in overall life satisfaction (column 1). This improvement starts immediately upon SSS announcement, and does not increase further after disbursement starts: life satisfaction rises by 0.09 (about 2.5% of the baseline mean / 0.11SD) upon announcement of the SSS, and stays elevated at this level upon commencement of the SSS payouts (see **Table 4.3** for results showing that the difference between β_{ann} and β_{disb} is not statistically significant). As one would expect, this rise in life satisfaction is at least partly driven by statistically significant increases in household income satisfaction (0.11, or 3.5% of the baseline mean / 0.13SD) and overall economic satisfaction (0.08, or 2.5% of the baseline mean / 0.10SD). In addition, there is some evidence of a marginally significant rise in social and family life satisfaction (0.07, or 1.9% of the baseline mean / 0.09SD). The coefficients of the treatment variables for health satisfaction and self-rated health condition are positive but insignificant, suggesting that the SSS has little, if any, overall effect on health. Consistent with the predictions of the PIH (see e.g. Cai & Park, 2016), there is no additional increase in life satisfaction (and its sub-domains) upon disbursement of the payouts – the difference in magnitude between β_{ann} and β_{disb} is small in magnitude and not statistically significant.

The direction of our main life satisfaction result is consistent with other published papers that look at the causal effect of income on life satisfaction (e.g. Frijters et al., 2004; Frijters et al., 2006; Boyd-Swan et al., 2016; Haushofer & Shapiro, 2016; Lachowska, 2016). In terms of magnitude, our main life satisfaction result is slightly smaller than the 0.17SD improvement in life satisfaction due to a large cash transfer[26] in Kenya reported in Haushofer and Shapiro (2016). Compared to East Germans who experienced an approximate 25% rise in real household incomes between 1991 and 1995, our life satisfaction result appears slightly larger as East Germans experienced a corresponding increase of 1.4 – 2.2% of their 1990 mean life satisfaction (Frijters et al., 2004).

4.6.2 Heterogeneity by financial preparedness for retirement and assets

The overall effects mask heterogeneity by subjective financial preparedness for retirement. **Table 4.4** shows statistically significant evidence of variation by financial preparedness for retirement in life satisfaction, household income satisfaction, economic satisfaction, as well as self-rated health condition. Consistent with intuition, individuals who felt less financially prepared for retirement in the baseline experienced larger increases in subjective well-being from receiving SSS payouts; on average, a one-point decrease (on a 5-point scale) in retirement preparedness increases the improvement in satisfaction from SSS payouts by about 0.06 to 0.08.

In addition, once we allow for heterogeneous effects, the effect of SSS on health for those who are least prepared for retirement surfaces. E.g., self-rated health condition for those who are least prepared for retirement improved by 0.13 points[27] during the post-disbursement period

[26] Between US$404 PPP to US$1,525 PPP was transferred to households in this experiment. The mean transfer was US$709 PPP, corresponding to almost two years of per-capita expenditure.
[27] 0.19 – 0.06 = 0.13

(5% of the mean, or 0.15SD). Our results are consistent with that of Jones and Schurer (2011), who also document heterogeneity in the effect of income on health, as well as past papers looking at the effect of income on health (Frijters et al., 2005; Behrman et al., 2011; Gunasekara et al., 2011).

Financial preparedness for retirement, however, may be correlated with flat-type and thus payout quantum (payout levels are determined entirely by flat-type). This would imply that the variation we find above could be driven purely by the fact that less financially prepared individuals received a higher payout quantum. We verify that this is not the case in **Table A2** of Appendix A, by investigating the existence of variation by financial preparedness within each flat-type. Generally, we still find that individuals who felt less financially prepared for retirement experienced larger increases in subjective well-being.

When we switch to heterogeneity by baseline asset quintiles in

Table 4.5, we see that coefficients for interactions between the policy and asset variables are mostly insignificant and small in magnitude when compared to those for financial preparedness[28]. In all, the results in this sub-section suggest the marginal utility of income is decreasing in subjective financial preparedness for retirement, but the same extent of heterogeneity is not observed for net assets.

There are a few potential explanations for this somewhat surprising result. First, subjective financial well-being may incorporate more information than objective measures of wealth, such as net assets. For example, subjective financial well-being is likely to capture information related to financial burden or differences in risk aversion, while net assets may not. For any given asset level, individuals with higher financial burden or risk aversion may be likely to benefit more from an increase in permanent income, due to an alleviation in financial burden and a decrease in the probability that retirement income will be insufficient in future respectively, and may be a reason why we observe more heterogeneity in responses by subjective financial preparedness for retirement.

Second, survey respondents may be more likely to give accurate answers to subjective financial preparedness for retirement than net assets. This potentially greater mismeasurement in net assets may then contribute to the lower heterogeneity in responses we observe for net assets.

Third, the variation in net assets among our treated, trimmed, sample is smaller than that in the full sample (see the third row in **Figure 4.4**). It is thus possible that the variation in assets within our sample is not sufficient for us to observe much heterogeneity in responses to a permanent income shock. However, we do not believe that this is likely to be an important

[28] We obtain similar results when we use baseline asset deciles instead of quintiles, and when we use non-housing assets instead of total assets (results available upon request).

explanation, as the same figure shows that the remaining variation in assets among the trimmed and treated group is still quite large.

Regardless of the potential explanations, our results in this section suggest that perceived financial well-being could be more important than net assets, as a summary measure, in understanding how individuals' utility change in response to income, at least among the age and socioeconomic group we study. As this result may have implications for the welfare effects of income transfers and how policymakers should target income transfers, it would be interesting to see if this result holds up in different contexts in future studies.

4.6.3 Heterogeneity by identity of recipients

Finally, we report results on how responses vary by the identity of the SSS recipient within a couple. Results in **Table 4.6** are based only on married respondents. Our first result is intuitive: subjective well-being rises slightly more if both the respondent and his/her spouse received SSS payouts, compared to the case where only the respondent received SSS. E.g., life satisfaction increases by 0.12 if only the respondent received SSS, and it increases by an additional 0.03 if both the respondent and his/her spouse receive SSS. When only the respondent's spouse received SSS payouts, the respondent's well-being does not appear to increase by much: the results are statistically insignificant across all domains of subjective well-being.

Our results suggest that having only a spouse receive SSS payouts does not improve a respondent's subjective well-being by much (if at all), and are consistent with a scenario of limited income pooling within the household. These results are also consistent with findings from past papers that focus on testing the unitary household model and the income pooling property using different outcome variables (see Donni and Chiappori (2011) for a review)[29].

[29] While we note that our results are consistent with a scenario of limited income pooling, we acknowledge that more work will need to be done if we wish to make a more definitive statement regarding the unitary household model. Such work, however, is beyond the scope of our current paper. Separately, we look at whether the effect of receiving SSS varies by gender as well, and find no evidence that it does (results available on request).

4.7 Robustness Checks

The overall impact of the SSS on each outcome we analyse is subjected to the following battery of robustness checks based on modifications of equation (1):

1. Removal of pre-announcement leads to verify that results are robust to changes in baseline period definition.
2. Restriction of sample to a "balanced" panel, where every individual has at least one observation in each of the pre-announcement, announcement-to-disbursement, and post disbursement periods, to verify that compositional changes in respondents across waves are not driving our results.
3. Addition of ethnicity- and flat-type-specific time fixed effects, to allow for differential time trends in different groups.
4. Addition of controls for additional welfare payments: whether the respondent received the Workfare Income Supplement (a wage subsidy) or GST vouchers (a modest one-off cash transfer)[30].
5. Estimating the effects of SSS using Abadie (2005)'s semiparametric DiD[31] and DiD matching using a 1-1 nearest neighbour match based on the propensity of receiving SSS[32], to verify that our results are robust to different methods of addressing imbalances in baseline characteristics that may affect the parallel trends assumption.

[30] We do not include controls for additional welfare payments in our main specification as these questions are only asked quarterly, so not all respondents reply to these questions. Including these controls in our main specification will reduce our sample size and reduce statistical power considerably.

[31] This method addresses the imbalance of baseline characteristics between the treated and control groups by reweighting control observations based on their propensity score; control observations with a higher propensity score are given a higher weight. We use the Stata package *absdid* described in Houngbedji (2015) to implement this estimator.

[32] This method addresses the same issue – imbalance in baseline covariates that might be correlated with trends – that Abadie (2005) and our main strategy tries to address. We match each treated individual to a control individual based on a nearest neighbor match using the propensity score, compute the DiD for each pair, then aggregate these results to obtain the impact of SSS receipt. We compute p-values using a permutation test, using methods similar to those described in Abadie et al. (2010), Robbins et al. (2016), and Chang and Lee (2011).

We find that our results are robust to the battery of checks we carry out. **Table A3** to **Table A6** in Appendix A report the results of these robustness checks for outcomes that show an overall response to receipt of SSS: life, social, household income, and economic satisfaction[33]. The estimated effect of receiving SSS payouts remains significant across almost all our robustness checks, except for two checks in social satisfaction, where we lose significance due to the smaller sample used in those checks. The magnitude of the estimated coefficients remains quite stable across different checks for each outcome variable.

In addition, the estimated effect of receiving SSS from using Abadie (2005)'s semi-parametric DiD on the full/untrimmed sample (in column (6) of **Table A3** to **Table A6**) are similar in magnitude to those from our trimmed sample, providing evidence that estimated ATTs from our trimmed and full sample are likely to be similar. We also present results from our full sample estimated using Equation (1) in **Table A7**. The estimated SSS effects are generally similar.

4.8 Conclusion

In this paper, we use a DiD strategy to study the causal effect of a permanent income increase on a broad range of satisfaction domains. The source of exogenous variation comes from a new national means-tested non-contributory pension in Singapore, the Silver Support Scheme (SSS), for which eligibility is pre-determined and payouts to all eligible individuals are disbursed automatically. Using a new monthly longitudinal dataset of elderly Singaporeans, we precisely time and estimate the SSS announcement and disbursement effects. We find that announcement effects are important: life satisfaction rises among recipients by about 0.11SD upon SSS announcement, and this rise is sustained (but does not increase further) after disbursement of the payouts. This improvement appears to be driven by social, household income, and economic satisfaction. Consistent with the predictions of the PIH (see e.g. Cai &

[33] Results for other outcomes are available upon request.

Park, 2016), there is no additional effect upon disbursement – the difference between the announcement and disbursement effects is small and statistically insignificant.

We also explore heterogeneity in the marginal utility of income. Consistent with intuition, recipients who reported being less financially prepared for retirement exhibited larger increases in well-being. Surprisingly, an analysis of how marginal utility of income varies by net assets shows little evidence of such heterogeneity, suggesting that subjective financial well-being may be more important than net assets in understanding how the marginal utility of income varies, at least among the age and socioeconomic group we study. In addition, once we account for heterogeneity, we discover that the self-rated health condition of those who are least financially prepared for retirement improved. Lastly, we find that well-being improved negligibly when an individual's spouse received SSS payouts while he/she did not, a result that is consistent with limited income pooling within the household.

Our findings add to the literature on the effect of income on subjective well-being, by separately estimating the announcement and disbursement effects of income on subjective well-being. We also document the presence of heterogeneity in responses to the SSS, which suggests that future policies designed with greater consideration of heterogeneity among individuals could reap higher welfare gains. One surprising result is that subjective financial well-being may be more important than net assets in understanding how marginal utility of income varies. It would be interesting to see if this result holds up in different contexts in future studies, as such a result may have implications for the welfare effects and targeting of income transfers. Lastly, our results also strengthen the evidence on the effects of non-contributory pensions on subjective well-being, and of income on health.

References

Abadie, A. (2005). Semiparametric Difference-in-Differences Estimators. *The Review of Economic Studies, 72*(1), 1-19.

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California’s Tobacco Control Program. *Journal of the American statistical Association, 105*(490), 493-505.

Agarwal, S., & Qian, W. (2014). Consumption and Debt Response to Unanticipated Income Shocks: Evidence from a Natural Experiment in Singapore. *The American Economic Review, 104*(12), 4205-4230.

Apouey, B., & Clark, A. E. (2015). Winning Big but Feeling No Better? The Effect of Lottery Prizes on Physical and Mental Health. *Health Economics, 24*(5), 516-538.

Bando, R., Galiani, S., & Gertler, P. (2016). The Effects of Non-Contributory Pensions on Material and Subjective Well Being. *NBER Working Paper No. 22995*.

Bayer, C., & Juessen, F. (2015). Happiness and the Persistence of Income Shocks. *American Economic Journal: Macroeconomics, 7*(4), 160-187.

Behrman, J., Calderon, M. C., Mitchell, O. S., Vasquez, J., & Bravo, D. (2011). First-Round Impacts of the 2008 Chilean Pension System Reform. *Michigan Retirement Research Center Research Paper No. WP, 245*.

Blanchflower, D. G., & Oswald, A. J. (2004). Well-Being over Time in Britain and the USA. *Journal of Public Economics, 88*(7), 1359-1386.

Blundell, R., Francesconi, M., & van der Klaauw, W. (2011). Anatomy of Welfare Reform Evaluation: Announcement and Implementation Effects. Iza Discussion Papers 6050. *Institute for the Study of Labor (IZA).*

Boyce, C. J., & Wood, A. M. (2011). Personality and the Marginal Utility of Income: Personality Interacts with Increases in Household Income to Determine Life Satisfaction. *Journal of Economic Behavior & Organization, 78*(1), 183-191.

Boyd-Swan, C., Herbst, C. M., Ifcher, J., & Zarghamee, H. (2016). The Earned Income Tax Credit, Mental Health, and Happiness. *Journal of Economic Behavior & Organization, 126*, 18-38.

Cai, S., & Park, A. (2016). Permanent Income and Subjective Well-Being. *Journal of Economic Behavior & Organization, 130*, 298-319.

Case, A. (2004). Does Money Protect Health Status? Evidence from South African Pensions *Perspectives on the Economics of Aging* (pp. 287-312): University of Chicago Press.

Chang, P.-L., & Lee, M.-J. (2011). The Wto Trade Effect. *Journal of International Economics, 85*(1), 53-71.

Cheng, L., Liu, H., Zhang, Y., & Zhao, Z. (2016). The Health Implications of Social Pensions: Evidence from China's New Rural Pension Scheme. *Journal of Comparative Economics*.

Clark, A. E., Diener, E., Georgellis, Y., & Lucas, R. E. (2008). Lags and Leads in Life Satisfaction: A Test of the Baseline Hypothesis. *The Economic Journal, 118*(529).

Department of Statistics. (2016). *Yearbook of Statistics Singapore*. Retrieved from Singapore: http://www.singstat.gov.sg/docs/default-source/default-document-library/publications/publications_and_papers/reference/yearbook_2016/yos2016.pdf

Department of Statistics. (2017). Latest Data. Retrieved from http://www.singstat.gov.sg/statistics/latest-data#22

Donni, O., & Chiappori, P.-A. (2011). Nonunitary Models of Household Behavior: A Survey of the Literature *Household Economic Behaviors* (pp. 1-40): Springer.

Easterlin, R. A. (1974). Does Economic Growth Improve the Human Lot? Some Empirical Evidence. *Nations and households in economic growth, 89*, 89-125.

Easterlin, R. A. (1995). Will Raising the Incomes of All Increase the Happiness of All? *Journal of Economic Behavior & Organization, 27*(1), 35-47.

Ferrer-i-Carbonell, A. (2005). Income and Well-Being: An Empirical Analysis of the Comparison Income Effect. *Journal of Public Economics, 89*(5), 997-1019.

Ferrer- i- Carbonell, A., & Frijters, P. (2004). How Important Is Methodology for the Estimates of the Determinants of Happiness? *The Economic Journal, 114*(497), 641-659.

Finkelstein, A., Luttmer, E. F., & Notowidigdo, M. J. (2013). What Good Is Wealth without Health? The Effect of Health on the Marginal Utility of Consumption. *Journal of the European Economic Association, 11*(s1), 221-258.

Fleurbaey, M., & Schwandt, H. (2015). Do People Seek to Maximize Their Subjective Well-Being? *IZA discussion paper no. 9450*.

Frey, B. S., & Stutzer, A. (2000). Happiness, Economy and Institutions. *The Economic Journal, 110*(466), 918-938.

Frey, B. S., & Stutzer, A. (2002). What Can Economists Learn from Happiness Research? *Journal of Economic Literature, 40*(2), 402-435.

Frijters, P., Geishecker, I., Haisken- DeNew, J. P., & Shields, M. A. (2006). Can the Large Swings in Russian Life Satisfaction Be Explained by Ups and Downs in Real Incomes? *The Scandinavian Journal of Economics, 108*(3), 433-458.

Frijters, P., Haisken-DeNew, J. P., & Shields, M. A. (2004). Investigating the Patterns and Determinants of Life Satisfaction in Germany Following Reunification. *Journal of Human Resources, 39*(3), 649-674.

Frijters, P., Haisken-DeNew, J. P., & Shields, M. A. (2005). The Causal Effect of Income on Health: Evidence from German Reunification. *Journal of Health Economics, 24*(5), 997-1017.

Frijters, P., Johnston, D. W., & Shields, M. A. (2011). Life Satisfaction Dynamics with Quarterly Life Event Data. *The Scandinavian Journal of Economics, 113*(1), 190-211.

Galiani, S., Gertler, P., & Bando, R. (2016). Non-Contributory Pensions. *Labour Economics, 38*, 47-58.

Gardner, J., & Oswald, A. J. (2007). Money and Mental Wellbeing: A Longitudinal Study of Medium-Sized Lottery Wins. *Journal of Health Economics, 26*(1), 49-60.

Gunasekara, F. I., Carter, K., & Blakely, T. (2011). Change in Income and Change in Self-Rated Health: Systematic Review of Studies Using Repeated Measures to Control for Confounding Bias. *Social Science & Medicine, 72*(2), 193-201.

Haushofer, J., & Shapiro, J. (2016). The Short-Term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya. *The Quarterly Journal of Economics, 131*(4), 1973-2042.

Houngbedji, K. (2015). Abadie's Semi-Parametric Difference-in-Difference Estimator. *The Stata Journal, 16(2),* 482 - 490.

Imbens, G. W. (2015). Matching Methods in Practice: Three Examples. *Journal of Human Resources, 50*(2), 373-419.

Jones, A. M., & Schurer, S. (2011). How Does Heterogeneity Shape the Socioeconomic Gradient in Health Satisfaction? *Journal of Applied Econometrics, 26*(4), 549-579.

Kahneman, D., & Krueger, A. B. (2006). Developments in the Measurement of Subjective Well-Being. *The Journal of economic perspectives, 20*(1), 3-24.

Kilburn, K., Handa, S., Angeles, G., Mvula, P., & Tsoka, M. (2016). Happiness and Alleviation of Income Poverty: Impacts of an Unconditional Cash Transfer Programme Using a Subjective Well-Being Approach. *Innocenti Working Paper WP-2016-23*.

Kronenberg, C., Jacobs, R., & Zucchelli, E. (2017). The Impact of the Uk National Minimum Wage on Mental Health. *SSM-Population Health*.

Lachowska, M. (2016). The Effect of Income on Subjective Well-Being: Evidence from the 2008 Economic Stimulus Tax Rebates. *Journal of Human Resources*.

Lindahl, M. (2005). Estimating the Effect of Income on Health and Mortality Using Lottery Prizes as an Exogenous Source of Variation in Income. *Journal of Human Resources, 40*(1), 144-168.

Oswald, A. J. (1997). Happiness and Economic Performance. *The Economic Journal, 107*(445), 1815-1831.

Reeves, A., McKee, M., Mackenbach, J., Whitehead, M., & Stuckler, D. (2016). Introduction of a National Minimum Wage Reduced Depressive Symptoms in Low- Wage Workers: A Quasi- Natural Experiment in the Uk. *Health Economics*.

Robbins, M. W., Saunders, J., & Kilmer, B. (2016). A Framework for Synthetic Control Methods with High-Dimensional, Micro-Level Data: Evaluating a Neighborhood-Specific Crime Intervention. *Journal of the American statistical Association*(just-accepted).

Tseng, F.-M., & Petrie, D. J. (2014). The Implications for Health, Depression, and Life Satisfaction from a Permanent Increase in Income for the Disadvantaged Elderly: Evidence from Taiwan. *Review of Social Economy, 72*(3), 311-336.

Vaithianathan, R., Hool, B., Hurd, M., & Rohwedder, S. (2017). High-Frequency Internet Survey of a Probability Sample of Older Singaporeans: The Singapore Life Panel. *CREA working paper*.

Table 4.1: Summary statistics for baseline demographics

	Full age-eligible sample						Trimmed sample (propensity score 0.2 – 0.8)					
	Received SS (N[1]=429)		No SS (N[1]=1,242)		Norm diff[2]	Diff in means[3]	Received SS (N[1]=324)		No SS (N[1]=580)		Norm diff[2]	Diff in means[3]
	Mean	SD	Mean	SD			Mean	SD	Mean	SD		
Age at 2016	68.86	3.32	68.04	2.66	0.27	0.83***	69.05	3.36	68.67	2.86	0.12	0.39*
Married	0.69	0.46	0.77	0.42	-0.18	-0.08***	0.66	0.47	0.73	0.45	-0.15	-0.07**
Male	0.40	0.49	0.54	0.50	-0.29	-0.15***	0.35	0.48	0.40	0.49	-0.11	-0.05
Chinese	0.87	0.33	0.87	0.34	0.00	0.00	0.87	0.34	0.90	0.30	-0.10	-0.03
Malay	0.07	0.25	0.07	0.25	0.00	0.00	0.07	0.25	0.06	0.24	0.03	0.01
Indian	0.05	0.22	0.05	0.22	0.01	0.00	0.05	0.22	0.03	0.18	0.09	0.02
No formal schooling	0.17	0.38	0.12	0.33	0.13	0.05**	0.17	0.38	0.16	0.36	0.05	0.02
Primary schooling	0.35	0.48	0.26	0.44	0.20	0.09***	0.41	0.49	0.37	0.48	0.07	0.04
Secondary schooling	0.36	0.48	0.41	0.49	-0.10	-0.05*	0.35	0.48	0.36	0.48	-0.01	-0.01
Lives in 1-room flat	0.05	0.22	0.01	0.11	0.22	0.04***	0.05	0.22	0.02	0.15	0.15	0.03**
Lives in 2-room flat	0.05	0.22	0.02	0.15	0.15	0.03**	0.05	0.22	0.03	0.17	0.11	0.02
Lives in 3-room flat	0.33	0.47	0.21	0.40	0.27	0.12***	0.39	0.49	0.32	0.47	0.16	0.07**
Lives in 4-room flat	0.36	0.48	0.37	0.48	-0.03	-0.01	0.36	0.48	0.42	0.49	-0.11	-0.05
Lives in 5-room flat	0.16	0.37	0.29	0.45	-0.32	-0.13***	0.10	0.30	0.15	0.36	-0.15	-0.05**
Owns home	0.83	0.38	0.88	0.33	-0.15	-0.05**	0.82	0.38	0.84	0.37	-0.05	-0.02
No. of hh members	2.95	1.49	3.27	1.51	-0.21	-0.32***	2.82	1.42	3.05	1.43	-0.16	-0.23**
No. of total children	2.18	1.43	2.15	1.25	0.02	0.03	2.16	1.47	2.21	1.31	-0.03	-0.05
Retirement preparedness[4]	2.17	0.90	2.33	0.88	-0.19	-0.17***	2.13	0.90	2.14	0.83	-0.02	-0.01

Notes:

[1] N refers to number of respondents.

[2] Normalised differences are computed as in Imbens (2015) (as the difference in means standardised by the square root of the mean variance of both groups)

[3] ***, **, * represent statistical significance at the 10%, 5% and 1% levels respectively

[4] This is a self-assessment on preparedness for retirement, captured on a scale of 1 to 5, with a higher value representing greater preparedness. This was captured during the baseline survey, which was conducted before the announcement of details on the Silver Support Scheme.

Table 4.2: Summary statistics for pre-announcement dependent variables[1]
(Trimmed sample with 0.20-0.80 propensity score)

	Age-eligible for Silver Support (age 65 and above)					
	Received SS			Didn't receive SS		
	N[2]	Mean	SD	N[2]	Mean	SD
Life satisfaction	1,947	3.51	0.80	3,593	3.54	0.78
Social / family satisfaction	1,944	3.65	0.74	3,594	3.66	0.71
Daily activities satisfaction	1,622	3.52	0.73	3,009	3.47	0.79
Household income satisfaction	1,944	3.18	0.86	3,591	3.19	0.84
Economic satisfaction	1,948	3.19	0.85	3,591	3.21	0.83
Health satisfaction	1,947	3.26	0.90	3,594	3.29	0.91
Self-rated health condition	1,947	2.52	0.89	3,592	2.60	0.88

Notes:
[1] These variables are rated from 1(worst) to 5(best). Options for the first 6 variables are "very dissatisfied", "dissatisfied", "neither satisfied nor dissatisfied", "satisfied", and "very satisfied", while the options for the last question are "poor", "fair", "good", "very good" and "excellent".
[2] N refers to the number of observations at the respondent-wave level.

Table 4.3: Overall impact of the Silver Support Scheme on subjective well-being

VARIABLES	(1) Life satis[fn]	(2) Social satis[fn]	(3) Daily activities satis[fn]	(4) HH income satis[fn]	(5) Economic satis[fn]	(6) Health satis[fn]	(7) Self-rated health cond
Received SS × Jan	0.0428	0.0378	0.00955	0.0121	-0.00109	-0.00428	-0.00601
	(0.0427)	(0.0403)	(0.0421)	(0.0418)	(0.0416)	(0.0436)	(0.0446)
Received SS × Feb	0.0165	-0.0116	0.0345	0.0719	0.0631	0.0425	0.0266
	(0.0435)	(0.0427)	(0.0474)	(0.0491)	(0.0476)	(0.0463)	(0.0462)
Received SS × Mar	0.0705	0.0174	0.0949**	0.0718	0.109**	0.0635	0.0721
	(0.0437)	(0.0410)	(0.0450)	(0.0454)	(0.0467)	(0.0508)	(0.0448)
Received SS × announce-to-disb	0.0858**	0.0643*	0.0226	0.0851**	0.0661*	-0.00177	0.0194
	(0.0359)	(0.0328)	(0.0352)	(0.0368)	(0.0349)	(0.0366)	(0.0352)
Received SS × post- disb	0.0892**	0.0695*	0.0549	0.113***	0.0825**	0.0383	0.0556
	(0.0409)	(0.0366)	(0.0400)	(0.0391)	(0.0381)	(0.0380)	(0.0380)
Testing if difference between β_{ann} and β_{disb} is statistically different from zero							
p-values	0.887	0.819	0.177	0.295	0.522	0.138	0.198
Mean	3.51	3.65	3.52	3.18	3.19	3.26	2.52
S.D.	0.80	0.74	0.73	0.86	0.85	0.90	0.89
Observations	12,652	12,651	11,742	12,646	12,650	12,651	12,649
R-squared	0.675	0.659	0.691	0.688	0.685	0.727	0.711

Notes:
[1] Standard errors clustered at the household level in parentheses. ***, ** and * represent statistical significance at the 1%, 5% and 10% level of significance respectively.
[2] Dependent variables shown at the top of each column are measured on a scale of 1 – 5. 1 represents the worst ("very dissatisfied" or "poor") and 5 the best ("very satisfied" or "excellent").
[3] Results are estimates of coefficients in eq(1). The sample is restricted to respondents who are age-eligible for SSS (i.e. aged 65 and above in 2016), live in public housing flats, and with a propensity score of 0.2 – 0.8.
[4]Mean and standard deviation statistics are based on pre-announcement levels of the dependent variable for respondents who received SSS payouts.

Table 4.4: Silver Support payouts and subjective well-being – heterogeneous effects by financial preparation for retirement

VARIABLES	(1) Life satis^{fn}	(2) Social satis^{fn}	(3) Daily activities satis^{fn}	(4) HH income satis^{fn}	(5) Economic satis^{fn}	(6) Health satis^{fn}	(7) Self-rated health cond
Received SS × announce-to-disb	0.164**	0.0459	0.0171	0.176**	0.123*	0.0653	0.160**
	(0.0687)	(0.0599)	(0.0599)	(0.0700)	(0.0629)	(0.0676)	(0.0674)
Received SS × post-disb	0.257***	0.0543	0.0793	0.246***	0.226***	0.130*	0.194***
	(0.0846)	(0.0732)	(0.0799)	(0.0785)	(0.0770)	(0.0721)	(0.0733)
Received SS × announce-to-disb× retirement prep	-0.0372	0.00876	0.00254	-0.0434*	-0.0271	-0.0319	-0.0671***
	(0.0234)	(0.0221)	(0.0207)	(0.0244)	(0.0224)	(0.0260)	(0.0257)
Received SS × post-disb× retirement prep	-0.0793**	0.00722	-0.0115	-0.0631**	-0.0677**	-0.0434	-0.0655**
	(0.0310)	(0.0274)	(0.0295)	(0.0299)	(0.0299)	(0.0273)	(0.0302)
Leads significant?	NO	NO	YES (MAR**)	NO	YES (MAR**)	NO	NO
Mean	3.51	3.65	3.52	3.18	3.19	3.26	2.52
S.D.	0.80	0.74	0.73	0.86	0.85	0.90	0.89
Observations	12,652	12,651	11,742	12,646	12,650	12,651	12,649
R-squared	0.675	0.659	0.691	0.689	0.686	0.728	0.712

Notes:
[1] Standard errors clustered at the household level in parentheses. ***, ** and * represent statistical significance at the 1%, 5% and 10% level of significance respectively.
[2] Dependent variables shown at the top of each column are measured on a scale of 1 – 5. 1 represents the worst (“very dissatisfied” or “poor”) and 5 the best (“very satisfied” or “excellent”).
[3] Results are estimates of coefficients in eq(2), where announcement and disbursement variables are also interacted with one’s self-assessed financial preparedness for retirement, which is rated from 1(Poor) to 5(Excellent). The sample is restricted to respondents who are age-eligible for SSS (i.e. aged 65 and above in 2016), live in public housing flats, and with a propensity score of 0.2 – 0.8.
[4]Mean and standard deviation statistics are based on pre-announcement levels of the dependent variable for respondents who received SSS payouts.

Table 4.5: Silver Support payouts and subjective well-being – heterogeneous effects by baseline assets quintile

VARIABLES	(1) Life satis^fn	(2) Social satis^fn	(3) Daily activities satis^fn	(4) HH income satis^fn	(5) Economic satis^fn	(6) Health satis^fn	(7) Self-rated health cond
Received SS × announce-to-disb	0.0840	0.0484	0.0644	0.0966*	0.0566	0.0363	0.0794
	(0.0561)	(0.0495)	(0.0511)	(0.0572)	(0.0548)	(0.0597)	(0.0543)
Received SS × post-disb	0.0723	0.0224	0.0621	0.0714	0.0711	0.0742	0.130**
	(0.0660)	(0.0603)	(0.0602)	(0.0631)	(0.0580)	(0.0595)	(0.0651)
Received SS × announce-to-disb× assets quintile	0.00130	0.00905	-0.0131	-0.00727	0.00275	-0.0147	-0.0243
	(0.0149)	(0.0145)	(0.0148)	(0.0173)	(0.0160)	(0.0189)	(0.0172)
Received SS × post-disb× assets quintile	0.00455	0.0157	-0.00118	0.0117	0.00362	-0.0185	-0.0326*
	(0.0188)	(0.0175)	(0.0187)	(0.0185)	(0.0164)	(0.0195)	(0.0193)
Leads significant?	NO	NO	YES (MAR**)	NO	YES (MAR**)	NO	NO
Mean	3.51	3.65	3.52	3.18	3.19	3.26	2.52
S.D.	0.80	0.74	0.73	0.86	0.85	0.90	0.89
Observations	11,893	11,892	11,092	11,887	11,892	11,892	11,892
R-squared	0.685	0.665	0.694	0.695	0.692	0.736	0.722

Notes:
[1] Standard errors clustered at the household level in parentheses. ***, ** and * represent statistical significance at the 1%, 5% and 10% level of significance respectively.
[2] Dependent variables shown at the top of each column are measured on a scale of 1 – 5. 1 represents the worst ("very dissatisfied" or "poor") and 5 the best ("very satisfied" or "excellent").
[3] Results are estimates of coefficients in eq(2), where announcement and disbursement variables are also interacted with one's net assets (captured before announcement and expressed in quintiles which are computed off respondents who are aged 65 in 2016 and who live in public housing flats). The sample is restricted to respondents who are age-eligible for SSS (i.e. aged 65 and above in 2016), live in public housing flats, and with a propensity score of 0.2 – 0.8. This sample is smaller because data on assets was collected only once before announcement, and not everyone responded in that wave.
[4]Mean and standard deviation statistics are based on pre-announcement levels of the dependent variable for respondents who received SSS payouts.

Table 4.6: Silver Support payouts and subjective well-being – heterogeneous effects by recipient identity

VARIABLES	(1) Life satis[fn]	(2) Social satis[fn]	(3) Daily activities satis[fn]	(4) HH income satis[fn]	(5) Economic satis[fn]	(6) Health satis[fn]	(7) Self-rated health cond
Only respondent rcv SS × announce-to-disb	0.139***	0.105**	0.0695	0.144**	0.112**	0.0933	0.0118
	(0.0493)	(0.0484)	(0.0564)	(0.0608)	(0.0554)	(0.0631)	(0.0628)
Only respondent rcv SS × post-disb	0.116*	0.127**	0.0571	0.167***	0.0853	0.0832	0.0682
	(0.0622)	(0.0569)	(0.0674)	(0.0625)	(0.0644)	(0.0737)	(0.0708)
Only spouse rcv SS × announce-to-disb	0.0216	0.0115	0.0168	0.0991	0.118	-0.0274	-0.0940
	(0.0705)	(0.0616)	(0.0681)	(0.0718)	(0.0766)	(0.0682)	(0.0758)
Only spouse rcv SS × post-disb	0.0226	0.0490	0.0380	0.0101	-0.00816	-0.00773	0.0348
	(0.0729)	(0.0636)	(0.0765)	(0.0915)	(0.0774)	(0.0799)	(0.0754)
Both rcv SS × announce-to-disb	0.0827	0.0877*	-0.00178	0.0615	0.0801	-0.00784	0.0191
	(0.0583)	(0.0484)	(0.0556)	(0.0584)	(0.0542)	(0.0542)	(0.0497)
Both rcv SS × post-disb	0.146**	0.136**	0.0928	0.134**	0.146***	0.0407	0.0775
	(0.0661)	(0.0537)	(0.0631)	(0.0597)	(0.0556)	(0.0562)	(0.0555)
Leads significant?	NO	YES (MAR*)	NO	YES (FEB*)	NO	NO	NO
Mean	3.51	3.65	3.52	3.18	3.19	3.26	2.52
S.D.	0.80	0.74	0.73	0.86	0.85	0.90	0.89
Observations	8,701	8,700	8,080	8,695	8,699	8,702	8,698
R-squared	0.679	0.670	0.697	0.697	0.704	0.723	0.706

Notes:
[1] Standard errors clustered at the household level in parentheses. ***, ** and * represent statistical significance at the 1%, 5% and 10% level of significance respectively.
[2] Dependent variables shown at the top of each column are measured on a scale of 1 – 5. 1 represents the worst ("very dissatisfied" or "poor") and 5 the best ("very satisfied" or "excellent").
[3] Results are estimates of coefficients in eq(4), where announcement and disbursement variables are also interacted with the identity of recipient. The sample is restricted to respondents who are age-eligible for SSS (i.e. aged 65 and above in 2016), live in public housing flats, and with a propensity score of 0.2 – 0.8. This sample is smaller because only respondents who are married are included.
[4]Mean and standard deviation statistics are based on pre-announcement levels of the dependent variable for respondents who received SSS payouts.

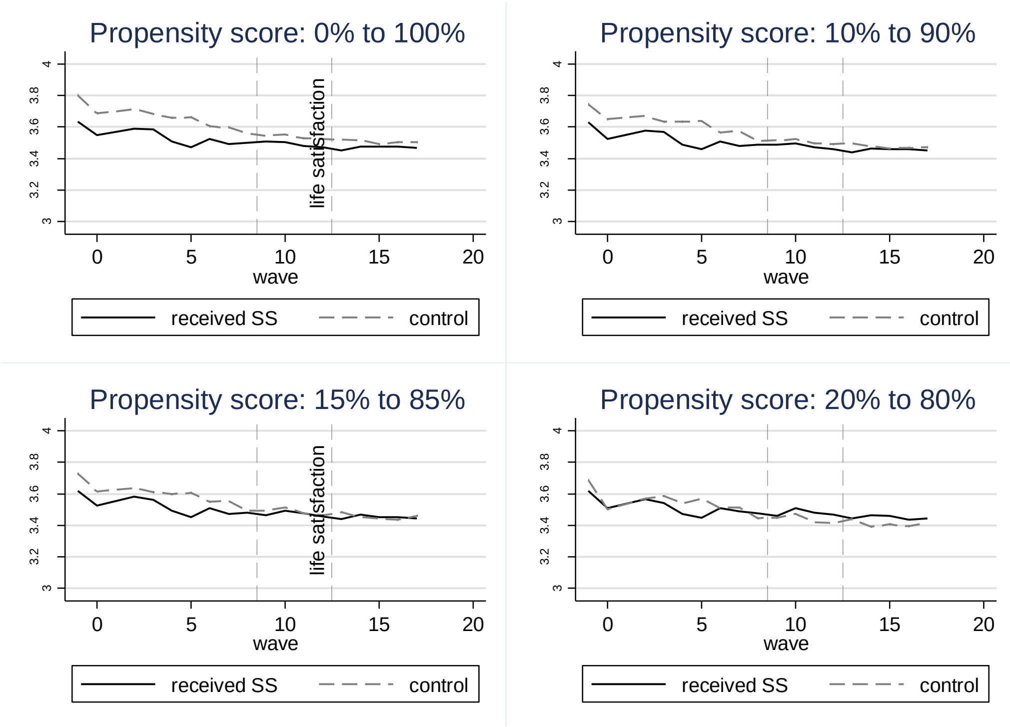
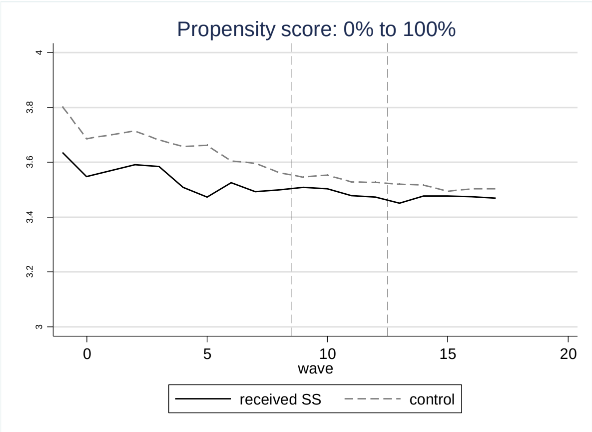
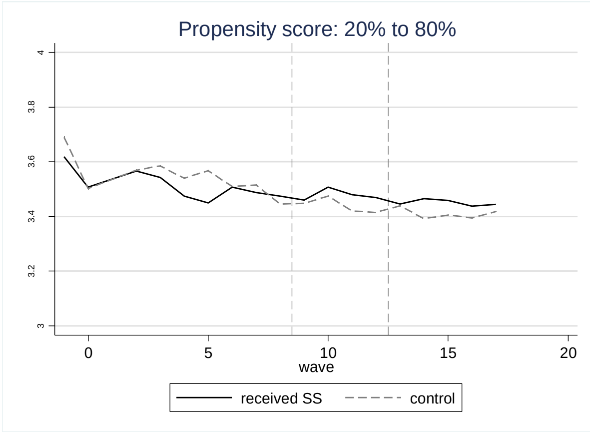


Figure 4.1: Trends for life satisfaction, for samples trimmed based on increasingly narrower ranges of propensity score

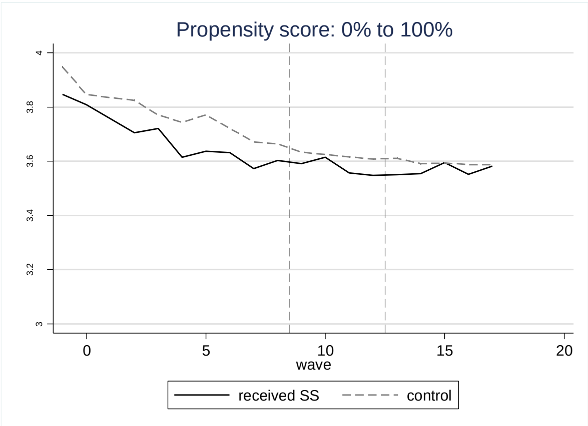
Before trimming / **After trimming**

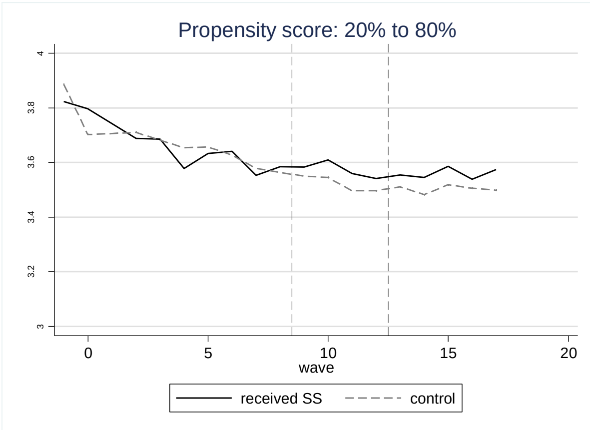
Life satisfaction



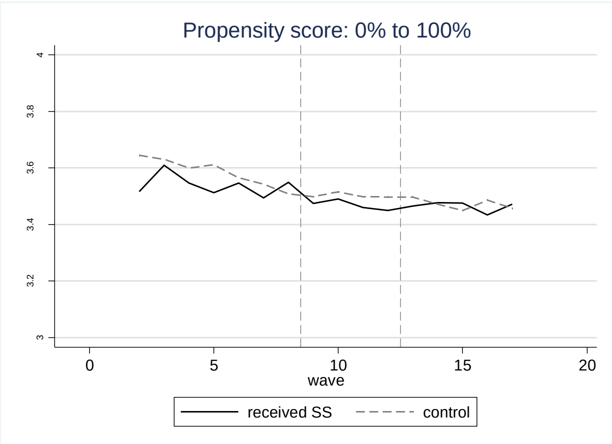


Social satisfaction





Daily activities satisfaction



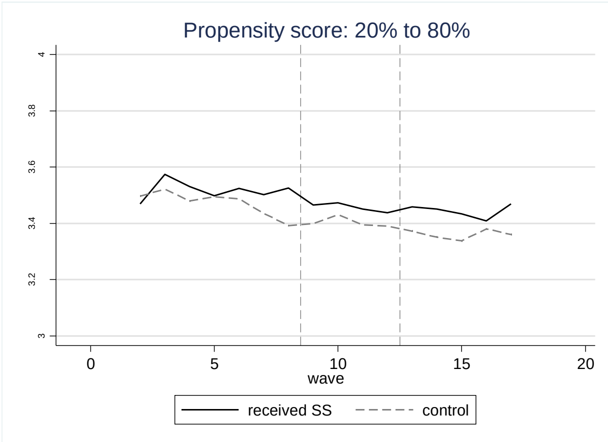


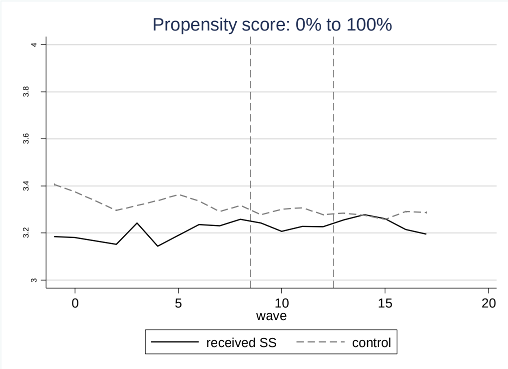
Figure 4.2: Trends for life, social, and daily activities satisfaction[95]

---

95 The periods to the left of the first and second dotted vertical lines refer to the pre-announcement and pre-disbursement periods respectively.

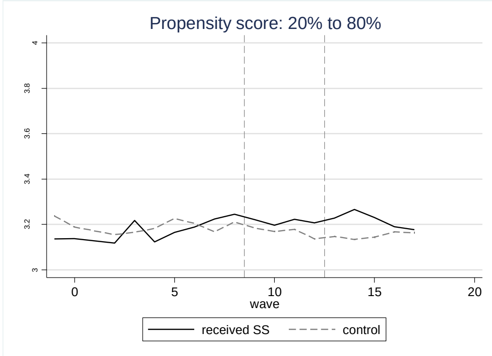
Before trimming

Household income satisfaction

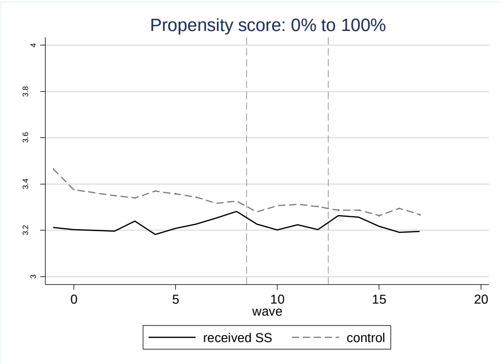


After trimming

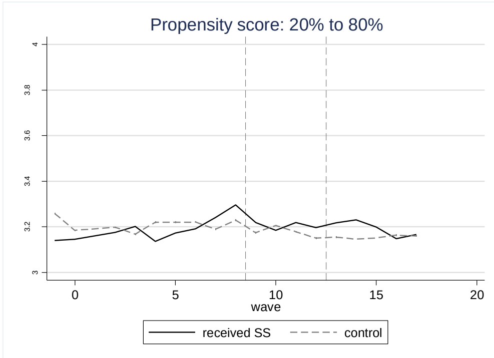
Household income satisfaction



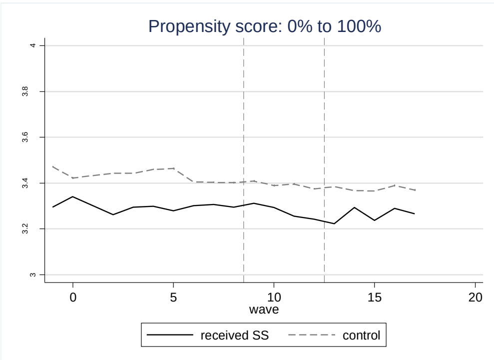
Economic satisfaction



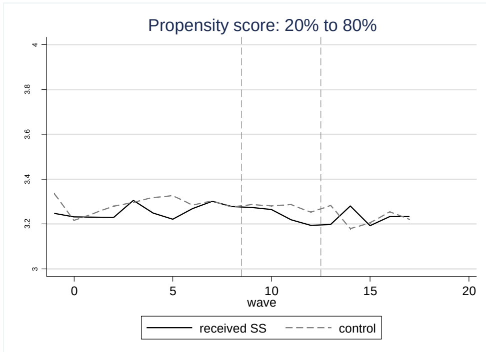
Economic satisfaction



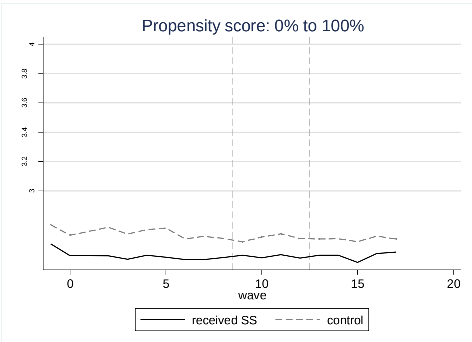
Health satisfaction



Health satisfaction



Self-rated health condition



Self-rated health condition

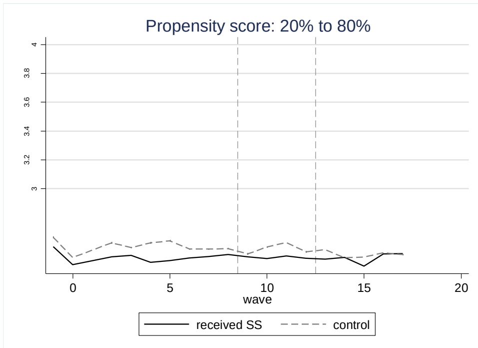


Figure 4.3: Trends for household income, economic, health satisfaction as well as health condition[96]

[96] The periods to the left of the first and second dotted vertical lines refer to the pre-announcement and pre-disbursement periods respectively.

Before trimming **After trimming**

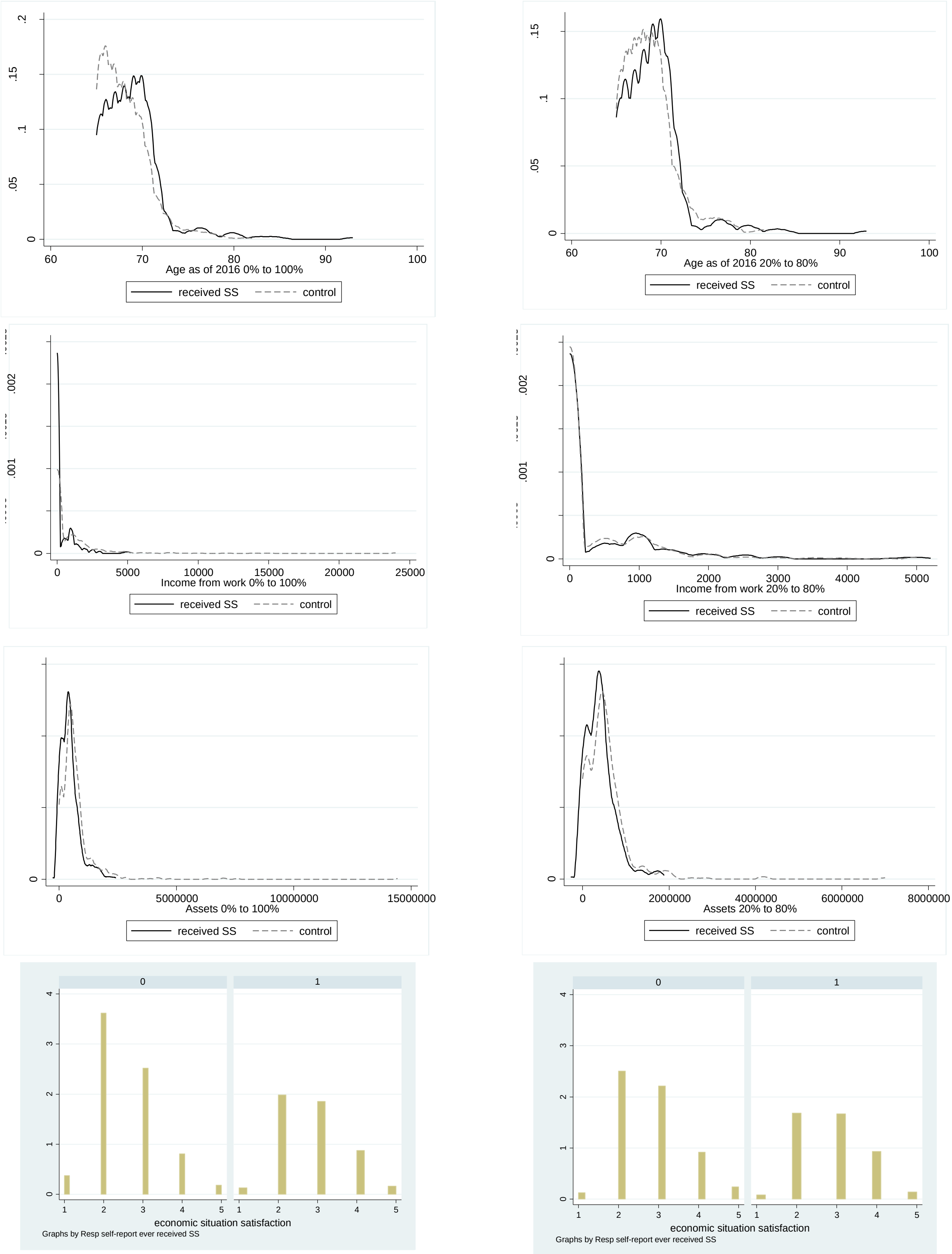


Figure 4.4: Density plots / histograms of key demographics before (left panel) and after (right panel) trimming[97]

[97] Note the change in scale of the horizontal axis after trimming.

Appendix A

Table A1: Summary statistics for baseline demographics

	Trimmed sample (propensity score 0.10 – 0.90)						Trimmed sample (propensity score 0.15 – 0.85)					
	Received SS (N[1]=406)		No SS (N[1]=929)		Norm diff[2]	Diff in means[3]	Received SS (N[1]=369)		No SS (N[1]=739)		Norm diff[2]	Diff in means[3]
	Mean	SD	Mean	SD			Mean	SD	Mean	SD		
Age at 2016	68.93	3.37	68.40	2.72	0.17	0.52***	68.95	3.32	68.55	2.74	0.13	0.40**
Married	0.68	0.47	0.75	0.43	-0.16	-0.07***	0.67	0.47	0.74	0.44	-0.16	-0.07**
Male	0.38	0.49	0.48	0.50	-0.21	-0.10***	0.36	0.48	0.44	0.50	-0.15	-0.08**
Chinese	0.87	0.33	0.89	0.31	-0.06	-0.02	0.87	0.34	0.89	0.31	-0.08	-0.03
Malay	0.07	0.25	0.06	0.23	0.04	0.01	0.07	0.25	0.06	0.24	0.03	0.01
Indian	0.05	0.22	0.05	0.21	0.02	0.00	0.05	0.23	0.04	0.20	0.06	0.01
No formal schooling	0.17	0.38	0.14	0.35	0.10	0.04	0.18	0.38	0.15	0.36	0.06	0.02
Primary schooling	0.36	0.48	0.32	0.47	0.08	0.04	0.38	0.49	0.35	0.48	0.06	0.03
Secondary schooling	0.37	0.48	0.38	0.48	-0.02	-0.01	0.35	0.48	0.36	0.48	-0.02	-0.01
Lives in 1-room flat	0.05	0.22	0.02	0.13	0.20	0.04***	0.05	0.22	0.02	0.14	0.17	0.03**
Lives in 2-room flat	0.05	0.21	0.02	0.16	0.12	0.02*	0.05	0.22	0.03	0.16	0.11	0.02*
Lives in 3-room flat	0.34	0.47	0.25	0.43	0.20	0.09***	0.36	0.48	0.29	0.45	0.16	0.07**
Lives in 4-room flat	0.36	0.48	0.40	0.49	-0.09	-0.04	0.37	0.48	0.41	0.49	-0.08	-0.04
Lives in 5-room flat	0.15	0.36	0.23	0.42	-0.21	-0.08***	0.13	0.34	0.19	0.39	-0.17	-0.06***
Owns home	0.83	0.38	0.87	0.34	-0.11	-0.04*	0.82	0.38	0.86	0.35	-0.10	-0.04
No. of hh members	2.91	1.47	3.16	1.43	-0.17	-0.24***	2.85	1.42	3.12	1.42	-0.19	-0.27***
No. of total children	2.18	1.42	2.18	1.26	0.00	0.00	2.19	1.45	2.22	1.29	-0.03	-0.03
Retirement preparedness[4]	2.15	0.91	2.26	0.85	-0.13	-0.11**	2.12	0.89	2.23	0.86	-0.12	-0.10*

Notes:

[1] N refers to number of respondents.

[2] Normalised differences are computed as in Imbens (2015) (as the difference in means standardised by the square root of the mean variance of both groups)

[3] ***, **, * represent statistical significance at the 10%, 5% and 1% levels respectively

[4] This is a self-assessment on preparedness for retirement, captured on a scale of 1 to 5, with a higher value representing greater preparedness. This was captured during the baseline survey, which was conducted before the announcement of details on the Silver Support Scheme.

Table A2: Silver Support payouts and subjective well-being – heterogeneous effects by flat-type and financial preparedness for retirement

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	Life satis[fn]	Social satis[fn]	Daily activities satis[fn]	HH income satis[fn]	Economic satis[fn]	Health satis[fn]	Self-rated health cond
Received SS × ann-to-disb × 1/2 rm × retirement prep	-0.103*	-0.0445	0.0342	0.0312	0.0600	-0.160	-0.0596
	(0.0556)	(0.120)	(0.0519)	(0.128)	(0.0980)	(0.125)	(0.0735)
Received SS × ann-to- disb × 3 rm × retirement prep	-0.0301	-0.00160	0.0371	-0.0676*	-0.0445	-0.0596	-0.143***
	(0.0363)	(0.0353)	(0.0418)	(0.0364)	(0.0386)	(0.0534)	(0.0472)
Received SS × ann-to- disb × 4 rm × retirement prep	-0.0868**	0.00709	-0.0593*	-0.0917**	-0.0573	0.0361	-0.00271
	(0.0405)	(0.0397)	(0.0349)	(0.0447)	(0.0455)	(0.0338)	(0.0506)
Received SS × ann-to- disb × 5 rm × retirement prep	0.0663	0.0496	0.0639	0.0330	0.000767	0.0202	-0.0171
	(0.0666)	(0.0477)	(0.0392)	(0.0469)	(0.0414)	(0.0411)	(0.0284)
Received SS × post- disb × 1/2 rm × retirement prep	-0.173*	-0.0415	0.0505	-0.0343	-0.0397	-0.219**	-0.133
	(0.0991)	(0.117)	(0.0873)	(0.102)	(0.0778)	(0.101)	(0.106)
Received SS × post- disb × 3 rm × retirement prep	-0.0796*	-0.0646	0.0262	-0.116**	-0.126**	-0.0529	-0.125**
	(0.0467)	(0.0472)	(0.0458)	(0.0452)	(0.0493)	(0.0435)	(0.0489)
Received SS × post- disb × 4 rm × retirement prep	-0.116**	0.0363	-0.0966**	-0.117**	-0.107**	-0.00333	-0.0288
	(0.0499)	(0.0470)	(0.0468)	(0.0467)	(0.0499)	(0.0468)	(0.0553)
Received SS × post- disb × 5 rm × retirement prep	0.0645	0.0640	0.137***	0.111	0.0982	0.0549	-0.00343
	(0.0681)	(0.0534)	(0.0528)	(0.0752)	(0.0676)	(0.0591)	(0.0689)
Leads significant?	NO	NO	YES (MAR**)	NO	YES (MAR**)	NO	NO
Mean	3.51	3.65	3.52	3.18	3.19	3.26	2.52
S.D.	0.80	0.74	0.73	0.86	0.85	0.90	0.89
Observations	12,652	12,651	11,742	12,646	12,650	12,651	12,649
R-squared	0.677	0.660	0.692	0.689	0.687	0.728	0.713

Notes:
[1] Standard errors clustered at the household level in parentheses. ***, ** and * represent statistical significance at the 1%, 5% and 10% level of significance respectively.
[2] Dependent variables shown at the top of each column are measured on a scale of 1 – 5. 1 represents the worst ("very dissatisfied" or "poor") and 5 the best ("very satisfied" or "excellent").
[3] Results are estimates of coefficients in eq(3), where announcement and disbursement variables are interacted not only with one's self-assessed financial preparedness for retirement (which is rated from 1(Poor) to 5(Excellent)) but also flat-type . The sample is restricted to respondents who are age-eligible for SSS (i.e. aged 65 and above in 2016), live in public housing flats, and with a propensity score of 0.2 – 0.8.
[4]Mean and standard deviation statistics are based on pre-announcement levels of the dependent variable for respondents who received SSS payouts.

Table A3: Robustness checks for overall life satisfaction

VARIABLES	(1) No leads	(2) Balanced	(3) Ethnic trends	(4) Flattype trends	(5) WIS/ GST	(6) Abadie	(7) 1-1 matching
Received SS × Jan	-	0.0360	0.0376	0.0294	0.0274	0.0283	-
		(0.0422)	(0.0422)	(0.0417)	(0.0467)	(0.0345)	
Received SS × Feb	-	0.0111	0.0147	0.00407	0.0158	-0.0281	-
		(0.0435)	(0.0438)	(0.0436)	(0.0468)	(0.0399)	
Received SS × Mar	-	0.0645	0.0691	0.0735*	0.0499	0.0375	-
		(0.0437)	(0.0443)	(0.0430)	(0.0471)	(0.0368)	
Received SS × Jan - Mar	-	-	-	-	-	-	0.0201
							P=0.58
Received SS × announce-to-disb	0.0685**	0.0819**	0.0841**	0.0729**	0.0911**	0.0460	0.0530
	(0.0289)	(0.0359)	(0.0363)	(0.0352)	(0.0412)	(0.0301)	P=0.11
Received SS × post-disb	0.0719**	0.0824**	0.0858**	0.0762*	0.0938**	0.0831**	0.0681*
	(0.0345)	(0.0409)	(0.0413)	(0.0401)	(0.0475)	(0.0343)	P=0.07
Observations	12,652	12,519	12,652	12,652	9,959	1,592	648
R-squared	0.675	0.679	0.677	0.678	0.693	-	-

Notes:

[1] Standard errors (SE) clustered at the household level in parentheses for columns (1) – (6); SE adjusted to account for uncertainty from propensity score estimation as in Abadie (2005) for column (7); p-values from a permutation test in parentheses in column (8). ***, ** and * represent statistical significance at the 1%, 5% and 10% level of significance respectively.

[2] The dependent variable is measured on a scale of 1 – 5. 1 represents the worst ("very dissatisfied") and 5 the best ("very satisfied").

[3] Columns (1) – (7) show results from additional robustness checks carried out. These checks are: (1) removal of the lead terms; (2) restricting the sample to a "balanced" panel, where each individual has at least one observation in the pre-announcement, announcement-to-disbursement, and post-disbursement periods; (3) allowing for ethnicity-specific time fixed effects; (4) allowing for flat-type-specific time fixed effects; (5) adding receipt of additional welfare payments (Workfare Income Supplement; GST Vouchers) as a control—sample is smaller because this data is not collected every wave and not everyone responds every wave; (6) reweighting each observation by their propensity of receiving SSS as in Abadie (2005); (7) DiD matching with a 1-1 nearest neighbour match. Eq(1) describes the baseline model used in these checks. The sample is restricted to those age-eligible for the Silver Support Scheme (i.e. aged 65 and above) as of end-Sep 2016 (the month of the most recent Silver Support payout in our data), and with a propensity score of 0.2 – 0.8 for checks in columns (1) – (5). The full/untrimmed sample is used in column (6). The number of observations in columns (1) – (5) refer to the number of individual-month observations; those in columns (6) and (7) refer to number of respondents.

Table A4: Robustness checks for social and family life satisfaction

VARIABLES	(1) No leads	(2) Balanced	(3) Ethnic trends	(4) Flattype trends	(5) WIS/ GST	(6) Abadie	(7) 1-1 matching
Received SS × Jan	-	0.0322	0.0296	0.0345	0.0263	0.00710	-
		(0.0398)	(0.0403)	(0.0406)	(0.0426)	(0.0337)	
Received SS × Feb	-	-0.0125	-0.0170	-0.00839	0.00159	-0.0488	-
		(0.0428)	(0.0423)	(0.0434)	(0.0455)	(0.0404)	
Received SS × Mar	-	0.0153	0.0113	0.0223	-0.00389	-0.0129	-
		(0.0410)	(0.0407)	(0.0401)	(0.0438)	(0.0353)	
Received SS × Jan - Mar	-	-	-	-	-	-	-0.0254
							P=0.44
Received SS × announce-to-disb	0.0581**	0.0632*	0.0594*	0.0570*	0.0667*	0.00651	0.0141
	(0.0264)	(0.0329)	(0.0328)	(0.0328)	(0.0367)	(0.0299)	P=0.62
Received SS × post-disb	0.0633**	0.0675*	0.0633*	0.0627*	0.0498	0.0523*	0.0458
	(0.0312)	(0.0366)	(0.0368)	(0.0362)	(0.0420)	(0.0313)	P=0.17
Observations	12,651	12,518	12,651	12,651	9,959	1,592	648
R-squared	0.659	0.661	0.661	0.662	0.677	-	-

Notes:

[1] Standard errors (SE) clustered at the household level in parentheses for columns (1) – (6); SE adjusted to account for uncertainty from propensity score estimation as in Abadie (2005) for column (7); p-values from a permutation test in parentheses in column (8). ***, ** and * represent statistical significance at the 1%, 5% and 10% level of significance respectively.

[2] The dependent variable is measured on a scale of 1 – 5. 1 represents the worst ("very dissatisfied") and 5 the best ("very satisfied").

[3] Columns (1) – (7) show results from additional robustness checks carried out. These checks are: (1) removal of the lead terms; (2) restricting the sample to a "balanced" panel, where each individual has at least one observation in the pre-announcement, announcement-to-disbursement, and post-disbursement periods; (3) allowing for ethnicity-specific time fixed effects; (4) allowing for flat-type-specific time fixed effects; (5) adding receipt of additional welfare payments (Workfare Income Supplement; GST Vouchers) as a control—sample is smaller because this data is not collected every wave and not everyone responds every wave; (6) reweighting each observation by their propensity of receiving SSS as in Abadie (2005); (7) DiD matching with a 1-1 nearest neighbour match. Eq(1) describes the baseline model used in these checks. The sample is restricted to those age-eligible for the Silver Support Scheme (i.e. aged 65 and above) as of end-Sep 2016 (the month of the most recent Silver Support payout in our data), and with a propensity score of 0.2 – 0.8 for checks in columns (1) – (5). The full/untrimmed sample is used in column (6). The number of observations in columns (1) – (5) refer to the number of individual-month observations; those in columns (6) and (7) refer to number of respondents.

Table A5: Robustness checks for household income satisfaction

VARIABLES	(1) No leads	(2) Balanced	(3) Ethnic trends	(4) Flattype trends	(5) WIS/ GST	(6) Abadie	(7) 1-1 matching
Received SS × Jan	-	0.00623	0.00205	0.0122	-0.0286	0.0394	-
		(0.0413)	(0.0419)	(0.0412)	(0.0454)	(0.0350)	
Received SS × Feb	-	0.0714	0.0621	0.0789*	0.0672	0.0484	-
		(0.0491)	(0.0490)	(0.0473)	(0.0521)	(0.0393)	
Received SS × Mar	-	0.0681	0.0579	0.0718	0.0580	0.0675*	-
		(0.0455)	(0.0454)	(0.0449)	(0.0499)	(0.0388)	
Received SS × Jan - Mar	-	-	-	-	-	-	0.0324
							P=0.36
Received SS × announce-to-disb	0.0654**	0.0865**	0.0753**	0.0823**	0.0931**	0.0623*	0.0712**
	(0.0304)	(0.0369)	(0.0371)	(0.0367)	(0.0405)	(0.0321)	P=0.05
Received SS × post-disb	0.0929***	0.114***	0.106***	0.105***	0.122***	0.109***	0.102***
	(0.0333)	(0.0393)	(0.0392)	(0.0389)	(0.0437)	(0.0332)	P=0.01
Observations	12,646	12,506	12,646	12,646	9,953	1,590	648
R-squared	0.688	0.691	0.691	0.691	0.710	-	-

Notes:

[1] Standard errors (SE) clustered at the household level in parentheses for columns (1) – (6); SE adjusted to account for uncertainty from propensity score estimation as in Abadie (2005) for column (7); p-values from a permutation test in parentheses in column (8). ***, ** and * represent statistical significance at the 1%, 5% and 10% level of significance respectively.

[2] The dependent variable is measured on a scale of 1 – 5. 1 represents the worst ("very dissatisfied") and 5 the best ("very satisfied").

[3] Columns (1) – (7) show results from additional robustness checks carried out. These checks are: (1) removal of the lead terms; (2) restricting the sample to a "balanced" panel, where each individual has at least one observation in the pre-announcement, announcement-to-disbursement, and post-disbursement periods; (3) allowing for ethnicity-specific time fixed effects; (4) allowing for flat-type-specific time fixed effects; (5) adding receipt of additional welfare payments (Workfare Income Supplement; GST Vouchers) as a control—sample is smaller because this data is not collected every wave and not everyone responds every wave; (6) reweighting each observation by their propensity of receiving SSS as in Abadie (2005); (7) DiD matching with a 1-1 nearest neighbour match. Eq(1) describes the baseline model used in these checks. The sample is restricted to those age-eligible for the Silver Support Scheme (i.e. aged 65 and above) as of end-Sep 2016 (the month of the most recent Silver Support payout in our data), and with a propensity score of 0.2 – 0.8 for checks in columns (1) – (5). The full/untrimmed sample is used in column (6). The number of observations in columns (1) – (5) refer to the number of individual-month observations; those in columns (6) and (7) refer to number of respondents.

Table A6: Robustness checks for economic satisfaction

VARIABLES	(1) No leads	(2) Balanced	(3) Ethnic trends	(4) Flattype trends	(5) WIS/ GST	(6) Abadie	(7) 1-1 matching
Received SS × Jan	-	-0.00844	-0.00825	-0.00151	-0.0306	0.0241	-
		(0.0412)	(0.0416)	(0.0421)	(0.0454)	(0.0396)	
Received SS × Feb	-	0.0610	0.0517	0.0588	0.0455	0.0382	-
		(0.0475)	(0.0480)	(0.0470)	(0.0515)	(0.0412)	
Received SS × Mar	-	0.106**	0.102**	0.108**	0.119**	0.0691*	-
		(0.0468)	(0.0477)	(0.0465)	(0.0523)	(0.0394)	
Received SS × Jan - Mar	-	-	-	-	-	-	0.0306
							P=0.40
Received SS × announce-to-disb	0.0443	0.0665*	0.0601*	0.0588*	0.0754*	0.0267	0.0454
	(0.0281)	(0.0350)	(0.0352)	(0.0348)	(0.0399)	(0.0307)	P=0.21
Received SS × post-disb	0.0608*	0.0854**	0.0787**	0.0706*	0.105**	0.0650*	0.0768**
	(0.0320)	(0.0383)	(0.0383)	(0.0380)	(0.0425)	(0.0332)	P=0.05
Observations	12,650	12,517	12,650	12,650	9,959	1,592	648
R-squared	0.685	0.688	0.687	0.689	0.706	-	-

Notes:

[1] Standard errors (SE) clustered at the household level in parentheses for columns (1) – (6); SE adjusted to account for uncertainty from propensity score estimation as in Abadie (2005) for column (7); p-values from a permutation test in parentheses in column (8). ***, ** and * represent statistical significance at the 1%, 5% and 10% level of significance respectively.

[2] The dependent variable is measured on a scale of 1 – 5. 1 represents the worst ("very dissatisfied") and 5 the best ("very satisfied").

[3] Columns (1) – (7) show results from additional robustness checks carried out. These checks are: (1) removal of the lead terms; (2) restricting the sample to a "balanced" panel, where each individual has at least one observation in the pre-announcement, announcement-to-disbursement, and post-disbursement periods; (3) allowing for ethnicity-specific time fixed effects; (4) allowing for flat-type-specific time fixed effects; (5) adding receipt of additional welfare payments (Workfare Income Supplement; GST Vouchers) as a control—sample is smaller because this data is not collected every wave and not everyone responds every wave; (6) reweighting each observation by their propensity of receiving SSS as in Abadie (2005); (7) DiD matching with a 1-1 nearest neighbour match. Eq(1) describes the baseline model used in these checks. The sample is restricted to those age-eligible for the Silver Support Scheme (i.e. aged 65 and above) as of end-Sep 2016 (the month of the most recent Silver Support payout in our data), and with a propensity score of 0.2 – 0.8 for checks in columns (1) – (5). The full/untrimmed sample is used in column (6). The number of observations in columns (1) – (5) refer to the number of individual-month observations; those in columns (6) and (7) refer to number of respondents.

Table A7: Overall impact of the Silver Support Scheme on subjective well-being for full sample

VARIABLES	(1) Life satisfn	(2) Social satisfn	(3) Daily activities satisfn	(4) HH income satisfn	(5) Economic satisfn	(6) Health satisfn	(7) Self-rated health cond
Received SS × Jan	0.0470	0.00342	0.0390	0.0457	0.0288	0.0120	-0.00775
	(0.0329)	(0.0325)	(0.0323)	(0.0324)	(0.0343)	(0.0348)	(0.0356)
Received SS × Feb	0.0149	-0.0181	-0.00302	0.0682*	0.0645*	0.0430	0.00273
	(0.0348)	(0.0333)	(0.0366)	(0.0377)	(0.0379)	(0.0372)	(0.0362)
Received SS × Mar	0.0582*	0.00581	0.0821**	0.0886**	0.105***	0.0347	0.0401
	(0.0350)	(0.0340)	(0.0342)	(0.0363)	(0.0389)	(0.0407)	(0.0364)
Received SS × announce-to-disb	0.0875***	0.0462*	0.0247	0.0957***	0.0712**	0.0185	0.0252
	(0.0291)	(0.0275)	(0.0276)	(0.0289)	(0.0295)	(0.0298)	(0.0283)
Received SS × post-disb	0.102***	0.0662**	0.0559*	0.133***	0.113***	0.0288	0.0456
	(0.0332)	(0.0300)	(0.0310)	(0.0315)	(0.0315)	(0.0314)	(0.0313)
Mean	3.51	3.65	3.52	3.18	3.19	3.26	2.52
S.D.	0.80	0.74	0.73	0.86	0.85	0.90	0.89
Observations	23,665	23,663	21,983	23,655	23,657	23,662	23,658
R-squared	0.684	0.658	0.693	0.707	0.698	0.720	0.724

Notes:
[1] Standard errors clustered at the household level in parentheses. ***, ** and * represent statistical significance at the 1%, 5% and 10% level of significance respectively.
[2] Dependent variables shown at the top of each column are measured on a scale of 1 – 5. 1 represents the worst (“very dissatisfied” or “poor”) and 5 the best (“very satisfied” or “excellent”).
[3] Results are estimates of coefficients in eq(1). The sample is restricted to respondents who are age-eligible for SSS (i.e. aged 65 and above in 2016), live in public housing flats. This sample is not trimmed based on propensity score.
[4]Mean and standard deviation statistics are based on pre-announcement levels of the dependent variable for respondents who received SSS payouts.

5 A pragmatic, randomized controlled trial of a cardiac hospital-to-home transitional care program

(First authorship shared with with Yanying Chen; other co-authors are: Jessica Sun, Cheng Zhan Chua, Jeffrey Yoo, Thomas Wong, Helen Chen, John Wong, Phillip Phan)

5.1 Introduction

Rehospitalisations are common in healthcare – 30% of Medicare beneficiaries discharged from a hospital were readmitted within 90 days (Jencks et al., 2009). They are costly for hospitals and patients, and a substantial percentage are preventable (Oddone et al., 1996), partly because hospital-to-community transitions are often poorly managed (Coleman, 2003). To reduce unplanned readmissions and costs, hospitals worldwide (including those in the United States (Centers for Medicare and Medicaid Services, 2016)) and Singapore (Wee et al., 2014)) have implemented transitional care programmes (TCPs). However, TCP effectiveness can vary greatly (Naylor et al., 1994; Rich et al., 1995; Naylor et al., 2004; Anderson et al., 2005; Wong et al., 2013) due to implementation challenges (Mistiaen et al., 2007; Naylor et al., 2011; Allen et al., 2014; Kronick et al., 2014; Jayakody et al., 2016). As more hospitals adopt TCPs, which are resource intensive, the need to measure programme effectiveness becomes more important.

In this study, we carry out a pragmatic randomised, controlled trial (Ford & Norrie, 2016) to evaluate the effectiveness of a new nurse-practitioner-led TCP called CareHub, which was piloted in Singapore's National University Hospital (NUH). CareHub merges existing TCPs by offering a protocolised 'one-stop shop' for post-discharge patient follow-ups. We examine a comprehensive set of primary outcomes on healthcare services utilisation, and an extensive list of secondary outcomes, including patients' quality of life, quality of transitional care, and selected clinical outcomes. We contribute to the research on TCPs by investigating if hospitals can further improve patient outcomes by integrating existing transitional care services.

5.2 Data and Methods

5.2.1 CareHub versus Usual Care

Here, we summarise the differences between CareHub (intervention) and usual care (control) (see **Table 5.1** for more details). Under usual care, patients can be enrolled in one or more of existing post-discharge services that operate independently, such as care-coordination, home care, telemonitoring, and home visits by pharmacists. CareHub patients deal with a single point of contact at the hospital whereas usual care patients might encounter several. CareHub oversees all post-discharge services offered to the patient, removes clinician discretion to enrol patients, coordinates the workflow across healthcare providers through systematic and regular multidisciplinary meetings during the inpatient and post-discharge periods, actively involves pharmacists and ward nurses in inpatient medication education and counselling based on a comprehensive assessment of needs, monitors and resolves emerging medical issues during telephone follow-ups by imposing structured questions on medication compliance and symptom checks, and provides a call centre during office hours with access to doctors' advice.

5.2.2 Study Population and Randomisation

The enrolment of patients into the study ran from July 2016 to November 2016, during a planned operational test-run of CareHub. Patients were enrolled based on criteria that NUH planned to use when CareHub is fully implemented. The pilot included patients who (i) were admitted into the Cardiology wards through the Emergency Department (ED), (ii) qualified for government subsidies of more than 50% (ward-classes B and C), and (iii) were assessed to have a ACE risk-score (derived from the LACE index) of nine or more, based on the number of existing co-morbidities and ED visits in the preceding six months (Wang et al., 2014; Low et al., 2015).

We excluded the ‘L’ (length-of-stay) in the LACE because the CareHub protocol needs to start at admission, and L is a number that can only be determined at the end of stay.

Non-residents and patients returning to institutionalised care after discharge were excluded. Non-residents such as tourists comprised a tiny share of the NUH patient population, and being transient, did not receive follow up. Patients returned to institutionalised care were not followed up by NUH. Since CareHub was provided for free as part of NUH’s care plan, patients did not explicitly opt-out of the service, though they could reject any service component at any point, say, by refusing to talk to the coordinator over the phone.

As resources designated for the pilot programme were not sufficient to meet the needs of the entire target population, assignment into CareHub followed a pragmatic randomisation protocol. To balance the day-to-day operational load, the computer generated an alphabetically-ordered list of eligible patients daily, and alternate patients were assigned to CareHub. (The first person on each day’s list was assigned to CareHub.) In all, 270 patients were included in the study. 150 were assigned to CareHub, while 120 remained in usual care, comprising the control group. Randomisation allowed a rigorous evaluation of CareHub’s effectiveness, avoiding the self-selection problem that often plagues non-randomised programme evaluations.

5.2.3 Data sources

Our data comes from three sources: National University Hospital’s (NUH) administrative data, National Healthcare Group’s (NHG) administrative data, and survey data. NUH provided data on patient demographics, healthcare utilisation six months before and after patients’ index admission, and patients’ clinical indicators. We use these data to analyse patients’ outcomes and compute back-of-the envelope calculations of CareHub’s cost-effectiveness. NHG’s data includes information on patients’ utilisation of nearby public hospitals in 2016. However, these data do not cover the full follow-up period through 2017, and thus could only be used for robustness checks.

Finally, to complement the primary outcome measures in the administrative data, we conducted a patient survey, six months after the last patient was enrolled. This survey collected data on quality of life and quality of transitional care (see supplementary materials). The survey response rate was 28% (76 patients with 33 from CareHub), which is high for this type of study.

5.2.4 Outcome Measures and Covariates

Our primary outcomes are observed in the six-month follow-up period, and focused on the utilisation of services in NUH. This allowed us to estimate the cost effectiveness of the programme. They include the number of all/unplanned cardiac-related readmissions, total number of days spent in all/unplanned cardiac-related readmissions, number of cardiac-related specialist outpatient clinic (SOC) visits, and number of emergency department (ED) visits. These variables take value zero if patients were not readmitted or did not make any visits after index admission (i.e. “total number of days spent in readmissions” includes patients who were not readmitted). Our secondary outcomes include survey data on quality of life and quality of transitional care from the hospital, and administrative data on relevant clinical indicators (diastolic/systolic blood pressure). The supplementary materials provide more details.

5.2.5 Statistical Analyses

5.2.5.1 Main analyses

The randomisation process allowed us to estimate an unbiased intent-to-treat effect of CareHub versus usual care. We recover only the intent-to-treat effect, rather than the average treatment effect, as patients can reject any component of CareHub at any point in time. We use ordinary least squares (OLS) regression with heteroskedasticity-robust standard errors to estimate the effect of CareHub on the primary outcomes. The key treatment covariate is a binary variable “CareHub” that takes value one if patient was assigned to CareHub. To control for observed differences between patients in both arms, we add covariates for:

(i) demographics: gender, age, square of age, ethnicity, marital status, month of index admission, and ward-class (which contains information on socio-economic status, as government subsidies vary by ward-class),

(ii) existing co-morbidities, and

(iii) baseline healthcare utilisation at NUH (which contains information about patients' physical health and social needs, which are captured by frequent readmissions due to lack of caregivers or self-care at home). Covariates on baseline healthcare utilisation include number of inpatient visits, total number of days spent in inpatient stays, number of SOC visits, and number of ED visits in the six months preceding their index admission.

We perform sensitivity analyses that include different subsets of the abovementioned covariates using OLS, and Cox proportional hazards regression. In general, our results are robust to, or consistent with the sensitivity checks. More details are in the supplementary materials.

We use ordered logit regressions and t-tests to assess the effect of CareHub in the survey outcomes. We do not include the additional controls mentioned above, due to the small sample size for our survey dataset (The supplementary materials contain a detailed discussion.) We apply the methodology used for the primary outcomes on blood pressure outcomes.

5.2.5.2 Subgroup analyses

NUH-related administrators and clinicians hypothesised that CareHub would be most effective for patients with more health problems and/or social needs, given the interaction between social- and health-related factors in the patient case mix. Hence, we conduct additional exploratory analyses of CareHub's impact on healthcare utilisation across different segments of patients. The relationships are correlational as we did not stratify our sample along these dimensions before randomisation. However, the results are highly suggestive of potential directions for future research.

To study whether effects vary by baseline physical health and social needs, we add the following interaction term into the main regression: "CareHub" × total number of inpatient days for each patient in the six months before index admission. Similarly, we examine differences in socio-economic status, gender, and age by interacting "CareHub" with "Ward C", "Female", and "Age". "Ward C" and "Female" are binary variables that take value one if the patient was in a C-class ward (the most heavily subsidised) or female respectively. Each heterogeneity analysis is conducted separately (i.e. only one interaction term was added to each regression). The values of interest are the coefficients of the interaction terms.

5.2.6 Study Approval

Our study protocol was approved by Singapore's National Healthcare Group's Domain-Specific Review Board.

5.3 Results

5.3.1 Summary Statistics and Balance Checks

Balance tests show that the CareHub and control patients are similar in demographics, co-morbidities, and pre-index admission hospital utilisation (**Table 5.2**). Differences between a few characteristics are statistically significant, but the absolute differences, standardised by variances, are small (less than 0.3). Nonetheless, given the statistically significant differences in some baseline characteristics between CareHub and control patients, we include the baseline characteristics as covariates in our main regression analyses.

5.3.2 Primary Outcomes

Table 5.3 reports the effect of CareHub on the primary outcomes, which measured *cardiac-related* utilisation (except the number of Emergency Department visits, which could not be classified). These estimates come from a regression that includes covariates for demographics, co-morbidities, and pre-index-admission hospital utilisation. In general, CareHub appears effective in reducing readmission outcomes in the six months after enrolment. This effect is strongest for unplanned readmissions.

CareHub reduced the number of unplanned readmissions by 0.23 readmissions ($p<0.05$). In other words, compared to patients under usual care, CareHub patients experienced 39% fewer readmissions on average (Column 1, **Table 5.3**: -0.234 ÷ 0.60 = -39%). CareHub's effect on all readmissions was very similar, which is a reduction of 0.20 readmissions ($p=0.10$; 31% lower than usual care). The total number of unplanned days spent in hospital were reduced by 2.2 days ($p<0.05$; 56% lower than usual care), while the total number of days spent in hospital fell by 2.0 days ($p<0.10$; 47% lower than usual care).

The fall in number of days spent in hospital can be attributed to three sources: fewer patients being readmitted, fewer readmissions amongst those readmitted at least once, and fall in length of stay per readmission. A rough decomposition shows that these three sources contributed about 29%, 46%, and 25% to the total effect, respectively (see supplementary materials).

There is little evidence of any change in outpatient specialist visits ($p=0.39$) or emergency department visits ($p=0.75$). All results in this section are robust to sensitivity analyses with different sets of covariates, including a specification with no covariates (see supplementary materials).

5.3.3 Heterogeneous Treatment Effects

NUH hypothesised that CareHub would be most effective for patients with more health problems and/or social needs, as they would probably benefit more from better planned care transitions. Consistent with that notion, our exploratory analyses show that CareHub's effects on cardiac-related readmission outcomes appear larger for patients who spent more days in the hospital prior to the index admission and patients who stayed in C-class wards. CareHub's effects did not vary by gender or age (see **Table 5.4)**.

Panels A and B report heterogeneous effects by the total number of days each patient spent in hospital prior to index admission and ward class, respectively. The coefficients of interest are negative across all the outcome variables, and statistically significant for the number of all/unplanned readmissions. These results suggest that CareHub was effective in reducing readmission utilisation for those with greater pre-existing health and social needs and those of lower socio-economic status. In other words, CareHub seems to have worked for the population it intended to focus its efforts on.

Panels C and D respectively suggest that the effect of CareHub may not vary by gender and age: the coefficients of interest are statistically insignificant, and inconsistent in sign across the outcomes studied.

5.3.4 Secondary Outcomes: Survey Results and Clinical Indicators

The survey results in **Table 5.5** suggest that CareHub improved quality of life and transitional care (see supplementary materials for more details). While our estimates are not sufficiently precise to achieve statistical significance for many survey outcome variables, the direction of the estimates consistently suggest that CareHub improved quality of life and transitional care. The strongest effects were in areas that CareHub was designed to address. CareHub patients reported lower anxiety/depression (odds ratio = 3.2, $p < 0.05$), were more likely to know who to contact at NUH (odds ratio = 2.6, $p < 0.05$); were more satisfied with the hospital's follow-up care (odds ratio = 3.6, $p < 0.01$), and were less likely to run out of medication (odds ratio = 4.33, $p < 0.10$). Finally, we find that CareHub made no difference to patients' diastolic and systolic blood pressure. The coefficients of interest are statistically and clinically insignificant (see supplementary materials).

5.4 Discussion

This study demonstrated that a one-stop patient-centric transitional care service can be effective for reducing hospital utilisation. In the context of CareHub, effects are largest among patients with more complex health and social needs (and hence at higher risk of readmission). This finding has been documented in the literature (Counsell et al., 2007; Brown et al., 2012; Peikes et al., 2012; Xing et al., 2015). The net cost reduction from fewer days of inpatient stay in the CareHub cohort in the 6 months after index admission, after accounting for startup and operational costs, was about SGD200,000 or SGD$1,300 per CareHub patient. These figures can be seen as an underestimate of the true savings, as they include some patients who were not part of the study, and as reduced readmissions among CareHub patients may continue beyond the six-month period we consider in this paper. Beyond healthcare utilisation, we find suggestive evidence that CareHub improved patients' quality of life in terms of reduced anxiety/depression.

5.4.1 Limitations

As with studies of this type, our study is subject to limitations. The healthcare utilisation data came only from NUH administrative data, which might bias our estimates if there was significant non-NUH utilisation. Using 2016 data for utilisation in the National Healthcare Group, which includes other public hospitals geographically closest to NUH, we find that the proportion of non-NUH utilisation for CareHub and control patients were similar before index admission, while the proportion of non-NUH utilisation fell for CareHub patients after admission to the programme (supplementary materials). This implies that the real effect of CareHub is likely higher than our effect estimated from NUH data.

Our survey results and heterogeneity analyses are only suggestive, due to the small sample (for surveys) and lack of *ex-ante* stratification (for heterogeneity). However, the consistency of the direction of the effects, coupled with the fact that the strongest improvements are seen in areas CareHub was designed to address, strongly suggests that CareHub's effects in these analyses are real (see supplementary materials and Table 5.4).

Finally, as our study included only two arms, we are unable to disentangle the relative effects of the different CareHub components, for example the impact of phone calls versus home visits. This could be the subject of future studies.

5.4.2 Potential Reasons for Improved Outcomes

Our survey results, interviews with NUH healthcare staff, and a comparison of CareHub's features with those of other successful transitional care / care-coordination programmes (Brown et al., 2012; Allen et al., 2014; Kronick et al., 2014) suggest possible mechanisms that explain CareHub's effectiveness.

First, CareHub enrolled patients using a risk scoring tool, rather than clinician referrals. The CareHub recruitment process was systematic compared to how patients were normally enrolled in these programmes, which increased the odds that CareHub would recruit patients that needed the programme most.

Second, CareHub adopted protocols to improve communication among providers involved in the patient's post-discharge care. CareHub's protocol mandated that the multidisciplinary team meet regularly (daily when the patient was in the hospital and weekly after discharge) to discuss each case. In contrast, before CareHub, providers only shared information whenever they felt it was necessary. The communication protocol reduced the odds that crucial information was missed because they were deemed unimportant. Such protocols are common in in-hospital settings such as surgical time-outs and hand-offs between shifts. They are less common in post-discharge care programmes.

Third, the post-discharge monitoring of patients was protocolised, relative to usual care. CareHub's telephone follow-ups involved structured checks on patients' symptoms and medication compliance. A telephone line in the call center was dedicated to CareHub patients so they had one number to call with questions or to report medical problems. Responses to patients' queries were also protocolised, so that patients' queries were directed to the correct personnel (e.g. doctors) or provider for quick action. Interviews with the CareHub staff and our survey results suggest that this practice improved medication compliance. Although there are standard protocols such as pill reconciliation, to monitor medication compliance, it is still not part of routine medical care (Zullig et al., 2016).

5.4.3 Implications

The extant literature has focused on the introduction of TCPs into settings without existing transitional care. Whether further gains can be made by attempts to improve existing programmes (e.g. the Care Transitions Intervention (Coleman et al., 2004)) remains understudied. Our results provide some answers to this question. First, hospitals with existing TCPs should continue to look for ways to increase efficiency and effectiveness by integrating and coordinating such programmes, especially if they live in separate departments in the hospital. Substantial improvements in patient outcomes may still be had by refining existing transitional care practices in the hospital.

A second insight from our study echoes Cook et. al.'s comment that bridging current gaps may lead to the formation of new gaps in care (Cook et al., 2000). In our setting, the introduction of TCPs for different, but sometimes overlapping, groups of patients may have opened new gaps, as they increased the complexity of the transitional care landscape in the hospital. These gaps made coordination between programmes difficult, so that patients often found themselves falling between the boundaries where such programmes did not serve. The effectiveness of CareHub may partially be due to its ability to bridge these gaps (e.g., enrolment into these programmes were all routed through the CareHub coordinator). With the increased interest in TCPs, different departments within a hospital may start to implement their own TCPs. Ensuring that they operate in a coordinated fashion is likely to improve patient outcomes and minimise waste.

5.5 Conclusion

We find that CareHub was effective in reducing healthcare utilisation. Exploratory analyses suggest that CareHub was most effective among patients with more complex health and social needs. Our survey results also suggest that CareHub reduced anxiety/depression, improved transitional care quality, and reduced the likelihood of patients running out of medicine. We note that these improvements took place in a setting where TCPs already exist. This suggests that additional value can be extracted in healthcare systems with existing TCPs through the continual identification and closure of service and process gaps. TCP proliferation may improve outcomes, but may also increase fragmentation of care, which can be bridged by improving care-coordination.

References

Allen, J., Hutchinson, A. M., Brown, R., & Livingston, P. M. (2014). Quality Care Outcomes Following Transitional Care Interventions for Older People from Hospital to Home: A Systematic Review. *BMC Health Services Research, 14*, 346.

Anderson, C., Deepak, B., Amoateng- Adjepong, Y., & Zarich, S. (2005). Benefits of Comprehensive Inpatient Education and Discharge Planning Combined with Outpatient Support in Elderly Patients with Congestive Heart Failure. *Congestive Heart Failure, 11*(6), 315-321.

Brown, R. S., Peikes, D., Peterson, G., Schore, J., & Razafindrakoto, C. M. (2012). Six Features of Medicare Coordinated Care Demonstration Programs That Cut Hospital Admissions of High-Risk Patients. *Health Affairs, 31*(6), 1156-1166.

Centers for Medicare and Medicaid Services. (2016). Community-Based Care Transitions Program. Retrieved 22 August 2017 from https://innovation.cms.gov/initiatives/CCTP/

Chen, Y., Tan, Y. J., Sun, J., Chua, C. Z., Yoo, J., Wong, S. H., . . . Phan, P. (2017). Online Supplementary Materials for the Paper "Can Refinements to Effective Transitional Care Services Improve Outcomes? Results from a Pragmatic, Randomized Controlled Trial".

Coleman, E. A. (2003). Falling through the Cracks: Challenges and Opportunities for Improving Transitional Care for Persons with Continuous Complex Care Needs. *Journal of the American Geriatrics Society, 51*(4), 549-555.

Coleman, E. A., Smith, J. D., Frank, J. C., Min, S. J., Parry, C., & Kramer, A. M. (2004). Preparing Patients and Caregivers to Participate in Care Delivered across Settings: The Care Transitions Intervention. *Journal of the American Geriatrics Society, 52*(11), 1817-1825.

Cook, R. I., Render, M., & Woods, D. D. (2000). Gaps in the Continuity of Care and Progress on Patient Safety. *British Medical Journal, 320*, 791 - 794.

Counsell, S. R., Callahan, C. M., Clark, D. O., Tu, W., Buttar, A. B., Stump, T. E., & Ricketts, G. D. (2007). Geriatric Care Management for Low-Income Seniors: A Randomized Controlled Trial. *JAMA, 298*(22), 2623-2633.

Ford, I., & Norrie, J. (2016). Pragmatic Trials. *New England Journal of Medicine, 375*(5), 454-463.

Jayakody, A., Bryant, J., Carey, M., Hobden, B., Dodd, N., & Sanson-Fisher, R. (2016). Effectiveness of Interventions Utilising Telephone Follow up in Reducing Hospital Readmission within 30 Days for Individuals with Chronic Disease: A Systematic Review. *BMC Health Services Research, 16*(1), 403.

Jencks, S. F., Williams, M. V., & Coleman, E. A. (2009). Rehospitalizations among Patients in the Medicare Fee-for-Service Program. *New England Journal of Medicine, 360*(14), 1418-1428.

Kronick, R., Slutsky, J., Chang, S., & Chang, C. (2014). *Transitional Care Interventions to Prevent Readmissions for People with Heart Failure*. Retrieved from Rockville, MD:

Low, L. L., Lee, K. H., Hock Ong, M. E., Wang, S., Tan, S. Y., Thumboo, J., & Liu, N. (2015). Predicting 30-Day Readmissions: Performance of the Lace Index Compared with a Regression Model among General Medicine Patients in Singapore. *BioMed Research International, 2015*.

Mistiaen, P., Francke, A. L., & Poot, E. (2007). Interventions Aimed at Reducing Problems in Adult Patients Discharged from Hospital to Home: A Systematic Meta-Review. *BMC Health Services Research, 7*(1), 47.

Naylor, M., Aiken, L. H., Kurtzman, E. T., Olds, D. M., & Hirschman, K. B. (2011). The Importance of Transitional Care in Achieving Health Reform. *Health Affairs, 30*(4), 746-754.

Naylor, M., Brooten, D., Jones, R., Lavizzo-Mourey, R., Mezey, M., & Pauly, M. (1994). Comprehensive Discharge Planning for the Hospitalized Elderlya Randomized Clinical Trial. *Annals of internal Medicine, 120*(12), 999-1006.

Naylor, M., Brooten, D. A., Campbell, R. L., Maislin, G., McCauley, K. M., & Schwartz, J. S. (2004). Transitional Care of Older Adults Hospitalized with Heart Failure: A Randomized, Controlled Trial. *Journal of the American Geriatric Society, 52*(5), 675-684.

Oddone, E. Z., Weinberger, M., Horner, M., Mengel, C., Goldstein, F., Ginier, P., . . . Asch, D. A. (1996). Classifying General Medicine Readmissions. *Journal of General Internal Medicine, 11*(10), 597-607.

Peikes, D., Peterson, G., Brown, R. S., Graff, S., & Lynch, J. P. (2012). How Changes in Washington University's Medicare Coordinated Care Demonstration Pilot Ultimately Achieved Savings. *Health Affairs, 31*(6), 1216-1226.

Rich, M. W., Beckham, V., Wittenberg, C., Leven, C. L., Freedland, K. E., & Carney, R. M. (1995). A Multidisciplinary Intervention to Prevent the Readmission of Elderly Patients with Congestive Heart Failure. *New England Journal of Medicine, 333*(18), 1190-1195.

Wang, H., Robinson, R. D., Johnson, C., Zenarosa, N. R., Jayswal, R. D., Keithley, J., & Delaney, K. A. (2014). Using the Lace Index to Predict Hospital Readmissions in Congestive Heart Failure Patients. *BMC Cardiovascular Disorders, 14*(1), 97.

Wee, S. L., Loke, C. K., Liang, C., Ganesan, G., Wong, L. M., & Cheah, J. (2014). Effectiveness of a National Transitional Care Program in Reducing Acute Care Use. *Journal of the American Geriatrics Society, 62*(4), 747-753.

Wong, F. K. Y., Chow, S. K. Y., Chan, T. M. F., & Tam, S. K. F. (2013). Comparison of Effects between Home Visits with Telephone Calls and Telephone Calls Only for Transitional Discharge Support: A Randomised Controlled Trial. *Age and Ageing, 43*(1), 91-97.

Xing, J., Goehring, C., & Mancuso, D. (2015). Care Coordination Program for Washington State Medicaid Enrollees Reduced Inpatient Hospital Costs. *Health Affairs, 34*(4), 653-661.

Zullig, L. L., Granger, B. B., & Bosworth, H. B. (2016). A Renewed Medication Adherence Alliance Call to Action: Harnessing Momentum to Address Medication Nonadherence in the United States. *Patient preference and adherence, 10*, 1189.

Table 5.1: Comparison of study intervention and usual care

Feature	Usual Care (Aged Care Transition Programme - ACTION[1])	Study Intervention (CareHub)
Selection into programme	Patients who meet pre-set criteria are subsequently selected at the discretion of the attending doctor and/or coordinator.	Patients are selected as long as pre-set criteria are met. No discretion is involved in the selection process.
Qualification of care coordinators	Nurses-by-training, physiotherapists, non-professional personnel.	Practicing nurse (team leader), nurse-by-training.
Customisation of discharge care plan	During hospitalisation, coordinator works with family members, other hospital staff on an ad-hoc one-on-one needs basis, and community partners.	During hospitalisation, coordinator (i)reaches out to family members; (ii)conducts a daily multidisciplinary huddle with all relevant professionals including the ward physician, nurse, pharmacist, allied healthcare workers such as dieticians and physiotherapists, and the medical social worker. Such high-frequency and simultaneous discussions among all relevant professionals allow for a more comprehensive care plan that considers patients' social set-up, financial situation, functional status, goals of care in addition to medical issues; and (iii)works with community partners.
Medication education and reconciliation	Conducted based on existing inpatient guidelines.	During hospitalisation, the coordinator will activate the ward pharmacist and/or nurse to conduct medication education and reconciliation based on a comprehensive assessment of the patient's needs, e.g. whether patient has existing medication, whether patient is literate etc.
Patient and caregiver education	Available during hospitalisation	Available during hospitalisation. Notes of the education sessions will also be given to the patient and caregiver as part of the discharge summary and care plan.

Feature	Usual Care (Aged Care Transition Programme - ACTION[1])	Study Intervention (CareHub)
Frequency of telephone follow-ups	First call is made within 48 hours of discharge. Other follow-up calls are made on an ad-hoc basis.	At least four times within six months, with the first call made within 48 hours of discharge. For the most complex cases, eight calls can be scheduled within the six-month period after discharge.
Content of telephone follow-ups	(i)Structured questions on functional issues e.g. whether patient can take care of himself, whether he needs community help, whether the home set-up is safe. (ii)Referrals to community partners if necessary.	(i)Follow up on outstanding medical issues, e.g. whether existing medical condition has been resolved, whether the existing medication dosage is effective and allowing adjustment of dosage (after consultation with doctor). (ii)Structured questions that monitor physical symptoms based on existing medical conditions. (iii) Structured questions on medication compliance and follow-up actions such as arranging a convenient collection of medication by patient/caregiver. (iv) Structural questions on functional issues. (v) Referrals to community partners if necessary. (vi) Reminder for upcoming appointment for specialist outpatient visits, on top of existing hospital-wide reminder system. For those who are illiterate, assistance is provided to reschedule appointments if required. For those who have more than one appointment due to multiple co-morbidities, the coordinator helps to consolidate appointments so that patient can make fewer trips to the hospital.
Home visits	Available based on needs	Available based on needs
Referral to community providers	Available	Available

Feature	Usual Care (Aged Care Transition Programme - ACTION[1])	Study Intervention (CareHub)
Post-discharge multidisciplinary discussion	Not available. Relevant personnel are contacted on an ad-hoc one-on-one basis. Rarely has access to a physician.	Any issues that patients experience post-discharge are discussed at a weekly multidisciplinary meeting which includes a physician.
Call centre	Not available. Patient can call coordinator-in-charge. Questions cannot be escalated to doctors.	Available during office hours. Questions from patients can be escalated to doctors if necessary.
Coordination across different post-discharge services	Coordinator does not have oversight over all other post-discharge services provided by the hospital, e.g. telehealth. Coordinator neither coordinates nor communicates with other service providers. As such, a patient who has enquiries regarding other services has to contact the respective service providers.	Coordinator has oversight of and coordinates all post-discharge services provided to patient. As such, patient needs to contact only the coordinator even if patient has enquiries about the other services.

[1]As mentioned, NUH offers many types of post-discharge services under usual care. We choose to contrast one of them (ACTION) against CareHub, as the design of ACTION is most similar to CareHub. Like CareHub, ACTION is also provided free-of-charge.

Table 5.2: Summary Statistics and Balance Check

Variables	Carehub (N=150)		Usual Care (N=120)		Difference	
	Mean	SD	Mean	SD	Diff in means[1]	Norm diff[2]
Demographics						
Age	70.15	12.82	68.49	14.03	1.66	0.12
Chinese (proportion)	0.67	0.47	0.54	0.5	0.13**	0.26
Malay (proportion)	0.18	0.39	0.21	0.41	-0.03	-0.07
Indian (proportion)	0.1	0.3	0.13	0.33	-0.03	-0.08
Other ethnicities (proportion)	0.05	0.23	0.13	0.33	-0.07**	-0.25
Male	0.57	0.5	0.57	0.5	0.01	0.01
Ward B	0.28	0.45	0.41	0.49	-0.13**	-0.27
Married[3]	0.91	0.29	0.86	0.35	0.05	0.14
Co-morbidities (proportion with disease)[4]						
Previous myocardial infarction	0.57	0.5	0.47	0.5	0.1	0.2
Cerebrovascular / peripheral vascular disease	0.19	0.4	0.3	0.46	-0.10**	-0.24
Diabetes	0.45	0.5	0.45	0.5	0	0
Congestive heart failure	0.43	0.5	0.42	0.5	0.01	0.02
Chronic pulmonary disease	0.18	0.39	0.25	0.44	-0.07	-0.18
Liver or renal failure	0.31	0.47	0.25	0.43	0.07	0.15
Connective tissue disease	0.85	0.36	0.9	0.3	-0.05	-0.15
Other diseases[5]	0.17	0.38	0.14	0.34	0.04	0.1
Hospital utilisation in the 6 months before enrollment						
No. of inpatient visits	0.78	1.89	0.87	1.33	-0.09	-0.05
No. of cardiac-related inpatient visits	0.33	0.62	0.38	0.79	-0.06	-0.08
Total days in hospital	5.24	11.88	4.87	9.3	0.37	0.03
Total cardiac-related days in hospital	2.45	6.47	2.98	7.19	-0.53	-0.08
No. of specialist visits	3.71	4.61	4.05	5.01	-0.34	-0.07
No. of cardiac-related specialist visits	1.13	1.91	1.4	1.87	-0.27	-0.14
No. of emergency department visits	0.96	2.24	1.02	1.43	-0.06	-0.03

Source: Authors' calculations using NUH/NHG administrative data and survey data. Notes:
[1] * $p < 0.1$ ** $p<0.05$ *** $p < 0.01$ based on t-tests.
[2] Normalised differences computed as the difference in means, standardised by the square root of the mean variance from both groups.
[3] We only have data on marital status for 97 in the treated group, and 87 in the control group.
[4] Data on co-morbidities is missing for 2 individuals in the control group.
[5] Other diseases include AIDS, cancer, and dementia.

Table 5.3: Effect of CareHub on Key Outcomes

Variable	(1) No. of readmissions (cardiac & unplanned)	(2) Total days readmitted (cardiac & unplanned)	(3) No. of readmissions (cardiac)	(4) Total days readmitted (cardiac)	(5) Outpatient specialist visits (cardiac)	(6) Emergency department visits
CareHub	-0.234**	-2.241**	-0.196	-2.017*	0.328	-0.0702
	(0.114)	(1.120)	(0.119)	(1.184)	(0.384)	(0.218)
Observations	270	270	270	270	270	270
R-squared	0.229	0.198	0.210	0.171	0.156	0.392
Post-index control group mean	0.60	4.01	0.63	4.32	2.50	1.36

Source: Authors' calculations using NUH/NHG administrative data and survey data. Notes:

[1] Heteroskedasticity-robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

[2] Estimates of the effect of CareHub are from a regression model that includes controls for demographics, co-morbidities, and healthcare utilisation in the 6 months before the index visit.

[3] For the outcome variables "total days readmitted" in columns (2) and (4), patients without any readmissions are assigned a value of 0. These outcome variables thus capture effects due to reduced readmissions as well as reduced length of stay for patients who are readmitted.

Table 5.4: Heterogeneity Analyses

Variable	(1) No. of readmissions (cardiac & unplanned)	(2) Total days readmitted (cardiac & unplanned)	(3) No. of readmissions (cardiac)	(4) Total days readmitted (cardiac)
Panel A: Heterogeneity by total days in hospital before index visit				
CareHub	0.0383	0.802	0.0729	0.992
	(0.116)	(1.426)	(0.123)	(1.515)
CareHub × total days	-0.0560***	-0.625	-0.0553***	-0.618
	(0.0190)	(0.393)	(0.0195)	(0.399)
Panel B: Heterogeneity by ward class				
CareHub	0.116	-1.024	0.208	-0.813
	(0.162)	(1.611)	(0.182)	(1.929)
CareHub × ward C	-0.545**	-1.897	-0.631***	-1.878
	(0.227)	(1.674)	(0.241)	(2.017)
Panel C: Heterogeneity by gender				
CareHub	-0.231	-1.056	-0.226	-1.020
	(0.154)	(1.021)	(0.159)	(1.233)
CareHub × female	-0.00748	-2.768	0.0698	-2.329
	(0.246)	(3.186)	(0.253)	(3.332)
Panel D: Heterogeneity by age				
CareHub	-0.796	0.151	-0.703	0.431
	(0.586)	(4.532)	(0.650)	(5.030)
CareHub × age	0.00812	-0.0346	0.00731	-0.0354
	(0.00813)	(0.0705)	(0.00892)	(0.0757)

Source: Authors' calculations using NUH/NHG administrative data and survey data. Notes:
[1] Heteroskedasticity-robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
[2] Coefficient estimates for CareHub and the interaction terms are from a regression model that includes controls for demographics, co-morbidities, and healthcare utilisation in the 6 months before the index visit.

Table 5.5: Survey Results

Variable	CareHub			Usual Care			Simple linear difference		Ordered logit results	
	N	Mean	SD	N	Mean	SD	Diff in means	Norm diff	Odds ratio	Std error
<u>Quality of Life</u>										
How well respondent can carry out usual activities	43	4.56	0.96	33	4.15	1.20	0.41	0.37	2.43*	1.21
Whether respondent is anxious or depressed	42	4.55	0.77	33	4.00	1.12	0.55**	0.57	3.20**	1.50
<u>Quality of Transitional Care</u>										
I know who to call at NUH	39	3.23	0.84	33	2.82	0.88	0.41**	0.48	2.62**	1.20
I am satisfied with the hospital's follow-up care	39	3.62	0.49	32	3.16	0.81	0.46***	0.69	3.56***	1.73
Do you sometimes run out of medicines because you did not collect them[5]	41	0.95	0.22	33	0.82	0.39	0.13*	0.42	4.33*	3.73

Source: Authors' calculations using NUH/NHG administrative data and survey data. Notes:
[1] *** p<0.01, ** p<0.05, * p<0.1.
[2] We flipped the values for all outcome variables so that a larger number represents a better outcome.
[3] We applied t-test for column "Diff in means". Normalised differences as captured in column "Norm diff" are computed as the difference in means, standardised by the square root of the mean variance from both groups.
[4] Odds ratios are from bivariate ordered logit regressions, where the only independent variable is a dummy variable assignment to CareHub. Robust standard errors for the odds ratios were obtained using the delta method.
[5] Unlike other questions in the transitional care section, the response to the last question is a binary variable that takes value 1 if patient replies no.
[6] Full survey results available in supplementary materials.

Supplementary Materials

Robustness Checks for Main Results

This section explores the sensitivity of our estimates to different regression specifications. Each table in this section reports coefficient estimates for the CareHub dummy variable under different specifications. Each table focuses on one outcome variable. We study the sensitivity of the coefficient estimates to the following four specifications:

1. A bivariate regression where the only independent variable is CareHub;
2. A regression where controls for the following demographic-type controls are added: gender, age, age^2, ethnicity, marital status, ward class, month of admission;
3. A regression where controls from (2), as well as the following controls for co-morbidities are added: previous myocardial infarction, vascular disease, diabetes, congestive heart failure, chronic pulmonary disease, liver/renal failure, connective tissue disease, and other diseases (AIDS, cancer, dementia);
4. This is the specification used for the main paper. In this regression, we include controls from (3), as well as controls for hospital utilisation in the 6 months before index admission: number of inpatient visits, total days spent in hospital, number of specialist visits, number of emergency department visits in the 6 months before index admission.

Our results indicate that the results in **Table 5.3** of the main tables are robust. The signs of the coefficients remain consistent across all specifications, and the coefficients that were significant in the main specification generally remain statistically significant in the robustness checks (with the exception of the total number of days readmitted for cardiac-related reasons, though the magnitude of the coefficient remains similar).

Table S1: Robustness Checks for No. of Unplanned, Cardiac-Related Readmissions

Variable	(1) No controls	(2) Demo controls	(3) Demo, comorbidity controls	(4) Demo, comorbidity, utilisation controls **(Main spec)**
CareHub	-0.273**	-0.258**	-0.216*	-0.234**
	(0.129)	(0.126)	(0.114)	(0.114)
Observations	270	270	270	270
R-squared	0.019	0.079	0.160	0.229

Notes:
[1] Parentheses report heteroskedasticity-robust standard errors for columns (1) – (4). *** p<0.01, ** p<0.05, * p<0.1
[2] Columns (1) – (4) report coefficient estimates for the CareHub variable (=1 if patient is in the CareHub group). Column (1) reports results from a regression without controls; (2) includes controls for demographics; (3) includes controls for demographics and comorbidities; (4) includes controls for demographics, comorbidities, and healthcare utilisation in the 6 months before index admission.

Table S2: Robustness Checks for Total Number of Days Readmitted for Unplanned, Cardiac-Related Reasons

Variable	(1) No controls	(2) Demo controls	(3) Demo, comorbidity controls	(4) Demo, comorbidity, utilisation controls **(Main spec)**
CareHub	-2.128*	-1.867*	-1.895*	-2.241**
	(1.206)	(1.021)	(0.980)	(1.120)
Observations	270	270	270	270
R-squared	0.013	0.052	0.125	0.198

Notes:
[1] Parentheses report heteroskedasticity-robust standard errors for columns (1) – (4). *** p<0.01, ** p<0.05, * p<0.1
[2] Columns (1) – (4) report coefficient estimates for the CareHub variable (=1 if patient is in the CareHub group). Column (1) reports results from a regression without controls; (2) includes controls for demographics; (3) includes controls for demographics and comorbidities; (4) includes controls for demographics, comorbidities, and healthcare utilisation in the 6 months before index admission.

Table S3: Robustness Checks for No. of Cardiac-Related Readmissions

Variable	(1) No controls	(2) Demo controls	(3) Demo, comorbidity controls	(4) Demo, comorbidity, utilisation controls **(Main spec)**
CareHub	-0.247*	-0.229*	-0.184	-0.196
	(0.131)	(0.128)	(0.117)	(0.119)
Observations	270	270	270	270
R-squared	0.014	0.078	0.153	0.210

Notes:
[1] Parentheses report heteroskedasticity-robust standard errors for columns (1) – (4). *** p<0.01, ** p<0.05, * p<0.1
[2] Columns (1) – (4) report coefficient estimates for the CareHub variable (=1 if patient is in the CareHub group). Column (1) reports results from a regression without controls; (2) includes controls for demographics; (3) includes controls for demographics and comorbidities; (4) includes controls for demographics, comorbidities, and healthcare utilisation in the 6 months before index admission.

Table S4: Robustness Checks for Total Number of Days Readmitted for Cardiac-Related Reasons

Variable	(1) No controls	(2) Demo controls	(3) Demo, comorbidity controls	(4) Demo, comorbidity, utilisation controls **(Main spec)**
CareHub	-1.877	-1.603	-1.685	-2.017*
	(1.251)	(1.068)	(1.041)	(1.184)
Observations	270	270	270	270
R-squared	0.009	0.049	0.109	0.171

Notes:
[1] Parentheses report heteroskedasticity-robust standard errors for columns (1) – (4). *** p<0.01, ** p<0.05, * p<0.1
[2] Columns (1) – (4) report coefficient estimates for the CareHub variable (=1 if patient is in the CareHub group). Column (1) reports results from a regression without controls; (2) includes controls for demographics; (3) includes controls for demographics and comorbidities; (4) includes controls for demographics, comorbidities, and healthcare utilisation in the 6 months before index admission.

Table S5: Robustness Checks for Total Number of Cardiac-Related Outpatient Specialist Visits

Variable	(1) No controls	(2) Demo controls	(3) Demo, comorbidity controls	(4) Demo, comorbidity, utilisation controls **(Main spec)**
CareHub	0.260 (0.368)	0.304 (0.378)	0.297 (0.394)	0.328 (0.384)
Observations	270	270	270	270
R-squared	0.002	0.074	0.108	0.156

Notes:
[1] Parentheses report heteroskedasticity-robust standard errors for columns (1) – (4). *** p<0.01, ** p<0.05, * p<0.1
[2] Columns (1) – (4) report coefficient estimates for the CareHub variable (=1 if patient is in the CareHub group). Column (1) reports results from a regression without controls; (2) includes controls for demographics; (3) includes controls for demographics and comorbidities; (4) includes controls for demographics, comorbidities, and healthcare utilisation in the 6 months before index admission.

Table S6: Robustness Checks for Total Number of Emergency Department Visits

Variable	(1) No controls	(2) Demo controls	(3) Demo, comorbidity controls	(4) Demo, comorbidity, utilisation controls **(Main spec)**
CareHub	-0.212 (0.300)	-0.114 (0.303)	-0.0238 (0.257)	-0.0702 (0.218)
Observations	270	270	270	270
R-squared	0.002	0.101	0.172	0.392

Notes:
[1] Parentheses report heteroskedasticity-robust standard errors for columns (1) – (4). *** p<0.01, ** p<0.05, * p<0.1
[2] Columns (1) – (4) report coefficient estimates for the CareHub variable (=1 if patient is in the CareHub group). Column (1) reports results from a regression without controls; (2) includes controls for demographics; (3) includes controls for demographics and comorbidities; (4) includes controls for demographics, comorbidities, and healthcare utilisation in the 6 months before index admission.

Decomposition of Total Post-Discharge Inpatient Days

As mentioned in the main paper, the outcome variable -- total number of post-discharge days spent in hospital takes value 0 if patients were not readmitted in the 6 months after discharge. This means that the estimated effect of CareHub on total post-discharge days in hospital presented in **Table 5.3** could be due to one of the following three effects:

(i) Fewer CareHub patients are readmitted after their index admission

(ii) A reduction in the number of readmissions among CareHub patients who were readmitted at least once

(iii) A reduction in the mean length of stay per readmission, among CareHub patients who were readmitted at least once.

Our aim in this section is to estimate the relative contributions of the three effects outlined above. To do so, we used methods similar in spirit to techniques used for decomposing changes in poverty[i, ii]. First, we defined the following terms:

- Let MTD_g represent the mean number of days patients in group g spent in hospital after index discharge, where $g \in (CH, UC)$ and CH represents CareHub while UC represents usual care. In other words, for patients in each group g, we averaged their total number of days spent in hospital after discharge.
- MTD_g can also be written as

$$MTD_g = P_g \times MV_g \times MD_g$$

where P_g is the proportion of patients in each group g with at least one readmission, MV_g is the mean number of readmissions for patients in each group g with at least one readmission, and MD_g is the mean length of stay per patient-visit in each group g for patients with at least one readmission.

- The percentage fall in total days spent in hospital after discharge induced by CareHub, $\%\Delta MTD$, can then be written as

$$\%\Delta MTD = 1 - \frac{MTD_{CH}}{MTD_{UC}} = 1 - \frac{P_{CH}}{P_{UC}} \times \frac{MV_{CH}}{MV_{UC}} \times \frac{MD_{CH}}{MD_{UC}}$$

$\frac{P_{CH}}{P_{UC}}, \frac{MV_{CH}}{MV_{UC}}, \frac{MD_{CH}}{MD_{UC}}$ represent the contribution of effects (i), (ii) and (iii) respectively.

We used the above equation to decompose $\%\Delta MTD$ and approximate the relative contributions of each effect by the following method:

1. First, we computed the contribution of effect (i), $\%\Delta MTD_P$, which is given by $\%\Delta MTD - \%\Delta MTD_{\sim P}$, where $\%\Delta MTD_{\sim P}$ is the $\%\Delta MTD$ in the hypothetical absence of a difference between P_{CH} and P_{UC} (i.e. set $\frac{P_{CH}}{P_{UC}} = 1$). $\%\Delta MTD_{\sim P}$ is computed as $1 - \frac{MV_{CH}}{MV_{UC}} \times \frac{MD_{CH}}{MD_{UC}}$.
2. Next, we computed the contribution of effect (ii), $\%\Delta MTD_{MV}$, in a similar way.
3. The contribution of effect (iii) is computed as a residual: $\%\Delta MTD_{MD} = \%\Delta MTD - \%\Delta MTD_P - \%\Delta MTD_{MV}$
4. Steps 1 – 3 were repeated twice, each time starting with either the computation of effect (ii)'s contribution first or the computation of effect (iii)'s contribution first. This gave us the contributions of effects (i) and (ii) as residuals.
5. In all, we obtained three estimates of each effect's contribution– two from direct computation, and one recovered as the residual. For each effect's contribution, we took the average across the three estimates, which we then used to compute their relative contributions to $\%\Delta MTD$.

Table S7 reports the unconditional means we used to compute $\frac{P_{CH}}{P_{UC}}, \frac{MV_{CH}}{MV_{UC}}, \frac{MD_{CH}}{MD_{UC}}$, and hence $\%\Delta MTD$ (i.e. $1 - \frac{P_{CH}}{P_{UC}} \times \frac{MV_{CH}}{MV_{UC}} \times \frac{MD_{CH}}{MD_{UC}}$), while **Table S8** reports the relative contribution of each effect.

Table S7: Summary Statistics for Variables Used in Computation

Variable	CareHub	Control	Ratio (CareHub / Control)
Cardiac-related			
Proportion with at least 1 readmission	0.28	0.33	0.84
Mean no. of visits for patients with at least 1 readmission	1.38	1.85	0.75
Mean LOS per visit for patients with at least 1 readmission	5.94	6.90	0.86
$\%\Delta MTD$ = 1 – 0.84 × 0.75 × 0.86 = 0.46			
Unplanned, cardiac-related			
Proportion with at least 1 readmission	0.25	0.31	0.80
Mean no. of visits for patients with at least 1 readmission	1.32	1.89	0.70
Mean LOS per visit for patients with at least 1 readmission	5.44	6.54	0.83
$\%\Delta MTD$ = 1 – 0.80 × 0.70 × 0.83 = 0.54			

Table S8: Relative Contribution of Each Effect

Variable	1st round of estimation	2nd round of estimation	3rd round of estimation	Mean estimate	Relative contribution to $\%\Delta MTD$
Cardiac-related					
$\%MTD_P$	0.10	0.19	0.10	0.13	29%
$\%MTD_{MV}$	0.18	0.18	0.27	0.21	46%
$\%MTD_{MD}$	0.17	0.09	0.09	0.12	25%
Unplanned, cardiac-related					
$\%MTD_P$	0.12	0.24	0.12	0.16	29%
$\%MTD_{MV}$	0.20	0.20	0.33	0.24	45%
$\%MTD_{MD}$	0.22	0.09	0.09	0.14	25%

Results from Cox Models

We estimated Cox proportional hazard models to study the instantaneous hazard of having (i) an unplanned, cardiac-related readmission; (ii)a cardiac-related readmission; (iii) a specialist outpatient visit; and (iv) and emergency department visit, in the 6 months after index admission. Our estimation considered only the time to first event of each type of hospital utilisation, and we included the same control covariates as those outlined earlier.

Our results indicate that at any point in time, the instantaneous risk of having a readmission or emergency department visit is lower for CareHub patients (see **Table S9**, **Table S10**, and **Table S12**), though the effects do not reach statistical significance. These results are consistent with the effects for outcomes we study in the main paper.

On the other hand, the hazard of having at least one specialist outpatient visit was higher for CareHub patients, and was marginally significant in some specifications (**Table S11**). The direction of this effect is consistent with results for the number of specialist outpatient visits in the main paper.

Table S9: Results from Cox Model Studying First Unplanned, Cardiac-Related Readmission

VARIABLES	(1) No controls	(2) Demo controls	(3) Demo, comorbidity controls	(4) Demo, comorbidity, utilization controls **(Main spec)**
Hazard Ratio	0.740 (0.170)	0.776 (0.186)	0.777 (0.201)	0.710 (0.189)
Observations	270	270	270	270

Notes:
[1] Parentheses report robust standard errors. *** p<0.01, ** p<0.05, * p<0.1
[2] The hazard ratios reported for columns (1) – (4) are from a Cox proportional hazards model, where the main variable of interest is the CareHub variable (=1 if patient is in the CareHub group). Column (1) reports results from a Cox regression without controls; (2) includes controls for demographics; (3) includes controls for demographics and comorbidities; (4) includes controls for demographics, comorbidities, and healthcare utilisation in the 6 months before index admission.
[3] Failure is defined here as having at least one unplanned, cardiac-related readmission in the 6 months post-discharge.

Table S10: Results from Cox Model Studying First Cardiac-Related Readmission

VARIABLES	(1) No controls	(2) Demo controls	(3) Demo, comorbidity controls	(4) Demo, comorbidity, utilisation controls **(Main spec)**
Hazard Ratio	0.796 (0.174)	0.824 (0.189)	0.819 (0.199)	0.768 (0.192)
Observations	270	270	270	270

Notes:
[1] Parentheses report robust standard errors. *** p<0.01, ** p<0.05, * p<0.1
[2] The hazard ratios reported for columns (1) – (4) are from a Cox proportional hazards model, where the main variable of interest is the CareHub variable (=1 if patient is in the CareHub group). Column (1) reports results from a Cox regression without controls; (2) includes controls for demographics; (3) includes controls for demographics and comorbidities; (4) includes controls for demographics, comorbidities, and healthcare utilisation in the 6 months before index admission.
[3] Failure is defined here as having at least one cardiac-related readmission in the 6 months post-discharge.

Table S11: Results from Cox Model Studying First Cardiac-Related Specialist Outpatient Visit

VARIABLES	(1) No controls	(2) Demo controls	(3) Demo, comorbidity controls	(4) Demo, comorbidity, utilisation controls **(Main spec)**
Hazard Ratio	1.246	1.304*	1.283*	1.261
	(0.176)	(0.184)	(0.189)	(0.186)
Observations	270	270	270	270

Notes:
[1] Parentheses report robust standard errors. *** p<0.01, ** p<0.05, * p<0.1
[2] The hazard ratios reported for columns (1) – (4) are from a Cox proportional hazards model, where the main variable of interest is the CareHub variable (=1 if patient is in the CareHub group). Column (1) reports results from a Cox regression without controls; (2) includes controls for demographics; (3) includes controls for demographics and comorbidities; (4) includes controls for demographics, comorbidities, and healthcare utilisation in the 6 months before index admission.
[3] Failure is defined here as having at least one cardiac-related specialist outpatient clinic visit in the 6 months post-discharge.

Table S12: Results from Cox Model Studying First Emergency Department Visit

VARIABLES	(1) No controls	(2) Demo controls	(3) Demo, comorbidity controls	(4) Demo, comorbidity, utilisation controls **(Main spec)**
Hazard Ratio	0.763	0.795	0.800	0.778
	(0.128)	(0.143)	(0.144)	(0.141)
Observations	269	269	269	269

Notes:
[1] Parentheses report robust standard errors. *** p<0.01, ** p<0.05, * p<0.1
[2] The hazard ratios reported for columns (1) – (4) are from a Cox proportional hazards model, where the main variable of interest is the CareHub variable (=1 if patient is in the CareHub group). Column (1) reports results from a Cox regression without controls; (2) includes controls for demographics; (3) includes controls for demographics and comorbidities; (4) includes controls for demographics, comorbidities, and healthcare utilisation in the 6 months before index admission.
[3] Failure is defined here as having at least one Emergency Department Visit in the 6 months post-discharge.

Patient Survey

This section outlines the methodology and results of a one-off survey conducted at the same timepoint (mid 2017)[98] for all study participants. Patients were contacted over the phone and asked to take part in a one-time survey. This survey collected information on quality of life and quality of transitional care for each patient.

A key limitation of this survey is that the time lapse between the index admission and the survey is different across patients. As data from this survey could not be linked to administrative data from NUH[99], we are unable to adjust survey responses for time since index readmission. Our results may thus be biased if the distribution of time since index admission differs for CareHub and control patients who responded to this survey. This concern is mitigated somewhat by the fact that both CareHub and control patients do not differ along most observable demographics in our sample. Nonetheless, findings from this survey should be interpreted with care, and should be seen as suggestive, rather than confirmatory.

Methodology

Survey Questions

Our survey questions were based off validated questionnaires, but were modified to account for the demographics of our study participants. The wording of the questions and responses were simplified to facilitate understanding, as they were generally elderly, less educated, and would be answering the survey over the phone (which posed additional comprehension problems for those

[98] At the start of the study, we planned to carry out four post-discharge surveys at 1 week, 1 month, 2 months, and 6 months after each patient's index admission, to collect information on quality of life and quality of transitional care for each patient. However, response rates were poor, and the eventual sample size of the survey was too low for any analysis to be carried out. We thus decided to carry out this one-off survey to enrich our analysis with additional information on non-utilisation measures that are unavailable in NUH's administrative database.

[99] Due to operational constraints, patients' consent could only be sought over the phone. The ethics board overseeing this study (the Domain Specific Review Board) thus advised us that the research team should not link data from this survey to administrative data from NUH, and that the identity of patients responding to the survey should be kept anonymous.

who were hard of hearing). A small number of questions were also added to collect information on outcomes that NUH was interested in. The questionnaires were translated from English into two other major languages used in Singapore, Mandarin and Malay. A copy of the full questionnaire can be obtained from the authors upon request. Shortened versions of these questions are used as variable labels in **Table S13** to provide information on the questions asked.

The questions on quality of life were based off the EQ-5D-5L[iii]. The wording of some questions was simplified. In addition, descriptors were provided for only the best and worst options for the 5-point scale. Minimal changes were made to the question on general health (i.e. the Visual Analogue Scale). In general, because of the changes we made, our results are not directly comparable to those using the actual questionnaire.

The questions on transitional care focused on patients' confidence in self-care, satisfaction with NUH's post-discharge services, as well as their knowledge of / compliance with prescribed medication. A subset of these questions were based loosely on the Care Transitions Measures®[100] questionnaire[iv] and on the Case Management Adherence Guidelines[v]. Respondents generally answered based on a scale of 1(strongly agree) to 4(strongly disagree). The only exception was a question asking whether patients ran out of medicine because they did not collect them; patients responded either "yes" or "no" for this question. Again, due to the changes we made, our results are not comparable to those using the actual questionnaires.

Outcome Measures

The outcomes of interest are the individual questions described in **Table S13**. Before analysing the survey outcomes, we flipped the response scale where relevant, so that better

[100] © Eric A. Coleman, MD, MPH, all rights reserved

outcomes are tagged with a larger number. E.g., for the question "Do you sometimes run out of medicines because you did not collect them?", "No" is coded as 1, and "Yes" is coded as 0.

Statistical Analysis

We analysed data from this survey by using both an ordered logistic regression and a t-test (to compare differences in means between the CareHub and control groups)[101, vi]. We did not adjust for baseline differences in demographics, as our sample size is relatively small, and including too many control covariates would likely reduce the degrees of freedom in our regression to an extent that affects statistical power. This lack of adjustment, however, is unlikely to be an issue, as both groups did not differ much in terms of baseline demographics.

Results

Table S13 reports results from the survey. The response rates were similar for both CareHub (29%) and control (28%) patients (which is high for this type of study). Both groups do not appear to differ much in terms of observable demographics either. These two observations give us more confidence that non-response bias is unlikely to be a large issue, and that both groups may be less likely to differ along unobservable dimensions (such as time since index admission) either.

Due to the small sample (n=76), most of the differences / odds ratios between CareHub and control patients are statistically insignificant. However, the direction of the estimates for all outcome variables consistently suggest that CareHub led to a higher quality of life, quality of transitional care, and a lower chance that patients would run out of medication.

The evidence of CareHub's positive effect is strongest for anxiety/depression, satisfaction with the hospital's follow-up care, patients knowing who to contact at NUH for their illness-related

[101] The t-test essentially assumes that the response scale is linear, but is unlikely to change our results substantially. Ferrer-i-Carbonell and Frijters show that results from analysing a similar scale variable – life satisfaction – do not depend much on whether the variable is treated as ordinal or cardinal. A comparison of the t-test results with the ordered logit results also show that the results remain the same qualitatively, regardless of method used.

issues, as well as a lower likelihood of running out of medication because patients did not collect their medication. While there is a possibility that the statistically significant improvements for individual outcomes may have arisen due to chance (as we compare many outcome variables), we deem it more likely that real improvements are driving these observed effects, as the improvements are in areas that CareHub was designed to address. E.g., it is reasonable to expect that regular follow-up calls and the provision of a contact point at NUH would reduce a patient's anxiety about managing the recovery process. In addition, we learnt from CareHub staff (in interviews before the statistical analysis was carried out) that they emphasised medication compliance and facilitated the collection of medication during the follow-up calls. Given this, it is not surprising to see that CareHub patients are less likely to run out of medication.

To sum up, the evidence from this section suggests that CareHub did improve patients' quality of life, and the quality of transitional care. The effects appear to be strongest for quality of transitional care, anxiety/depression, and whether patients run out of medication due to not collecting them.

Table S13: Comparison of Survey Outcomes for CareHub vs Non-CareHub Groups

Variable	CareHub			Usual Care			Simple linear difference		Ordered logit results	
	N	Mean	SD	N	Mean	SD	Diff in means	Norm diff	Odds ratio	Std error
Demographics										
Number of household members	40	1.73	1.47	33	2.42	1.58	-0.70*	-0.46	-	-
Number of children	34	2.71	1.36	29	3.03	1.30	-0.33	-0.25	-	-
Has domestic helper	41	0.29	0.46	33	0.30	0.47	-0.01	-0.02	-	-
Presence of main caregiver other than domestic helper	41	0.61	0.49	32	0.59	0.50	0.02	0.03	-	-
At least secondary education	43	0.33	0.47	33	0.33	0.48	-0.01	-0.02	-	-
Housing type									-	-
1-room flat	37	0.11	0.31	32	0.06	0.25	0.05	0.16	-	-
2-room flat	37	0.00	0.00	32	0.03	0.18	-0.03	-0.25	-	-
3-room flat	37	0.30	0.46	32	0.31	0.47	-0.02	-0.03	-	-
4-room flat	37	0.35	0.48	32	0.31	0.47	0.04	0.08	-	-
5-room flat/bigger	37	0.24	0.43	32	0.28	0.46	-0.04	-0.09	-	-
Executive condominium	38	0.03	0.16	33	0.00	0.00	0.03	0.23	-	-
Private housing	38	0.03	0.16	33	0.03	0.17	0.00	-0.02	-	-
Quality of Life										
How well respondent can walk	43	4.14	1.06	33	4.12	1.19	0.02	0.02	0.89	-0.41
How well respondent can wash/dress	43	4.72	0.80	33	4.52	1.09	0.21	0.22	1.90	1.14
How well respondent can carry out usual activities	43	4.56	0.96	33	4.15	1.20	0.41	0.37	2.43*	1.21
Level of respondent's pain / discomfort	43	4.23	1.00	33	4.03	1.10	0.20	0.19	1.45	0.63
Whether respondent is anxious or depressed	42	4.55	0.77	33	4.00	1.12	0.55**	0.57	3.20**	1.50
How good or bad general health is today	42	78.26	20.40	33	74.61	21.24	3.66	0.18	-	-
Quality of Transitional Care										
I know what I need to do to take care of my health	41	3.56	0.55	33	3.42	0.61	0.14	0.23	1.56	0.73

I clearly understand need for each type of medication	41	3.54	0.67	33	3.33	0.65	0.20	0.31	2.08	0.97
I know who to call at NUH	39	3.23	0.84	33	2.82	0.88	0.41**	0.48	2.62**	1.20
I am satisfied with the hospital's follow-up care	39	3.62	0.49	32	3.16	0.81	0.46***	0.69	3.56***	1.73
Hospital made sure post-discharge needs taken care of before discharge	39	3.54	0.64	33	3.36	0.70	0.17	0.26	1.71	0.80
Given clear and complete instructions on what to do before leaving hospital	39	3.46	0.72	33	3.30	0.68	0.16	0.23	1.77	0.83
Do you sometimes run out of medicines because you did not collect them	41	0.95	0.22	33	0.82	0.39	0.13*	0.42	4.33*	3.73

Notes:
[1] *** p<0.01, ** p<0.05, * p<0.1.
[2] We flipped the values for all outcome variables so that a larger number represents a better outcome.
[3] We applied t-test for column "Diff in means". Normalised differences as captured in column "Norm diff" are computed as the difference in means, standardised by the square root of the mean variance from both groups.
[4] Odds ratios are from bivariate ordered logit regressions, where the only independent variable is a dummy variable assignment to CareHub. Robust standard errors for the odds ratios were obtained using the delta method.
[5] Unlike other questions in the transitional care section, the response to the last question is a binary variable that takes value 1 if patient replies no.

Clinical indicators – diastolic and systolic blood pressure

Table S5.1: Effect of CareHub on diastolic and systolic blood pressure

Variable	(1) Mean diastolic blood pressure	(2) Mean systolic blood pressure
CareHub	-1.445	-1.074
	(1.210)	(2.427)
Observations	237	238
R-squared	0.200	0.200

Notes:
[1] Heteroskedasticity-robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
[2] Estimates of the effect of CareHub are from a regression model that includes controls for demographics, co-morbidities, healthcare utilisation in the 6 months before the index visit, and mean number of days from index discharge date on which blood pressure is measured. The last control covariate is added as blood pressure can be measured more than once post index discharge.
[3] The dependent variables were computed by averaging each patient's multiple measurements of blood pressure.

Comparison of NUH vs non-NUH utilisation for CareHub and non-CareHub groups

Our results in **Table 5.3** and **Table 5.4** of the main paper are based only on NUH's administrative data, and do not capture healthcare utilisation at non-NUH institutions. These estimates could be biased if the ratio of NUH to non-NUH healthcare utilisation differs for CareHub and usual care patients. In this section, we use National Healthcare Group (NHG) administrative data[102] from 2016 to study whether such a bias exists. (We do not use this dataset for our main analysis, as it does not allow us to follow the vast majority of our patients for 6 months after their index admission.)

Table S14 shows that before patients' index admission, the ratio of NUH to non-NUH healthcare utilisation was very similar for both CareHub and usual care patients, across all types of utilisation studied. After patients are enrolled on CareHub, however, NUH healthcare utilisation as a proportion of total healthcare utilisation increases for CareHub patients, but not for control patients.

This suggests that our estimates in the main paper are attenuated. We illustrate this using a simple numerical example on post-index readmissions. Suppose that NUH administrative data indicates there is, on average, 5 readmissions for each CareHub patient, and 10 for each control patient. The estimated effect of CareHub would then be -5; i.e. CareHub leads to 5 fewer readmissions per patient over a 6-month follow-up period. However, NUH data captures only 90% of all cardiac-related hospital readmissions for CareHub patients, and 80% for non-CareHub patients. Once we take non-NUH admissions into account, there are 5.5 readmissions per CareHub patient and 12.5 per control patient. The estimated effect of CareHub would then be -7, implying that using only NUH data would attenuate our estimates of CareHub's effects.

[102] The NHG is a regional healthcare system for the central region in Singapore. NUH administrators believe that most non-NUH healthcare use by NUH patients is concentrated in the two other nearby hospitals in the NHG system: Ng Teng Fong General Hospital (NTFGH) and Tan Tock Seng Hospital (TTSH). Thus, even though we do not have data on healthcare utilisation for all hospitals in Singapore, our analysis in this section should capture most of the non-NUH healthcare utilisation by patients in our sample.

Table S14: Comparison of Non-NUH Healthcare for CareHub and Non-CareHub Groups

Pre-CareHub Enrollment				Post-CareHub Enrollment			
Hospital	Usual Care	Care-Hub	Total	Hospital	Usual Care	Care-Hub	Total
All inpatient admissions				**All inpatient admissions**			
NTFGH	43	36	79	NTFGH	24	11	35
NUH	144	148	292	NUH	72	99	171
TTSH	8	1	9	TTSH	3	1	4
Total	195	185	380	Total	99	111	210
% NUH	0.74	0.80		% NUH	0.73	0.89	
Relative % (Carehub / usual care)		1.08		Relative % (Carehub / usual care)		1.23	
Cardiac-related inpatient admissions				**Cardiac-related inpatient admissions**			
NTFGH	9	15	24	NTFGH	7	3	10
NUH	67	64	131	NUH	41	40	81
TTSH	5	0	5	TTSH	3	1	4
Total	81	79	160	Total	51	44	95
% NUH	0.83	0.81		% NUH	0.80	0.91	
Relative % (Carehub / usual care)		0.98		Relative % (Carehub / usual care)		1.13	
Unplanned, cardiac-related inpatient admissions				**Unplanned, cardiac-related inpatient admissions**			
NTFGH	9	15	24	NTFGH	7	2	9
NUH	63	64	127	NUH	40	32	72
TTSH	5	0	5	TTSH	3	1	4
Total	77	79	156	Total	50	35	85
% NUH	0.82	0.81		% NUH	0.80	0.91	
Relative % (Carehub / usual care)		0.99		Relative % (Carehub / usual care)		1.14	
All specialist outpatient visits				**All specialist outpatient visits**			
IMH	3	2	5	IMH	2	0	2
NTFGH	24	44	68	NTFGH	7	14	21
NUH	430	483	913	NUH	199	338	537
TTSH	13	5	18	TTSH	3	4	7
Total	470	534	1,004	Total	211	356	567

Pre-CareHub Enrollment				Post-CareHub Enrollment			
Hospital	Usual Care	Care-Hub	Total	Hospital	Usual Care	Care-Hub	Total
% NUH	0.91	0.90		% NUH	0.94	0.95	
Relative % (Carehub / usual care)		0.99		Relative % (Carehub / usual care)		1.01	
Cardiac-related specialist outpatient visits				**Cardiac-related specialist outpatient visits**			
NUH	125	139	264	NUH	80	121	201
Total	125	139	264	Total	80	121	201
Emergency Department visits				**Emergency Department visits**			
NTFGH	49	48	97	NTFGH	28	17	45
NUH	267	306	573	NUH	94	95	189
TTSH	12	6	18	TTSH	4	1	5
Total	328	360	688	Total	126	113	239
% NUH	0.81	0.85		% NUH	0.75	0.84	
Relative % (Carehub / usual care)		1.04		Relative % (Carehub / usual care)		1.13	

Abbreviations:
NTFGH: Ng Teng Fong General Hospital
NUH: National University Hospital
TTSH: Tan Tock Seng Hospital
IMH: Institute of Mental Health

References for Supplementary Materials

[i] Shorrocks AF. Decomposition procedures for distributional analysis: a unified framework based on the Shapley value. Journal of Economic Inequality. 2013:1-28.
[ii] Datt G, Ravallion M. Growth and redistribution components of changes in poverty measures: A decomposition with applications to Brazil and India in the 1980s. Journal of development economics. 1992;38(2):275-95.
[iii] The EuroQol Group. EuroQol-a new facility for the measurement of health-related quality of life. Health policy. 1990;16(3):199-208.
[iv] Coleman EA, Mahoney E, Parry C. Assessing the quality of preparation for posthospital care from the patient's perspective: the care transitions measure. Medical care. 2005:246-55.
[v] Case Management Society of America. Case Management Adherence Guidelines 2012.
[vi] Ferrer- i- Carbonell A, Frijters P. How important is methodology for the estimates of the determinants of happiness? The Economic Journal. 2004;114(497):641-59.