THE UNIVERSITY OF QUEENSLAND
AUSTRALIA

Essays in Welfare Economics

Thi Thao Nguyen

BEcon, MEcon, MAdvEcon

Thi Thao Nguyen's ORCID

A thesis submitted for the degree of Doctor of Philosophy at The University of Queensland in 2022

School of Economics

Abstract

This dissertation consists of three separate essays on resource allocation that provide valuable empirical and theoretical contributions to the welfare economics literature. In the first essay, we develop economic models to identify at-risk communities that assist governments and nongovernmental organisations (NGOs) to best target resources. Similar models developed to date neglect the impact of weather shocks. Because climate change is predicted to increase the frequency and severity of weather shocks that adversely affect agricultural crops, in this paper, we study the impact of weather shocks on household welfare and how it exacerbates household income inequality via increasing crop income inequality. We first recognise the limitations of existing measures of weather shocks and propose an absolute measure of weather shocks that is more robust to the length of time over which weather samples are obtained. Next, we study the impact of the newly constructed weather shocks on household welfare measured by different income sources and different types of consumption. The findings suggest that weather shocks reduce crop revenue significantly, and that their impact varies across households with different characteristics. Next, we consider how the diverse impact of weather shocks affects household income inequality. The Gini decomposition of income sources suggests crop income reduces income inequality in rural areas. Overall, weather shocks reduce income from crops and therefore they increase income inequality.

Although income inequality largely dominates the debate on inequality, non-income disparities do exist. In the second essay, we study gender inequality in mental health scores using Australian panel data. We show that men have significantly higher mean outcomes, and lower variances, such that the left tail of the combined distribution is disproportionately female. Using regression-based decompositions, we examine the degree that economic inequalities account for this phenomenon. We find that disparities in income play a very substantial role, enough to account for the gender gap amongst individuals with very poor psychological wellbeing. We also examine the mental health effects of various negative life events, such as the death of a family member or being a victim of violence. At the individual level, these characteristics have large effect sizes but are not correlated strongly enough with gender to explain the aforementioned mental health disparities.

In the third essay, we examine a more general framework of resource allocation by studying the probabilistic allocation of finitely many indivisible objects to finitely many agents. Well-known allocation rules for this problem include random priority, the market mechanism proposed by Hylland and Zeckhauser (1979), and the probabilistic serial rule of Bogomolnaia and Moulin (2001). We propose a new allocation rule, which we call the *lexicographic serial rule*, that is tailored for situations in which each agent's primary concern is to maximise the probability of receiving her favourite object. Three axioms – *lex efficiency*, *lex fairness*, and *lex envy-freeness* – are proposed. We also discuss how our axioms and the lexicographic serial rule are related to

other allocation rules, particularly the probabilistic serial rule.

Declaration by author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, financial support and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my higher degree by research candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the policy and procedures of The University of Queensland, the thesis be made available for research and study in accordance with the Copyright Act 1968 unless a period of embargo has been approved by the Dean of the Graduate School.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate I have obtained copyright permission from the copyright holder to reproduce material in this thesis and have sought permission from co-authors for any jointly authored works included in the thesis.

Publications included in this thesis

No publications included.

Submitted manuscripts included in this thesis

- Nguyen, T. T. *Weather Shocks and Income Inequality.*

Other publications during candidature

Conference abstracts

- **Nguyen, T.T.**, Takayama, S., and Phan, B. N. (2019). Inequality of Education: A Model of School Choices. Presented at the Inequality of Opportunity Conference, UQ and Griffith University.
- **Nguyen, T.T.**, Takayama, S., and Yeo, T. (2020). Individual Manipulability in Large Elections. Presented at the 33 PhD Conference in Business and Economics, Monash University.
- **Nguyen, T.T.** (2022). Weather Shocks and Income Inequality. Presented at the International TVSEP Conference on Shocks and Resilience in Rural Southeast Asia.

Contributions by others to the thesis

- Chapter 2: *Weather Shocks and Income Inequality.* Nguyen, T. T. and Rao, D. S. Prasada. developed the initial ideas and designed the study (5%). Nguyen, T. T. conducted data analysis, interpreted results, prepared tables and figures, and wrote the paper (90%). Rao, D. S. Prasada., McLennan, A., and Takayama, S. provided supervision and guidance and critically reviewed the chapter (5%).
- Chapter 3: *Economic Disparity, Life Events and the Gender Mental Health Gap.* Nguyen, T. T., Rohde, N., and Nguyen, K. H. conceptualised and designed the project (10%). Nguyen, T. T. conducted data analysis, interpreted results, prepared tables and figures, and wrote the paper (80%). Rohde, N., and Nguyen, K. H. critically reviewed the chapter (10%).
- Chapter 4: *The Lexicographic Serial Rule* - Nguyen, T. T., McLennan, A., and Takayama, S. conceived and designed the study (10%). Nguyen, T. T. and Takayama, S. proposed motivating examples and initial results (20%). Nguyen, T. T., McLennan, A.,

and Takayama, S. wrote the paper (40%). Nguyen, T. T., developed additional lemmas and finalised the chapter (30%).

Statement of parts of the thesis submitted to qualify for the award of another degree

No works submitted towards another degree have been included in this thesis.

Research involving human or animal subjects

No animal or human subjects were involved in this research.

Acknowledgments

I would like to thank my current and former advisers Kim-Huong Nguyen, Jason Pole, Prasada Rao, Shino Takayama, and Andrew McLennan for their assistance, encouragement and general camaraderie throughout my academic career. Without their advice and support, this thesis would be much impoverished.

Special thanks to Nicholas Rohde and Trong Anh Trinh for our invaluable academic discussions. Our exchanges were instrumental in uncovering useful research directions and honing ideas.

I would like to thank the data custodians of the Thailand Vietnam Socio Economic Panel (TVSEP) and the Household, Income and Labour Dynamics in Australia (HILDA) projects for making the datasets available for this thesis. Specifically, I thank Trung Thanh Nguyen, Manh Hung Do and Florian Heinrichs at Leibniz University Hannover for granting me access to the TVSEP data and assisting me in understanding the dataset, which was used in Chapter 2 of this thesis. The unit record data from the HILDA Survey used in Chapter 3 were obtained from the Australian Data Archive, which is hosted by the Australian National University. The HILDA Survey was initiated and is funded by the Australian Government Department of Social Services (DSS) and is managed by the Melbourne Institute of Applied Economic and Social Research (Melbourne Institute).

I am grateful for the IT and ergonomic support that I received at the University of Queensland, from Nathanael Elcock and Chris Pye in particular. I would like to acknowledge the writing advice received from the learning advisor David Rowland and thank Campbell Aitken, who provided professional editing services in accordance with the Institute of Professional Editors' Guidelines for editing research theses.

Outside the academic sphere, I am grateful for the psychological support of my family and friends. Thanks to all friends and mentors that I have been lucky to have in the last few years of this program: Hien Pham, Anh Tran, Dao Nguyen, Tim Ludlow, Hoang Nguyen, Ngoc Nguyen, Thanh Ha, Nhan Phan, Toan Le, Toan Nguyen, Alice Tran, Giang Pham, Bhagya Gunawardena, Aarushi Dhingra, Elcin Tuzel, Catherine Belward, and Louisa Coglan.

I am immensely grateful to my parents and brother for always being supportive of my choices. Even though far away, their encouragement and emotional support made me feel their presence all along.

I cannot thank my fiancé, Hieu Tran, enough for his love, support and endurance throughout the years of this program. Thanks for always being by my side and believing in me even when I feel lost and doubted myself. The last few years have been so difficult and without your love and support, this thesis would not exist.

Financial support

This research was supported by a University of Queensland Research Training Stipend, University of Queensland Research Training Tuition Fee Offset, and Research Higher Degree Top Up Scholarship.

Keywords

weather shocks, income inequality, decomposition method, welfare, mental health, gender gap, lexicographic serial rule

Australian and New Zealand Standard Research Classifications (ANZSRC)

- ANZSRC code: 140104, Microeconomic Theory, 30%
- ANZSRC code: 140219, Welfare Economics, 40%
- ANZSRC code: 140304, Panel Data Analysis, 30%

Fields of Research (FoR) Classification

- FoR code: 1401, Economic Theory, 33%
- FoR code: 1402, Applied Economics, 67%

To Hieu Tran

Table of Contents

List of Figures

List of Tables

List of Abbreviations and Symbols

Abbreviations	
CPI	Consumer Price Index
CQR	Conditional Quantile Regression
DFG	Deutsche Forschungsgemeinschaft – German Research Foundation
ELWM	Egalitarian Lexicographic Welfare Maximiser
EM-DAT	Emergency Events Database
FE	Fixed Effects
sd	(First Order) Stochastic Dominance
HILDA	Household, Income and Labour Dynamics in Australia
MHI	Mental Health Inventory
NASA	National Aeronautics and Space Administration
NGOs	Nongovernmental Organisations
OECD	Organisation for Economic Co-operation and Development
PPP	Purchasing Power Parity
RIF	Recentered Influence Function
SF-36	36-Item Short Form Survey
TRMM	Tropical Rainfall Measuring Mission
TVSEP	Thailand Vietnam Socio Economic Panel
UQR	Unconditional Quantile Regression
USD	United States dollars
vNM	von Neumann-Morgenstern
WELWM	Weak Egalitarian Lexicographic Welfare Maximizer

Symbols	
I	Finite set of agents with cardinality of n
A	Finite set of objects with cardinality of m
$\alpha \in \Re^I_{++}$	Vector of positive appetites
$q \in \Re^A_{++}$	Vector of positive quotas
p_{ia}	Probability that agent i receives object a
$p = (p_{ia})_{n\times m}$	An $n \times m$ matrix of random assignment
p_i	Row i of p, a vector in $\Re^A_+$, representing the allocation for agent i
$a \succ_i b$	Object a is preferred to object b by agent i
$\mathcal{P}$	Set of preferences
$\succ = (\succ_i)_{i\in I} \in \mathcal{P}^I$	A preference profile
$f_k(\succ_i)$	The k^{th} favourite of i
$I(a, k, \succ)$	Set of agents for whom a is the k^{th} favourite.
$c_i(r)$	Agent i's total consumption at r
$d_a(r)$	The total consumption of a at r
$U(\succ_i, a) = \{b \in A : b \succsim_i a\}$	The weak upper contour set of a
$F(\succ_i, a, p_i)$	i's consumption of objects that are at least as good as a
$\lambda(\succ)$	The random assignment under Lexicographic Serial Rule

Chapter 1

Introduction

1.1 Motivation and Objectives

This thesis includes three separate essays that use a mixture of empirical studies and theoretical analyses to address a fundamental issue in economics – allocating scarce resources. Examples of allocating scarce resources include students deciding how to divide their time for study, work, and leisure, and governments deciding how budgets are spent on priorities such as the military, health and education. The study of scarce resource allocation in economics centres on three questions.

1. How does resource allocation occur (positive economic question)?
2. How should resource allocation occur (normative economic question)?
3. Are there institutions to make what should occur to be achieved in equilibrium (design question)?

Economists study these resource allocation questions by analysing data from real-world situations, or building theoretical models that are an abstraction from the real world but allows us[1] to study the most essential factors and relationships by removing many real-world complications. As Alpha C Chiang states in the book, *Fundamental Methods of Mathematics*:

> ...empirical studies and theoretical analyses are often complementary and mutually reinforcing. On the one hand, theories must be tested against empirical data for validity before they can be applied with confidence. On the other, statistical work needs economic theory as a guide, in order to determine the most relevant and fruitful direction of research. (pp. 5-6)

[1] In this thesis, "us" means economists in general.

We[2] take advantage of both approaches, with the first two essays being empirical in nature and the last essay being theoretical work. The two empirical essays apply different econometric techniques in analysing real-world data to assist policymakers in allocating resources. The theoretical model provides us with tools to draw a set of conclusions from a given set of assumptions, utilising a wealth of existing mathematical theorems. It also forces us to be explicit about the assumptions we make and allows us to address the more general cases.

The purpose of the research described in this thesis is to show policymakers how to improve people's lives in terms of welfare measures; therefore, we need to discuss what welfare means and how to evaluate welfare changes. Whilst positive economics is about understanding where supply and demand come from and their implications for resource allocation, normative economics is concerned with making a judgment about whether an allocation is good or bad according to certain criteria. To decide how resources should be allocated (i.e., answer the normative question), we need to decide on appropriate criteria to evaluate alternative allocations.

Pareto efficiency is considered a reasonably uncontroversial criterion to assess the success of an allocation in economics; however, Pareto efficiency alone is often insufficient. Under a Pareto efficient allocation, it is impossible to make anyone better off without making someone else worse off. The two fundamental theorems of welfare economics (see, e.g., Mas-Colell et al., 1995) assure us that through appropriate redistribution of initial endowments and the operation of a market, it is possible to achieve a Pareto efficient allocation. However, questions have been raised about the degree to which the assumptions underlying these two fundamental theorems are realistic. For example, people are not always rational and are capable of acting in seemingly irrational ways, especially when experiencing scarcity (Mullainathan and Shafir, 2013). Furthermore, the fundamental theorems do not account for externalities. For example, a new irrigation system for a farmer upstream could cause a loss in output for a farmer downstream due to reduced water flow. Since the idealised assumptions of the two fundamental theorems are often not met to ensure efficient allocation, policy intervention is often required. Most importantly, Pareto optimality has little implication on the distribution of welfare across individuals because for any given distribution with fixed total resources, it will always be considered Pareto optimal because any distribution that makes someone better off will make someone else worse off (Kakwani and Son, 2016).[3] Nevertheless, efficiency is still an important criterion in allocating resources and is discussed further in Chapter 4.

Another criterion for evaluating alternative allocations and allocation mechanisms, which is the focus of Chapters 2 and 3, is *equity*, which, roughly speaking, means "equal treatment of equals"; yet there often exists a tradeoff between equity and efficiency in many economic problems.

[2]In this thesis, "we" means myself and my co-authors.

[3]This is due to a reasonable assumption in economics called "nonsatiation" which essentially means that "more of a good is always better".

Economic inequity is often understood as the ethically unjustifiable differences between people in their command over economic resources. The value we place on equity is the main reason for our concern with inequality.[4]

To discuss inequality, it is necessary to impose some normative judgements on how individuals' utilities should be compared and aggregated into "social welfare". This is a controversial issue, discussion of which relies heavily on the philosophical and ethical assumptions about the value to be placed on different individuals' wellbeing (see, e.g., Fleurbaey and Maniquet, 2011; Fleurbaey, 2012; Backhouse et al., 2021; Sen, 1984). The most important two principles in developing practical social decision-making are *utilitarianism* and *egalitarianism*. The utilitarian principle asserts that the best policy is the one that maximises the sum of individual utilities. The egalitarian principle asserts that the best policy is the one that maximises welfare as long as all individuals enjoy equal benefits. When different individuals enjoy different levels of utility from consuming an object, the egalitarian principle would lead to the same social choices as the Rawlsian principle, which maximises the utility of the most unfortunate individuals in society (also called the *maximin* principle). As stated by Myerson (1981):

> Translated into the practical debates of daily life, the utilitarian principle asserts that "you should do something for me if it will hurt you less than it will help me," whereas the egalitarian principle asserts that "you should do something for me if you are better off than I am (or if you have gained more from our cooperation than I have)." (p. 883)

Underlying any measure of inequality is some concept of social welfare, and thus, the form of the social welfare function employed (Atkinson, 1970). Social welfare functions – developed by Bergson in 1938 and further refined by Samuelson in 1947 – are considered the most appropriate tools for aggregating individual utilities into social welfare (Kakwani and Son, 2016). In general, a social welfare function can be defined as a function of individual utilities. However, to be useful in evaluating public policy, in which normative judgments are unavoidable, a particular functional form of social welfare function needs to be specified. The utilitarian principle discussed above underpins one of the most widely used approaches to aggregate individual utilities – that is, to construct social welfare functions by summing up all individual utility functions.

One could argue that if the gain from the resource reallocation is more than enough to compensate those who have been harmed by it, then by employing the utilitarian approach, there will be a net benefit to society – that is, an improvement in social welfare. This is due to another standard assumption in economics that every individual has a concave utility function, which

[4] Economic inequality is not the same concept as economic inequity. Economic inequality is the differences between people's command over economic resources.

means that for a given increase in resources, a poor person would derive more satisfaction from it than a rich one; therefore, it is possible to increase overall welfare by redistributing resources. Consequently, many social policies, such as progressive income taxation, are attempts of governments to redistribute resources from the rich to the poor to increase social welfare. This is also the reason why the focus of Chapters 2 and 3 is the less fortunate individuals in society.

Rising inequality worldwide has been well-documented (Dabla-Norris et al., 2015), and should be a great concern to the public if it is deemed "unfair". If individuals are not responsible for the level of endowment they receive; this raises the question: What should be done to improve their welfare once the possibility of government transfer exists? For example, many governments have a fixed budget that can be used to provide targeted help to particular categories of people, such as those who suffer from natural disasters or experience physical violence, which are the topics we consider in Chapters 2 and 3.

Extensive literature has examined the impact of income inequality and its causes. Weather shocks are likely to be a cause of income inequality, but have received little investigation to date. In Chapter 2, we aim to establish a link between weather shocks and income inequality. If weather shocks drive income inequality, there is a strong case for governments to ameliorate their impacts, because this inequality is due to circumstances that fall outside of an individual's control rather than being due to a lower level of effort (Roemer, 1998). Governments need to know if weather shocks contribute to increasing income inequality and how to identify the affected communities to intervene effectively.

To understand the impact of weather shocks on household welfare and eventually, on income inequality, we seek to answer the following four research questions.

1. What are the impacts of weather shocks on the income sources of rural households? This is of interest because different households might have different strategies to diversify their income sources. Poor households might not have access to credit and thus, struggle to escape the poverty trap.

2. How do weather shocks affect households' consumption? This is relevant to inequality because households which lack assets to offset lost income may be forced to lower their consumption. Then, the question is, what do they stop consuming – entertainment, education, or medical treatment? If they have to take their children out of school or marry off their daughters early (depending on the culture), then weather shocks have an indirect but significant negative impact on their children's lives and intergenerational mobility.

3. How do the impacts of weather shocks vary for households with different characteristics? If all household incomes are altered by the same proportion, then relative inequality would

stay the same even though absolute inequality might rise sharply.[5] How can we capture both relative and absolute measures of inequality?

4. How do weather shocks exacerbate income inequality? This final research question recognises that weather shocks are likely to have doubly negative impacts on household welfare, both reducing income and increasing income inequality, making some households worse off.

Although income inequality tends to dominate the debate on inequality, non-income disparities exist and often have strong interactions with income inequality. In Chapter 3, we study the gender mental health gap, for several reasons. First, society should be concerned with inequality in multiple dimensions of wellbeing, especially health (Sen, 2002). Mental health is an integral and essential component of health, and mental illness has wide-ranging consequences for both individuals and communities. For example, depression and anxiety cost the global economy around USD 1 trillion annually in lost productivity (Chisholm et al., 2016). In addition, mental illnesses are common in the Australian community, with nearly one in two (46%) Australians aged 16–85 having experienced a mental disorder during their lifetime (Australian Bureau of Statistics, 2007).

Another reason for studying the gender mental health gap is because gender equality is a fundamental human right. Achieving gender equality is essential to achieving peaceful societies that maximise human potential and enjoy sustainable development. On the one hand, gender inequality often involves women and girls being deprived of access to various services, such as education, health care or proper nutrition.[6] On the other hand, gender inequality affects men, who may be unable to live up to traditional stereotypes, less likely than women to seek professional help or talk about their problems with friends or family, and more likely to commit suicide.[7]. Therefore, understanding the mental health distribution and determinants of the gender mental health inequalities is essential to developing effective and equitable economic policies.

We observe that men have significantly higher mental wellbeing on average compared to women in Australia (measured by the Mental Health Inventory [MHI-5] score, from the Household, Income and Labour Dynamics in Australia [HILDA] survey (Watson and Wooden, 2021)). In addition, a higher proportion of women than men have low psychological well-being. Since there are strong correlations between health and socioeconomic variables, such as income and

[5] Relative inequality measures are means-independent, i.e., they stay constant if every income is altered by the same proportion. In contrast, absolute inequality measures will remain constant if every income is increased or decreased by the same amount. For example, if we have an economy with just two households with an annual income of $20,000 and $100,000, relative inequality would stay the same if both household incomes were doubled. However, the absolute difference would have risen from $80,000 to $160,000.

[6] https://www.un.org/sustainabledevelopment/wp-content/uploads/2018/09/Goal-5.pdf

[7] https://www.vic.gov.au/gender-inequality-affects-everyone

education (see, e.g., Butterworth et al., 2009; Morasae et al., 2012), we ask if economic differentials between men and women are sufficient to explain the observed gender mental health inequality. In addition, since negative life events, such as being a victim of physical violence, tend to harm mental health, we also consider their role in driving the gender mental health gap.

We aim to exploit an econometric technique that goes beyond the mean and provides a full-distribution decomposition to determine the structure of the mental health inequality. This is particularly relevant in the case of mental health, because the social and economic cost of very poor mental health is often catastrophic; that is, people at the very low end of the mental health spectrum can be suicidal or incapable of engaging in productive activity, including education, work and social engagements, and they often require a much higher level of healthcare interventions (inclusive of medications, therapies and social support). Our paper represents (to our knowledge), the first piece of research to examine quantile effects on the gender mental health gap.

Chapter 4 is concerned with a specific type of resource allocation problem – allocating indivisible objects when monetary transfers are not allowed.[8] Various real-life problems require distributing indivisible objects, which it is ethically unacceptable for people to use money to purchase, to a set of individuals. Examples include the assignment of teaching responsibilities to faculty members, the provision of social housing, or the assignment of treatments to patients, just to name a few.

In such cases we would expect the assignment of resources to be Pareto optimal, and if the agents' preferences of objects are known, it could easily be achieved using a job-assignment algorithm. When the preferences are unknown, we would like to find a mechanism that, when agents respond to it (and reveal their preferences), produces an efficient outcome and satisfies prescribed distributional objectives, whether it is treating everybody equally or favouring certain individuals systematically. An assignment problem can be deterministic or random. We focus exclusively on random assignments because the classical Birkhoff-von Neumann theorem implies that any random assignment can be induced by a probabilistic distribution over deterministic assignments (Budish et al., 2011).

The oldest rule used for this setting is *random priority*, in which an equiprobable lottery chooses a first agent, who receives her favourite object, if there is one she prefers to receiving no object. After that, another equiprobable lottery chooses a second agent, who receives her favourite among the objects not chosen by the first agent if one of them is preferred to receiving no object, and so forth. Despite exhibiting several advantages such as being ex-post Pareto

[8]Objects do not care who they are allocated to so this is not the well-known marriage problem.

efficient,[9] strategy-proof,[10] and fair,[11] random priority is not envy-free[12] and may be inefficient in the sense that another random assignment can stochastically dominate the one produced by random priority for all agents, and strictly for some (Bogomolnaia and Moulin, 2001).

A second solution to the random assignment problem is the well-known market mechanism proposed by Hylland and Zeckhauser (1979) that depends on agents' cardinal preferences.[13] It views a von Neumann-Morgenstern (vNM) utility function over random allocations of objects as a linear utility over vectors of the probability of receiving objects in the eventual assignment. Although this solution is envy-free and efficient with respect to the profile of vNM utility functions, it is not strategy-proof (Bogomolnaia and Moulin, 2001). In fact, there is no efficient and equitable assignment rule based on cardinal preferences that is strategy-proof (Zhou, 1990). In addition, the representation of preferences over uncertain outcomes by vNM utility functions can be inadequate because untrained agents often do not know what they are. Also, the formulation of rational preferences over a given set of lotteries is a complex process, and most agents choose to avoid it if they can. Consequently, in Chapter 4, we relax this assumption and focus on ordinal mechanisms.

With a focus on ordinal preferences, the *probabilistic serial rule* was proposed to overcome the limitations of the random priority rule (Bogomolnaia and Moulin, 2001). It is described as a "cake-eating" algorithm, in which objects are thought of as "cakes" of the probability of unit size. Over the unit interval of time, the agents "eat" cake at unit speed, at each point in time consuming the favourite object whose cake has not yet been exhausted if there is one that is preferred to receiving no object. The probabilistic serial rule is *sd efficient*,[14] sd envy-free, and fair, but it is not strategy-proof. If an agent is almost indifferent between two objects, she might do better by first eating the cake of the one she likes less well. Furthermore, although the probabilistic serial rule is welfarist in the sense that for each profile, the outcome maximises a sum of expected utilities for some vector of vNM utility functions that is consistent with the given preferences,[15] there is no consistency across profiles: changing one agent's ordinal ranking can change the utility functions of the other agents.

The objective of Chapter 4 is to propose another assignment rule that depends only on ordinal preferences and is tailored for lexicographic preferences. In other words, agents are primarily concerned with maximising the probability of receiving their most favourite object; they care

[9] There is no ex-post trade that makes all the traders better off.

[10] It is never possible to benefit by reporting a false preference.

[11] Agents and objects are treated symmetrically.

[12] Each agent's probabilistic assignment first-order stochastically dominates (for their preferences) each other agent's probabilistic assignment. See Appendix C.2 for a revision of stochastic dominance.

[13] It is not, strictly speaking, an allocation rule because it is possible that the market has multiple equilibria.

[14] There is no reallocation of probability that gives each agent a first-order stochastically dominating probabilistic assignment.

[15] Bogomolnaia (2015) characterises the probabilistic serial rule as a kind of Rawlsian welfare maximiser.

much more about the difference between the second and third favourite than the difference between the third and fourth favourite, and so on.

In conclusion, the main objective of this thesis is to show how to improve people's lives (in terms of welfare measures) by allocating scarce resources efficiently and equitably. In Chapter 2 we develop a model for identifying communities that are most affected by weather shocks, and in Chapter 3 identify factors that contribute to the gender mental health gap. Findings in these two chapters can guide policymakers in designing effective policies to support those in need. Chapter 4 proposes an allocation rule that improves welfare when agents exhibit lexicographic preferences, which can then be applied to the context of Chapters 2 and 3.

1.2 Thesis Structure

Each of the following three chapters presents a standalone essay that begins with greater detail about our motivation and our identification of the research gap from existing literature. The empirical chapters (2 and 3) follow by giving descriptions of data and variables of interest. After that, we justify our choice of econometric techniques to answer the research questions. Then the results are presented, together with a discussion and conclusion. In the theoretical chapter (4), a model is presented and the proposed allocation rule defined. This is followed by a discussion of axioms that the new allocation rule satisfies in relation to other allocation rules existing in the literature.

Chapter 5 summarises the achievements of the research with respect to the objectives laid out in Section 1.1 and evaluates how successful they are in meeting those objectives. It also discusses how findings are relevant to policymakers, how the allocation rule proposed in Chapter 4 can be applied in the context of Chapters 2 and 3, and how our results can be useful for future economic research.

Chapter 2

Weather Shocks and Income Inequality

2.1 Introduction

Growing income inequality has become a major concern for governments and policymakers over the last few decades due to its impact on the economy, democracy and justice systems. First, societies with high levels of income inequality do not function efficiently, and their economies are neither stable nor sustainable in the long run (Stiglitz, 2012). Second, an increase in income inequality lowers consumption because rich people tend to consume a smaller proportion of their income than do the poor (Dynan et al., 2004); this leads to more unemployment in the short run, because total demand is lower than what the economy is capable of supplying. Thus, for example, if governments wish to stimulate the economy during a recession, giving money to the rich, who consume a smaller portion of their income, is ineffective.

At the household level, if economic status is associated with a given outcome, then an increase in economic inequality will lead to an increase in inequality in that outcome (Neckerman and Torche, 2007). For example, if higher income means people are happier, then higher income inequality would result in a wider gap in happiness measures. The negative association between income inequality and happiness is reported not only in advanced economies, but in emerging ones (Oishi et al., 2011; Tran et al., 2018). Furthermore, high income inequality is associated with high prevalences of mental illness and drug misuse in rich societies (Pickett and Wilkinson, 2010). Hence, reducing income inequality by redistributing income to those for whom it is more efficacious may improve many important social outcomes, such as health and education.

Attempts have been made to identify key factors that drive income inequality to assist governments in developing targeted programs to reduce inequality. A crucial factor is technological change, which disproportionately raises the demand for skilled labour over low-skilled and unskilled labour. Technological change eliminates many routine jobs via automation or by requiring higher skill levels to attain or remain in jobs (Acemoglu, 1998; Card and DiNardo, 2002). Thus, new information technology has driven up the skill premium, resulting in an increase in labour income inequality. Other factors have been identified as contributing to increased income inequality, including international trade, the flow of foreign direct investment from advanced to emerging economies, and changes in labour market institutions (Feenstra and Hanson, 1996, 2001; Figini and Görg, 2011).

Weather shocks are likely to be a cause of income inequality, but there is little investigation of the impact of such shocks on income inequality in the existing literature. If weather shocks drive income inequality, there is a strong case for governments to rectify their impacts because this is inequality due to different circumstances rather than different effort levels (Roemer, 1998). Governments need to know if weather shocks contribute to increasing income inequality and how to identify the affected communities to intervene effectively. To my knowledge, the only study that attempts to establish a link between weather shocks and income inequality is Marx (2018). However, Marx only shows that the impact of local temperatures on income varies by income deciles, and does not draw any conclusions about whether global warming increases income inequality. In this chapter, I study the impact of weather shocks on different layers of household welfare, including income and consumption, and examine how it might exacerbate income inequality, which in turn has implications for households' happiness and fulfillment.

Although many studies show that natural disasters (including weather shocks) reduce household income and consumption and increase poverty (Nguyen et al., 2020b), the focus has been on the poor versus the non-poor. Poor households suffer more from shocks than wealthier households because their livelihoods are highly dependent on natural conditions, and their stocks of assets (which are often small) are more vulnerable (Tran, 2015). Furthermore, wealthier households can sell assets to smooth consumption, whereas poorer households may have to suffer falls in consumption when facing weather shocks (Hoddinott, 2006). When the poor experience economic stress, they eat less or take their children out of school (Banerjee and Duflo, 2007). In regions where marriage payments are customary and female children are considered tradeable assets, households might cope with temporary aggregate income shocks by marrying off their daughters earlier (Corno et al., 2020). Overall, it appears that weather shocks widen the gap between the poor and the wealthy, leading to a reduction in life satisfaction.

Some might argue that people could protect themselves from the financial impacts of weather shocks by buying insurance or hedging using weather derivatives. However, such a trading market for weather derivatives does not exist in many emerging economies; for instance, Vietnam lacks such a market (Tran and Otake, 2020). Furthermore, even when an insurance market exists, the cost might be too high for poor households, which may source their income from growing crops on small family plots; thus, take-up rates are low (King and Singh, 2020). In addition, households in rural areas might not be capable of understanding and participating in such complex financial contracts.

Given the above background, in this chapter I seek to answer the following four research questions (presented in Chapter 1, but repeated here for the sake of clarity).

1. What are the impacts of weather shocks on the income sources of rural households? This is of interest because different households might have different strategies to diversify their income sources. Poor households might not have access to credit and thus, struggle to escape the poverty trap.

2. How do weather shocks affect households' consumption? This is relevant to inequality because households which lack assets to offset lost income may be forced to lower their consumption. Then, the question is, what do they stop consuming – entertainment, education, or medical treatment? If they have to take their children out of school or marry off their daughters early (depending on the culture), then weather shocks have an indirect but significant negative impact on their children's lives and intergenerational mobility.

3. How do the impacts of weather shocks vary for households with different characteristics? Specifically, I will examine households that derive their income predominantly from agriculture, and also examine characteristics including crop area, ethnicity, and household size. As this indicates, my focus is on households in rural villages. If all household incomes are altered by the same proportion, then relative inequality would stay the same even though absolute inequality might rise sharply. How can we capture both relative and absolute measures of inequality?

4. How do weather shocks exacerbate income inequality? This final research question recognises that weather shocks are likely to have doubly negative impacts on household welfare, both reducing income and increasing income inequality, making some households worse off.

In investigating these research questions, I use Vietnam's economy as a case study. Vietnam makes a useful case study because it is quite vulnerable to weather shocks. Although Vietnam has been experiencing rapid economic change,[1] more than half of its workforce remains employed in agriculture (Quyen, 2019), which is increasingly affected by extreme weather conditions, such as storms, floods and droughts. Vietnam experiences various types of natural disasters due to its location in a tropical monsoon region and its diverse and complex topography. As shown in Figure 2.1.1, which was constructed using Emergency Events Database (EM-DAT) data,[2]

[1] Vietnam has transformed from one of the world's poorest nations into a low middle-income country over the last 30 years.

[2] EM-DAT is an international disaster database developed and maintained by the Centre for Research on the Epidemiology of Disasters, Department of Public Health, Université Catholique de Louvain (Brussels, Belgium).

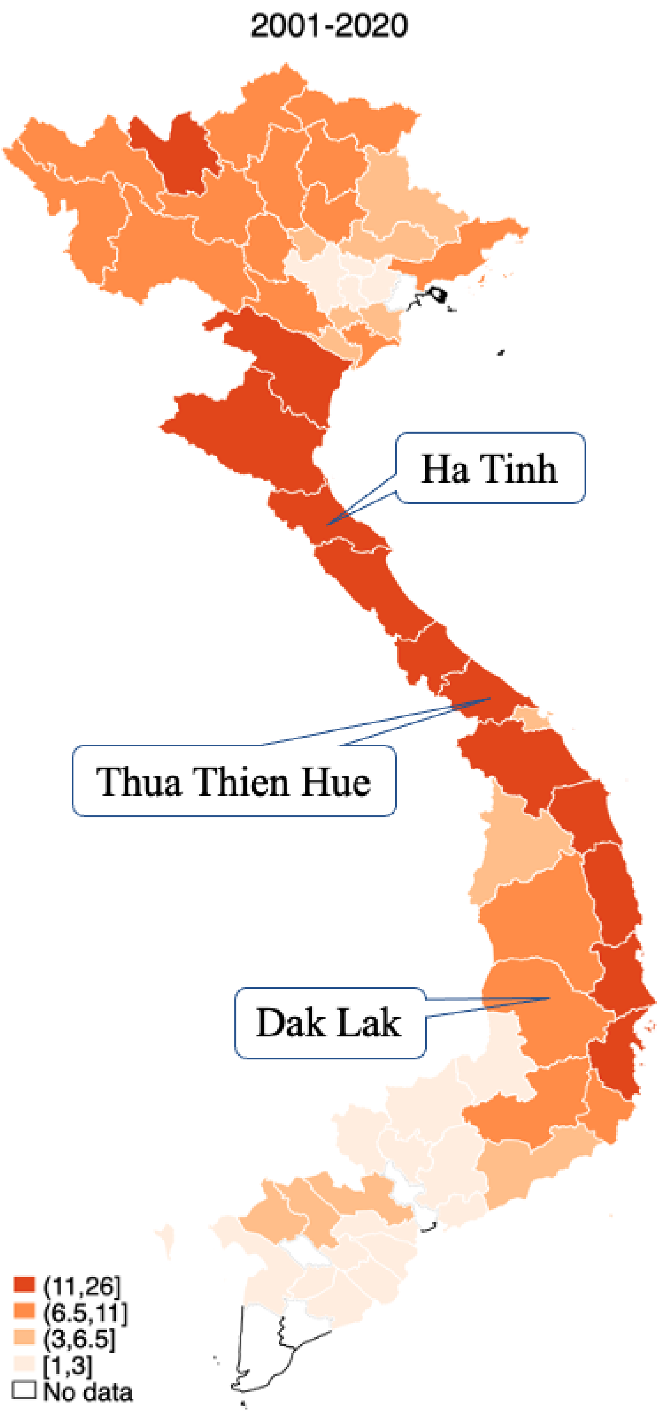
much of the country suffers from natural disasters, particularly floods and storms, with the three studied areas experiencing a high number of natural disasters over the 20-year period 2001-20.

As economic growth continues, average income rises, and the social impact of inequality is increasingly understood, there is increasing interest in analysing why some groups are falling behind. Although Vietnam has witnessed a significant reduction in poverty over the last few decades,[3] without a sustainable source of income, many families are likely to fall back into poverty following shocks. Most of Vietnam's remaining poor tend to be ethnic minorities living in mountainous areas. Poor households are often characterised by large household size, low education, a lack of supporting infrastructure, and high dependency on agriculture (Quyen, 2019). Because this group is at greater risk from weather shocks than other households, a focus of this study is understanding how weather shocks might affect farmers, ethnic minority groups, and large households differently.

My main contributions in conducting this research are to propose a new measure of weather shocks and to establish a link between weather shocks and income inequality. First, I recognise the limitations of existing measures of weather shocks, which tend to be dependent on the length of time over which weather samples are observed, and propose an absolute measure to overcome this limitation. This measure is defined as the total number of days with rainfall of at least 100 mm where there is also at least two such days in a row in the period that coincides with the household survey. Apart from being an absolute measure that does not depend on the period of weather data observations, this measure also provides a more accurate indication of weather shocks for short data series. Furthermore, weather shocks are likely to destroy crops and reduce the income of rural households, with the reduction varying between households with different characteristics; therefore, such shocks may exacerbate income inequality in locations where an increase in crop income would otherwise contribute to reducing income inequality. The results of this study will help governments and NGOs identify at-risk communities and design more effective and equitable assistance programs.

[3]The earliest survey conducted in Vietnam in 1992 indicated that about 64% of the population was considered poor, as measured by their income being below the international poverty line of $1.25 per day. Twenty years later, less than 3% of the population was considered poor by the same standard (Vu, 2015).

Figure 2.1.1: Number of natural disasters in provinces of Vietnam,2001-20



Disaster Type	
Flood	69
Storm	67
Drought	4
Epidemic	6
Landslide	2
Wildfire	1
Total	149

Note: EM-DAT defines a disaster as “a situation or event which overwhelms local capacity, necessitating a request to the national or international level for external assistance, or is recognized as such by a multilateral agency or by at least two sources, such as national, regional, or international assistance groups and the media”. Accordingly, an event is considered a disaster if there are (i) 10 or more people reported killed and/or (ii) 100 or more people reported affected and/or (iii) calls for international assistance/declaration of a state of emergency.

2.2 Data

2.2.1 Household Data

This study uses data collected in two research projects funded by the German Research Foundation (DFG) (Nguyen et al., 2020a), entitled *Impact of shocks on the vulnerability to poverty: Consequences for development of emerging South-East Asian Economies* (DFG FOR 756/1) and *Poverty dynamics and sustainable development: A long-term panel project in Thailand and Vietnam* (DFG FOR 756/2) (Do et al., 2021).[4] The researchers have collected data from rural areas of Thailand and Vietnam since 2007, with the purpose of examining and comparing the economic dynamics and vulnerability of rural households to poverty in these countries. The Thailand Vietnam Socio Economic Panel (TVSEP) panel dataset covers many important aspects of rural households' lives, including demography, income, expenditure and shocks experienced. The average attrition rate across the panel is below 5% (Parvathi et al. 2019, as cited in Do et al. (2021)). In this study, I focus on three rural provinces of Vietnam, Ha Tinh, Thua Thien Hue (Hue), and Dak Lak, shown in Figure 2.1.1. These provinces were chosen for their high incidence of poverty and high dependence on agriculture, which is increasingly being affected by extreme weather events.[5] As can be seen from Figure 2.1.1, these three provinces suffer frequent natural disasters, with from 11 to 26 natural disasters for Ha Tinh and Thua Thien Hue, and 6.5 to 11 for Dak Lak in the period between 2001 and 2020. Although Dak Lak does not experience weather shocks of the severity of those in Ha Tinh and Thua Thien Hue, observations from Dak Lak are included in the analysis because some variation in the treatment helps us estimate the coefficients of interest more accurately.

I include around 6,000 households in the main specification, which covers three years (2008, 2010, and 2013) because of the short span of weather data and the fact that only Thua Thien Hue was surveyed in 2011. The year here refers to the year when the household survey was completed; they often begin in May of the previous year and finish in April of that year.[6] The panel includes a section for village heads in some years. However, because the village head survey was not implemented in 2008, I do not include village-level variables as controls. The numbers of households interviewed in each year in each province are similar, as indicated in Table 2.2.1.

Most of the 6,000 households in the study are dependent on agriculture, with 67% of household heads (about 30% of all household members) reporting agricultural activities as their main occupation. The kernel densities of the log of total income and crop revenue per capita for

[4]For more information, see `https://www.tvsep.de/`. Household and village questionnaires can be downloaded free of charge from this page.

[5]For details about the sampling procedure, please refer to Do et al. (2021)

[6]Some sections of the survey (e.g., self-reported shocks) cover the period from the last survey until the current survey.

Table 2.2.1: Estimated populations of the three studied Vietnamese provinces in 2020 and numbers of observations by year

Province	Population (million)	2008	2010	2013
Ha Tinh	1.5	713	701	659
Thua Thien Hue	1.3	696	683	648
Dak Lak	2.1	735	715	703
Total observations		2144	2099	2010

the three provinces are approximately normal, as shown in Figure 2.2.1. Table 2.2.2 shows descriptive statistics for the three provinces in 2013. Similar statistics for 2008 and 2010 can be found in Appendix A.1.

Welfare measures such as income and consumption are recorded at the household level in TVSEP dataset; most studies using this data set to study the impact of natural disasters on household welfare use the household-level data, controlling for household heads' characteristics. I also do this, and the results are presented in Appendix A.2. However, household heads might change over time and thus, variables such as gender or ethnicity are not time-invariant for household heads. Presenting a fixed-effect model from which supposedly time-invariant variables like gender or ethnicity do not drop out may appear peculiar. Therefore, I use the individual-level data to investigate the impact of weather shocks on household welfare measures such as income or consumption.

Descriptive statistics of the variables of interest are presented in Table 2.2.2, and tests for the significance of differences between selected statistics for the three provinces are presented in Table 2.2.3. The variables are divided into four groups: demographic variables, income variables, consumption variables and household asset variables. With regard to demographic variables, the average household size in all three provinces is between five and six members. With ethnicity as a dummy variable that equals 1 if the individual belongs to the ethnic majority (Kinh), we observe that most of Ha Tinh's population are from an ethnic majority, as is the case in Thua Thien Hue (76%). The population of Dak Lak is more heterogeneous, with many indigenous ethnic minorities, which account for about 40% of the province's population. Around two thirds of the population in these three provinces has not been educated beyond primary school level. Many individuals (mostly men) are members of socio-political organisations; Ha Tinh has the highest percentage of adult women who are members of socio-political organisations, at 6%.[7]

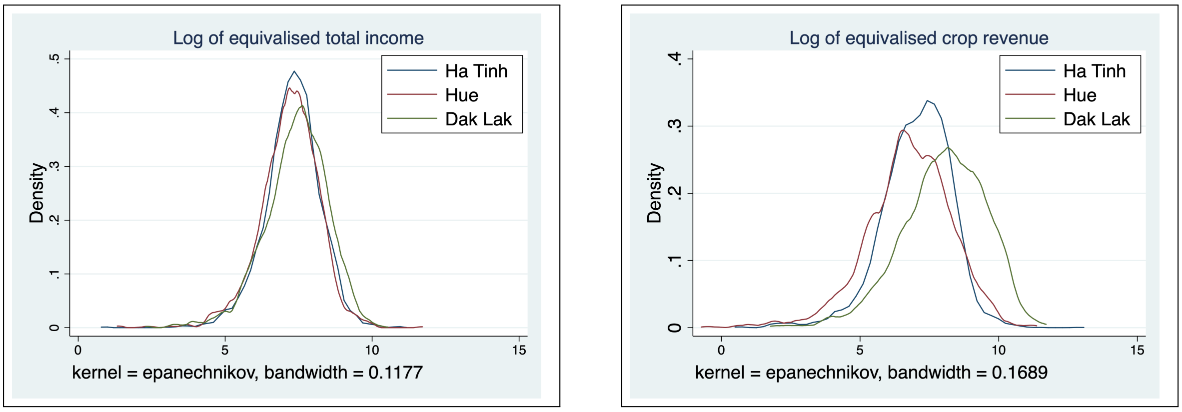
[7]It would be interesting to know whether having more women as members of socio-political organisations means those women receive more support from such organisations and thus, cope better with weather shocks (in terms of not having to reduce consumption). However, the low percentage of women as members of such organisations casts doubt on the meaningfulness of such an analysis.

To calculate the equivalised income, I use the OECD-modified scale,[8] following the formula

$$\text{Equivalised income} = \frac{\text{Total household income}}{1 \times \text{first adult} + 0.5 \times \text{additional adults} + 0.3 \times \text{children}}.$$

With regard to consumption, I follow the convention and use consumption per capita. The gap between equivalised total income and per capita total consumption is mostly due to saving. All monetary values are converted to 2005 PPP USD using the appropriate consumer price index (CPI) ratios and PPP factors, as noted in the TVSEP documentation.

Figure 2.2.1: Kernel density of log of total income and crop revenue per capita



The main sources of income for individuals in these three provinces are crops; livestock; hunting, collecting, and logging in the forests; off-farm employment activities; and remittances from family members and friends. Of the three sources of income that are most likely to be affected by weather shocks (crop, livestock and hunting income), each province has one that dominates. Crop income contributes more to total household income in Dak Lak in both relative and absolute terms in all three years. Livestock yields more income to households in Ha Tinh, and hunting in Hue in all three years. Compared to statistics of previous years (see Appendix A.1), remittances from family members have been increasing consistently in Ha Tinh and Hue, but not Dak Lak. It is possible that household members in the former two provinces have pursued migration as a deliberate strategy in response to weather shocks, as Imbert et al. (2018) suggest. It would be interesting to determine whether receiving remittances helps households to smooth consumption when facing adverse weather conditions.

Regarding consumption, it appears that individuals in Dak Lak spends more on main consumption categories such as food, nonfood items and education than individuals in Ha Tinh and Hue. They also have larger crop areas and more phones per household.

2.2.2 Weather Data

The rainfall data used in this study come from the Tropical Rainfall Measuring Mission (TRMM), a joint space mission between the National Aeronautics and Space Administration (NASA) and

[8] https://www.oecd.org/els/soc/OECD-Note-EquivalenceScales.pdf

Table 2.2.2: Descriptive statistics for the three studied Vietnamese provinces, 2013 (all values are annual means)

	All	Ha Tinh	Hue	Dak Lak
Demographic variables				
Household size	6.4180	5.8494	6.6881	6.6408
Number of children	1.1955	1.0079	1.1865	1.3597
Gender (Female=1)	0.5108	0.5203	0.5134	0.5004
Age	32.6041	35.2805	32.8209	30.1824
Ethnicity (Kinh=1)	0.7690	0.9988	0.7573	0.5889
Marital status (Married=1)	0.3659	0.3754	0.3544	0.3687
Education: Primary school	0.6277	0.5125	0.6931	0.6630
Education: Secondary school	0.2189	0.2888	0.1773	0.1991
Education: High school	0.1016	0.1256	0.0887	0.0937
Education: Bachelor or higher	0.0518	0.0731	0.0409	0.0442
Farmer	0.2930	0.2909	0.2288	0.3541
Member of socio-political organisation	0.5177	0.7505	0.4388	0.4043
Percentage of socio-political women	0.0412	0.0600	0.0377	0.0287
Income variables (2005 PPP USD)				
Equivalised total income	2277.5437	2115.6071	2492.9875	2212.9235
Equivalised crop income	410.3031	228.9850	201.2899	754.3516
Equivalised crop revenue	3822.1095	2384.7374	2466.1898	6271.1244
Equivalised livestock income	265.5030	417.6850	198.8936	200.5851
Equivalised hunting income	149.1123	16.1386	393.0601	34.0510
Equivalised off-farm employment income	588.8645	543.0971	671.7589	550.2505
Equivalised remittance: family/friends	314.2706	492.9429	338.9819	142.8717
Consumption variables (2005 PPP USD)				
Per capita total consumption	1080.3641	979.6208	1061.8917	1181.2042
Per capita food consumption	539.4329	466.8852	551.9569	588.1650
Per capita nonfood consumption	403.0316	372.1514	390.9424	439.8855
Per capita education consumption	72.6053	68.8062	65.6077	82.2354
Per capita health consumption	44.8644	50.2809	35.6335	48.8982
Per capita rent	20.9330	21.8850	18.9017	22.0201
Household asset variables				
Household crop area ($1000m^2$)	0.7386	0.3516	0.8458	0.9763
Number of tractors	1.1198	1.1201	1.1769	1.1047
Number of vehicles	1.6830	1.4361	1.8278	1.7402
Number of phones	2.5151	2.1863	2.7413	2.5856

Table 2.2.3: Socioeconomic differences between the three studied Vietnamese provinces, 2013

	Ha Tinh - rest	Hue - rest	Dak Lak - rest
Demographic variables			
Household size	-0.8141***	0.4066***	0.3497***
Number of children	-0.2686***	-0.0135	0.2577***
Gender (Female=1)	0.0137	0.0040	-0.0163*
Age	3.8357***	0.3254	-3.8071***
Ethnicity (Kinh=1)	0.3294***	-0.0176**	-0.2832***
Marital status (Married=1)	0.0136	-0.0174*	0.0044
Education: Primary school	-0.1650***	0.0985***	0.0554***
Education: Secondary school	0.1002***	-0.0625***	-0.0310***
Education: High school	0.0343***	-0.0195***	-0.0125**
Education: Bachelor or higher	0.0305***	-0.0164***	-0.0120***
Farmer	-0.0030	-0.0966***	0.0959***
Member of socio-political organisation	0.3295***	-0.1200***	-0.1785***
Percentage of socio-political women	0.0270***	-0.0052***	-0.0196***
Income variables (2005 PPP USD)			
Equivalised total income	-231.8806***	324.2421***	-101.4152
Equivalised crop income	-259.6334***	-314.5641***	539.9511***
Equivalised crop revenue	-2058.2047***	-2040.6548***	3843.4935***
Equivalised livestock income	217.9126***	-100.2470***	-101.8825***
Equivalised hunting income	-190.4079***	367.1408***	-180.5777***
Equivalised off-farm employment income	-65.5353***	124.7558***	-60.6010***
Equivalised remittance: family/friends	255.8448***	37.1905**	-268.9940***
Consumption variables (2005 PPP USD)			
Per capita total consumption	-144.2566***	-27.8008	158.2589***
Per capita food consumption	-103.8827***	18.8486***	76.4803***
Per capita nonfood consumption	-44.2180***	-18.1943	57.8386***
Per capita education consumption	-5.4399**	-10.5314***	15.1136***
Per capita health consumption	7.7560***	-13.8924***	6.3308**
Per capita rent	1.3632***	-3.0571***	1.7062***
Household asset variables			
Household crop area ($1000m^2$)	-0.5684***	0.1518***	0.3879***
Number of tractors	0.0005	0.0635***	-0.0249*
Number of vehicles	-0.3433***	0.2135***	0.0949***
Number of phones	-0.4726***	0.3364***	0.1117***

Note: The symbols ***, ** and * denote p<0.01, p<0.05 and p<0.10, respectively.

the Japan Aerospace Exploration Agency. TRMM relies on a satellite designed to measure the interactions between water vapor, clouds and precipitation, which are central to regulating the Earth's climate. TRMM was in operation between 1997 and 2015. It officially ended on April 15, 2015 after the spacecraft depleted its fuel reserves.[9] This data source provides 642,840 observations of daily rainfall, covering 220 villages in the three provinces of interest from 2007 to 2014. However, to construct the weather shocks variable so that its time span matches that of the household surveys, I use only data from May of the previous year to April in each year of interest. For this reason, I do not have enough weather data to construct weather shocks for the year 2007 (because it would require data from May 2006 to April 2007), or the lags of weather shocks, and thus focus my analysis on the years 2008, 2010 and 2013.

2.3 Data Analysis

2.3.1 Identifying Dependent and Independent Variables

When analysing the impacts of weather shocks on rural households' welfare, the dependent variable is often household consumption or household income (Nguyen et al., 2020b). Although consumption might be more closely related to household wellbeing, it tends to be smoothed over time. Because the choice of wellbeing measure matters empirically, I use both income and consumption as welfare indicators to examine the impacts of weather shocks (Decanq and Neumann, 2016). This allows me to cross-check the results for errors in measuring household income and consumption.

In the Vietnamese households under study, household income includes farm income, off-farm income, and remittances. Farm income comes mainly from crops, livestock and hunting, whereas off-farm income is derived from employment and self-employment. Household income could be negative as a result of income losses from investing in crops, livestock, or nonfarm self-employment. The numbers of households with income less than or equal to zero across provinces are shown in Table 2.3.1. Because the households with zero or negative incomes are small in number and appear to be randomly distributed across provinces, I drop such values and transform the remaining values into the logarithmic form to symmetrise the residuals and reduce potential outliers, as shown in Figure 2.2.1.

As Table 2.2.2 shows, household consumption includes food and nonfood items, education, health services and rent, with the largest share spent on food. I analyse the log values of consumption for the same reason as for income. As explained in Section 2.2.1, the dependent variables include equivalised income and per capita consumption for all individuals and for farmers specifically.

[9] For more information, see https://gpm.nasa.gov/missions/trmm

Table 2.3.1: Households with nonpositive income by province and year

Province	2008	2010	2013	Total
Ha Tinh	6	12	20	38
Thua Thien Hue	10	2	13	25
Dak Lak	19	43	45	107
Total	35	77	78	170

Farmers are classified as such because their main occupation is growing crops, fishing, hunting or collecting, or because they are permanently employed in agriculture. I did not use the proportion of income derived from agricultural activities to distinguish farmers from nonfarmers, because this percentage is likely to be low when households face weather shocks. Therefore, the share of income derived from agriculture activities does not accurately reflect the relative importance of farming income and other income sources. Based on this classification, nonfarming households may still engage in agricultural activities, but the income generated from these activities is not their main income source.

The identification of explanatory variables for regression models is based on the sustainable livelihoods framework, in which a livelihood is defined as the capabilities, assets and activities of a means of living (Nguyen et al., 2020b). Control variables include demographic variables such as household size, number of children, age, education level, whether an individual is a member of a socio-political organisation, and household assets.

2.3.2 Constructing Measured Weather Shocks

The current literature measures the exposure to weather shocks using two approaches. The first approach, "self-reported weather shocks", involves directly asking households to report whether they have experienced weather shocks. Under the second approach, "measured weather shocks", the times and places at which the shocks occurred are traced using weather data, which are matched with the locations of the surveyed households.

The limitations of using self-reported data to study the impacts of extreme weather events have been well documented. There is evidence that perception of shocks is endogenous to a household. People experience similar shocks differently and tend to adapt to their average environment Guiteras et al. (2015). Households affected by the same weather shocks might report the shocks differently depending on their level of engagement in agricultural activities and their perceived resilience, which depends on their application of coping strategies (Nguyen and Nguyen, 2020; Lohmann and Lechtenfeld, 2015). The TVSEP asks households if they had experienced any major shocks since the last survey. In the first survey, in 2007, households were asked to recall events over the previous five years, which is likely to result in inaccuracies.

Researchers construct weather shocks using the measured weather shocks approach in several ways. The first way is to define a "rain shocks" variable as equal to one if the annual rainfall is above the 80^{th} percentile for the district, as zero if it is between the 80^{th} and 20^{th} percentiles, and −1 if it is below the 20^{th} percentile (Jayachandran, 2006; Shah and Steinberg, 2017; Kaur, 2019). This measure is appropriate for India, where more rain often benefits crops. However, Vietnam often has too much rain. Using only one categorical variable to identify a rain shock becomes problematic when interpreting the results in different contexts. For instance, a negative coefficient for the rain shock variable would mean more rain reduces income even if the country is currently experiencing drought. Conversely, a positive coefficient would mean that a hurricane that destroys all crops would increase income if the country is currently experiencing normal weather. To rectify these issues, I create two dummies for weather shocks: one for too much rain (flood) and one for too little rain (drought). This construction shows that drought is almost 12 times more frequent than flood, which is inconsistent with the EM-DAT data used in constructing Figure 2.1.1. EM-DAT data record no droughts for Ha Tinh and Hue during 2001-2020, and only one drought in Dak Lak – in 2015, which is outside my period of analysis. Because most of the natural disasters experienced by the three provinces from 2007 to 2014 are storms and floods, I focus on identifying incidents of too much rain as weather shocks for the remainder of the chapter.[10]

The second way of identifying weather shocks in the existing literature is to count the number of times that monthly rainfall is three standard deviations away from the mean (see, e.g., Nguyen and Nguyen, 2020). The drawback of this approach is its dependence on the duration of the sample, meaning that the identification of shocks could vary from sample to sample. A month could be counted as including shocks using one sample, but be considered a normal month using a different sample. Another way to identify weather shocks is to use the deviation of yearly rainfall from the norm of the location, which is calculated as the natural log of a year's rainfall minus the natural log of mean annual rainfall in a given village (Maccini and Yang, 2009). Again, this measure is dependent on the mean, which varies depending on the duration of the sample observed. Therefore, I sought to identify an absolute measure, defined by the absolute amount of rainfall, that conveys the level of severity of the rainfall events.

Initially, I considered a weather shock as at least two consecutive days with rainfall of more than 100 mm.[11] However, this measure has two issues. First, if there are two days of continuous rainfall above the cutoff, followed by one day below the cutoff, then another two days above the

[10] Although Quiñones et al. (2021) use TVSEP and show significant impacts of drought, their paper combines provinces of Vietnam and Thailand and covers a longer period from 2007 to 2017. They also state that exceptional dry spells fell occurred in 2007, 2013 and 2016, which mostly fall outside my period of analysis.

[11] This cutoff was used based on information from the official website of the Thua Thien Hue meteorology department.

cutoff, using this measure would result in counting two flood incidents, whereas it is actually one. Second, the measure does not reflect the level of severity accurately. For example, it cannot distinguish between a three-day rainy incident and a 100-day rainy incident.

To resolve these issues, I construct the weather shocks variable as the total number of days with rainfall of at least 100 mm where there was also at least two such days in a row in the period that coincides with the household survey. Table 2.3.2 presents the frequency of this constructed weather shocks variable across the three provinces in three years. As mentioned in Section 2.2.1, the year in this table refers to the year in which the household survey was undertaken, which often covers the period from the previous May through to April of the reference year. As can be seen from Table 2.3.2, of the three provinces, Hue experienced the most severe rain shocks.

Table 2.3.2: The frequency of the weather shocks variable

	2008			2010			2013		
	Ha Tinh (713)	Hue (696)	Dak Lak (735)	Ha Tinh (701)	Hue (683)	Dak Lak (715)	Ha Tinh (659)	Hue (648)	Dak Lak (703)
0	175	10	735	334	39	656	39	477	703
2	0	370	0	367	46	59	620	171	0
3	538	0	0	0	0	0	0	0	0
4	0	0	0	0	347	0	0	0	0
5	0	153	0	0	0	0	0	0	0
6	0	163	0	0	0	0	0	0	0
10	0	0	0	0	251	0	0	0	0

Note: The number in the far left column is the total number of days across the year where there was both at least 100 mm of rain and at least two such days in a row. Thus, a frequency of 6 could indicate three instances of two continuous days with rainfall of at least 100 mm, or two instances of three such days continuously, or one instance of two continuous days together with another instance of four continuous days. The numbers in the remaining columns indicate the number of households in the province experiencing these adverse rain events. The total number of households in each province in each survey is given in parentheses under each province name to provide an understanding of how widespread the shocks are.

2.3.3 Model Specification

I use econometric regressions to estimate the effects of independent variables, including weather shocks, on dependent variables (welfare outcome variables). Following the livelihood framework, the basic form of the econometric model is:

$$Y = f(S, H), \tag{2.1}$$

where Y denotes the outcome (dependent) variables, S represents the shocks that the individual faced, and H is a vector representing the individual characteristics and household assets as controls.

The advantage of the model is that weather shock can be treated as an exogenous variable and thus, its causal interpretation is clear. However, several econometric challenges need to be taken into account. First, because I use panel data, either fixed effects or random effects regressions can be chosen. Hausman tests indicated that a fixed effects regression is the appropriate specification. Furthermore, to control for econometric heteroscedasticity, the standard errors are clustered at the village level. Thus, the model is further specified as follows:

$$y_{ivt} = \sigma_i + \gamma_t + \mathbf{h}'_{ivt}\boldsymbol{\beta} + s_{vt}\theta + \epsilon_{ivt}, \tag{2.2}$$

where y_{ivt} is a welfare measure of individual i in village v in year t; σ_i is individual fixed effects; γ_t is the year fixed effects; $\mathbf{h}_{ivt}$ is a vector capturing individual characteristics and household assets; s_{vt} is the weather shock faced by village v in year t; $\boldsymbol{\beta}$ and θ are the corresponding coefficients; and ϵ_{ivt} is the error term.

Although it is possible that severe weather shocks might have impacts on households lasting for several years, I do not include lag of weather shocks in the model because I do not have household-level data for every single year. Sometimes there are two or three-year gaps between the surveys, as explained in Section 2.3.1.

2.3.4 Effect of Crop Income on Household Income Inequality

As Table 2.2.2 shows, crops are one of the most important income sources of rural households. As a result, changes in crop income lead to changes in total income and income inequality among households. Many measures have been used to quantify income inequality, including the Gini coefficient, Atkinson index and Theil index. In this chapter, I use Gini coefficient, one of the most well-known inequality measures to capture both relative and absolute inequality (Kakwani and Son, 2016). To determine the contribution of crop income to overall income inequality in each province, I employ the Gini decomposition method proposed by Shorrocks (1982) and later extended by Lerman and Yitzhaki (1985).

In the Gini decomposition method, the Gini coefficient (G) is written as:

$$G = \sum_{k=1}^{K} S_k G_k R_k, \tag{2.3}$$

where S_k refers to the share of income source k, G_k is the Gini coefficients of income source k, and R_k is the Gini correlation of income source k with the distribution of total income.

There are different ways to calculate S_k in (2.3). For example, one could use total crop income divided by total income. In this chapter, I follow Stark et al. (1986) and compute

$S_k = \bar{Y}_k/\bar{Y}$, where $\bar{Y}_k$ is the mean of income source k and $\bar{Y}$ is the mean of total income. R_k =Cov$[Y_k, F(Y)]/$Cov$[Y_k, F(Y_k)]$, where $F(Y)$ and $F(Y_k)$ are the cumulative distributions of total income Y and of income source k (Y_k).

An important issue arises when calculating the Gini coefficient for distributions such as crop income that include negative values. The Gini coefficient is defined by:

$$G = \frac{S}{2(M-1)(T_a - T_n)}, \tag{2.4}$$

where M is the number of observations, S denotes the sum of absolute differences,

$$S = \sum_{i=1}^{M} \sum_{j=1}^{M} |Y_i - Y_j|,$$

T_a is the sum of positive values, and T_n is the sum of absolute negative values.

When dealing with negative values, the original Gini coefficient can be greater than one and is no longer a concentration index, making interpretation troublesome. I follow De Battisti et al. (2019) in computing the G_a by dropping all the negative values. When disaggregating overall income distributions into their sources, the number of negative values may no longer be considered a negligible phenomenon, and the G_p that is a normalised version of the original Gini should be computed. G_p can be evaluated from the original Gini coefficient using the following transformation:

$$G_p = G \cdot \left[\frac{T_a - T_n}{T_a + T_n}\right]. \tag{2.5}$$

Following Stark et al. (1986), the partial derivative of G with respect to a 1% change (e) in income source k is:

$$\frac{\partial G}{\partial e} = S_k(G_k R_k - G), \tag{2.6}$$

and therefore the marginal percentage change of income source k in income inequality is:

$$\frac{\partial G/\partial e}{G} = S_k\left(\frac{G_k R_k}{G} - 1\right). \tag{2.7}$$

There are three channels through which an income source could contribute to total income inequality, as shown in equation (2.7). First, if an income source accounts for a large share of total income (large S_k), it is likely to have a large effect on inequality. Second, if that income source is unequally distributed (large G_k), it may increase or decrease inequality, depending on where the households earning that income source are on the income distribution. Third, inequality could increase if the income source is unequally distributed and skewed toward those at the top of the income distribution (large positive R_k).

The fact that weather shocks amplify inequality can be illustrated by a simple model. Let X be the hypothetical income random variable in the absence of weather shocks. The realized income is:

$$Y = \begin{cases} X + c \text{ with probability } 1/2 \text{ if weather is good} \\ X - c \text{ with probability } 1/2 \text{ if weather is bad,} \end{cases}$$

where c is a positive scalar. It is easy to see that Var$Y$ = Var$X + c^2$, so that even shocks of neutral quality could amplify inequality. A natural question to ask is why Vietnamese farmers do not use hedging to lower c and reduce inequality. I will return to this question after showing that weather shocks reduce households' income in Section 2.4.3.

We note that if weather shocks reduce the income of all households by the same amount (proportion), they would have no impact on absolute (relative) inequality. Although these two cases almost never coincide, I will check if income loss is the same for all households. After confirming that weather shocks have differential impacts on households with different characteristics, I apply the Gini decomposition method to evaluate the impact of weather shocks on income inequality.

2.4 Results and Discussion

2.4.1 Impact of Weather Shocks on Household Welfare

As described in Section 2.3.1, the impacts of weather shocks on household welfare are measured by considering the effects on income and consumption. The impact of weather shocks varies depending on income sources, as indicated in Table 2.4.1. Annual equivalised income from hunting is most affected, with a decrease of approximately 5.53% on average if households experience one more day of rain with at least 100 mm, given that they have already experienced at least two continuous days of rain of similar magnitude. This result is consistent with Le (2020), who found that hunting and aquaculture income were most affected when a village is flooded. One reason for this large reduction in annual hunting income could be that households that hunt rely more on available natural resources.

If a household has already experienced at least two continuous days of rain with rainfall of at least 100 mm, then having one more day of such rainfall reduces equivalised remittances by about 3.63%, and average annual equivalised crop revenue by approximately 2.91%. It can be seen that the coefficient of crop income is of smaller magnitude compared with that of crop revenue, at 2.45% on average. We expect that this is because households adjust their input costs. For example, in response to a weather shock, they might reduce the amount of fertilisers used on crops or spend less time harvesting. Although the percentage reductions in crop revenue and crop income are smaller than those for remittances and hunting income, it is worth emphasising that crop income makes up a larger share of total income. Hence, a small percentage reduction could mean a larger value in absolute terms. On average, having an additional day of heavy rain after the first two days reduces the annual equivalised total income by about 0.65%.

Table 2.4.1: Impact of weather shocks on equivalised income (ln)

	Total income	Crop income	Crop revenue	Livestock	Hunting	Off Farm	Remittance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Rain shock	-0.0065	-0.0248	-0.0295***	-0.0008	-0.0569***	0.0175	-0.0369**
	(0.0070)	(0.0155)	(0.0107)	(0.0117)	(0.0145)	(0.0118)	(0.0162)
Household size	-0.1346***	-0.1155*	-0.1366***	-0.2072***	0.0075	-0.0898	0.0118
	(0.0360)	(0.0592)	(0.0329)	(0.0702)	(0.0503)	(0.0783)	(0.0778)
Number of children	0.0364	0.1141***	0.0482*	0.1483**	-0.0075	0.0819	-0.0298
	(0.0276)	(0.0433)	(0.0275)	(0.0636)	(0.0505)	(0.0528)	(0.0887)
Age	0.0006	-0.0053	0.0016	0.0143	0.0150**	0.0033	-0.0210
	(0.0040)	(0.0062)	(0.0043)	(0.0088)	(0.0062)	(0.0132)	(0.0159)
Marital status (Married=1)	0.0841**	-0.0478	0.0242	0.0587	0.2081***	-0.1644	0.1354
	(0.0393)	(0.0489)	(0.0369)	(0.0784)	(0.0713)	(0.1040)	(0.0900)
Education (ref: Primary school)							
Education: Secondary school	0.0662*	-0.0041	0.0443	0.0651	-0.0769	0.0232	0.0436
	(0.0393)	(0.0545)	(0.0432)	(0.0836)	(0.0616)	(0.0691)	(0.1062)
Education: High school	0.0523	-0.0591	0.0072	-0.0580	-0.1113	0.0653	-0.3237***
	(0.0385)	(0.0662)	(0.0518)	(0.0780)	(0.0759)	(0.0762)	(0.1097)
Education: Bachelor or higher	-0.0513	-0.1888**	-0.1288	-0.1247	-0.1404	-0.2145*	-0.1500
	(0.0526)	(0.0827)	(0.0785)	(0.1186)	(0.1855)	(0.1096)	(0.1467)
Member of socio-political organisation	0.0015	0.0008	0.0069	0.0070	0.0842**	0.0159	-0.2810***
	(0.0211)	(0.0383)	(0.0290)	(0.0578)	(0.0398)	(0.0478)	(0.0678)
Household crop area ($1000m^2$)	0.0232	0.0184	0.0676	-0.0679	0.0180*	0.0040	0.0125
	(0.0161)	(0.0161)	(0.0474)	(0.0574)	(0.0105)	(0.0074)	(0.0220)
Number of tractors	0.0933***	0.0388	0.1212***	0.1353**	-0.0297	0.0480	-0.2285**
	(0.0336)	(0.0505)	(0.0438)	(0.0660)	(0.0797)	(0.0835)	(0.0992)
Number of vehicles	0.1633***	0.0716	0.0503	0.1431**	-0.0236	0.1033**	-0.1071*
	(0.0257)	(0.0480)	(0.0368)	(0.0627)	(0.0548)	(0.0461)	(0.0632)
Number of phones	0.0439***	-0.0157	0.0282	0.0411	-0.0100	0.0425	0.0492
	(0.0141)	(0.0277)	(0.0175)	(0.0271)	(0.0329)	(0.0304)	(0.0441)
Constant	7.6402***	6.1974***	7.9656***	5.3981***	3.0306***	6.4220***	6.2521***
	(0.2451)	(0.4082)	(0.2566)	(0.5046)	(0.3325)	(0.5689)	(0.7369)
Observations	29994	24179	27507	20480	16831	17865	14262
R^2	0.0355	0.1135	0.4754	0.0208	0.0427	0.3323	0.0268
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors (in parentheses) are clustered at the village level. The symbols ***, ** and * denote p<0.01, p<0.05 and p<0.10, respectively.

The impact of weather shocks on individuals whose main occupations are related to agricultural activities can be expected to be larger. The larger absolute values of coefficients of the rain shocks on crop income, crop revenue, livestock, and hunting, as shown in specification (2) to (5) of Table 2.4.2, reaffirm this expectation. If a farmer has already experienced at least two continuous days of rain with rainfall of at least 100 mm, then having one more day of such rainfall reduces equivalised crop revenue by about 3.6% on average, an increase of 0.7% compared with the impact on an individual on average. The reduction of weather shocks on equivalised hunting income is also stronger for farmers, at 6.2%, than for a general individual, at 5.53%. However, the magnitude of weather impacts on remittance income is smaller for farmers compared with general households (and not statistically significant). This is plausible, because farmers may have fewer family members working away from home who send remittances. Rainfall shocks also reduce income from off-farm employment activities for farmers, possibly due to higher input prices or shortages of inputs, as Grabrucker and Grimm (2021) suggest.

Table 2.4.2: Impact of weather shocks on equivalised income (ln) for farmers

	Total income	Crop income	Crop revenue	Livestock	Hunting	Off Farm	Remittance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Rain shock	-0.0041	-0.0357**	-0.0367***	-0.0115	-0.0641***	-0.0043	-0.0281
	(0.0100)	(0.0156)	(0.0128)	(0.0140)	(0.0187)	(0.0157)	(0.0213)
Household size	-0.1220***	-0.1305*	-0.1501***	-0.2150***	-0.0087	-0.0328	0.0790
	(0.0458)	(0.0722)	(0.0347)	(0.0816)	(0.0549)	(0.0983)	(0.0839)
Number of children	0.0320	0.1534***	0.0884***	0.1806***	-0.0173	0.0512	0.0891
	(0.0331)	(0.0518)	(0.0304)	(0.0683)	(0.0536)	(0.0744)	(0.1131)
Age	0.0091	0.0057	0.0124	0.0171	0.0319	0.0132	-0.0343
	(0.0138)	(0.0197)	(0.0105)	(0.0278)	(0.0228)	(0.0278)	(0.0343)
Marital status (Married=1)	0.0552	-0.1310	-0.0262	0.2754	0.1409	-0.1317	-0.2172
	(0.0871)	(0.1095)	(0.0843)	(0.2324)	(0.1057)	(0.2218)	(0.2370)
Education (ref: Primary school)							
Education: Secondary school	0.0733	0.0176	0.0816	0.0886	-0.1476	-0.1365	-0.0853
	(0.0881)	(0.1073)	(0.0818)	(0.1924)	(0.1307)	(0.1750)	(0.1940)
Education: High school	0.1473	0.1589	0.1316	0.0006	-0.2521	0.0151	-0.6442**
	(0.1257)	(0.2242)	(0.1836)	(0.2488)	(0.2517)	(0.2531)	(0.2907)
Education: Bachelor or higher	0.1888	0.3457	0.0957	0.0604	0.8614***	-0.7494	-0.2185
	(0.1950)	(0.3477)	(0.2601)	(0.3924)	(0.2815)	(0.4680)	(0.5133)
Member of socio-political organisation	0.0580	0.0037	-0.0193	-0.0357	0.1705***	-0.0727	-0.2943**
	(0.0378)	(0.0528)	(0.0466)	(0.1092)	(0.0639)	(0.0907)	(0.1368)
Household crop area ($1000m^2$)	0.0243	0.0119	0.0462	-0.0554	0.0163*	0.0065	0.0021
	(0.0161)	(0.0114)	(0.0356)	(0.0701)	(0.0086)	(0.0079)	(0.0123)
Number of tractors	0.0813**	0.0097	0.0928*	0.1243*	-0.0663	-0.0032	-0.1558
	(0.0373)	(0.0526)	(0.0501)	(0.0694)	(0.0865)	(0.1059)	(0.1103)
Number of vehicles	0.1452***	0.1081**	0.0694*	0.0647	-0.0154	0.0655	-0.0691
	(0.0326)	(0.0511)	(0.0364)	(0.0643)	(0.0600)	(0.0625)	(0.0839)
Number of phones	0.0307	-0.0158	0.0309	0.0585*	-0.0055	0.0157	0.0705
	(0.0187)	(0.0296)	(0.0195)	(0.0315)	(0.0358)	(0.0489)	(0.0642)
Constant	7.0184***	5.9895***	7.6967***	5.2072***	2.8206***	4.3237***	6.3703***
	(0.6630)	(1.0019)	(0.4840)	(1.2091)	(0.8827)	(0.9789)	(1.5289)
Observations	10404	9184	10410	7603	6776	6034	4616
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors (in parentheses) are clustered at the village level. The symbols ***, ** and * denote p<0.01, p<0.05 and p<0.10, respectively.

Regarding the impact of lost income on household consumption, it appears that households employ some coping mechanisms to smooth consumption when facing adverse events, as indicated in Tables 2.4.3 and 2.4.4. Food, nonfood and education consumption are not reduced by statistically significant amounts when the households experience weather shocks. However, an additional day of heavy rain after at least two consecutive days reduces annual health consumption per capita by about 3.3% for an individual on average, and by about 3.63% for farmers. It may be that heavy rains result in rivers flooding and transportation disruptions, making it more difficult for people to obtain medical treatment, which may have long-term negative consequences. However, an additional day of heavy rain after at least two consecutive days results in an increase in annual rent consumption per capita by about 2.2%. This is probably due to the floods displacing people from their homes, forcing them to seek shelter elsewhere.

It is important to note that the impact of weather shocks is different for households with different characteristics, as indicated in Table 2.4.5. First, ethnic majorities appear to fare better than minority groups when facing weather shocks, possibly because the latter face greater disadvantages in accessing formal credit, which limits their ability to diversify income sources, as Nguyen et al. (2020a) suggest. Second, the impact of weather shocks on annual equivalised crop revenue is stronger for larger households and for farmers. Furthermore, Thua Thien Hue and Dak Lak are statistically more vulnerable to weather shocks than Ha Tinh. As discussed in Section 2.2.1, it is possible that Ha Tinh residents already employ strategies to diversify their income sources, such as migration of some family members, or diversifying away from crops to raising livestock.

As noted above, it has been observed that households employ mechanisms to smooth consumption when experiencing weather shocks. A natural question is why Vietnamese households do not employ some strategies to “smooth” income, such as buying insurance against bad weather or using weather derivatives. The first reason is that there are no active trading markets for weather derivatives in Vietnam or other developing countries (Tran and Otake, 2020). Second, agricultural practices in Vietnam, where farmers grow crops on small family-run paddy fields, could explain why Vietnamese farmers do not think investment in insurance is worthwhile. Vietnamese farmers tend to use private transfers as a substitute for agricultural insurance. The individual riskiness and lack of trust in insurers have been found to contribute to the low take-up rates (King and Singh, 2020).

2.4.2 Robustness Check

I use two alternative weather shock variables to check the validity of the newly constructed weather shock variables described in Section 2.3.2 and used to produce the results in Section 2.4.1.

Table 2.4.3: Impact of weather shocks on per capita consumption (ln)

	Total	Food	NonFood	Education	Health	Rent
	(1)	(2)	(3)	(4)	(5)	(6)
Rain shock	0.0003	0.0033	-0.0004	0.0090	-0.0330**	0.0244***
	(0.0030)	(0.0036)	(0.0039)	(0.0084)	(0.0135)	(0.0048)
Household size	-0.0876***	-0.0882***	-0.0804***	-0.2762***	-0.0670	-0.1425***
	(0.0148)	(0.0171)	(0.0183)	(0.0390)	(0.0508)	(0.0182)
Number of children	-0.0308**	-0.0003	-0.0614***	0.1219***	-0.0227	-0.0240
	(0.0154)	(0.0166)	(0.0189)	(0.0350)	(0.0458)	(0.0174)
Age	0.0016	0.0015	0.0008	0.0091	0.0011	0.0018
	(0.0020)	(0.0023)	(0.0026)	(0.0059)	(0.0079)	(0.0031)
Marital status (Married=1)	-0.0041	0.0171	-0.0019	-0.1138**	-0.0116	-0.0026
	(0.0185)	(0.0180)	(0.0259)	(0.0540)	(0.0592)	(0.0238)
Education (ref: Primary school)						
Education: Secondary school	-0.0185	-0.0101	0.0154	-0.2420***	-0.0304	0.0151
	(0.0179)	(0.0190)	(0.0217)	(0.0429)	(0.0651)	(0.0243)
Education: High school	-0.0173	0.0146	0.0377*	-0.3872***	0.0441	-0.0129
	(0.0160)	(0.0185)	(0.0205)	(0.0527)	(0.0662)	(0.0226)
Education: Bachelor or higher	-0.0590**	-0.0112	-0.0253	-0.4087***	0.0664	0.0261
	(0.0246)	(0.0274)	(0.0372)	(0.0898)	(0.0885)	(0.0434)
Member of socio-political organisation	0.0203**	0.0010	0.0497***	0.0307	-0.0559	-0.0084
	(0.0096)	(0.0098)	(0.0133)	(0.0269)	(0.0383)	(0.0158)
Household crop area ($1000m^2$)	0.0080	0.0079**	0.0076	0.0254	0.0450	-0.0204***
	(0.0055)	(0.0035)	(0.0083)	(0.0178)	(0.0478)	(0.0055)
Number of tractors	0.0754***	0.0720***	0.0809***	0.0518	0.0522	0.0846***
	(0.0174)	(0.0182)	(0.0251)	(0.0452)	(0.0588)	(0.0281)
Number of vehicles	0.1483***	0.0981***	0.2636***	0.0249	-0.0028	0.0233
	(0.0134)	(0.0145)	(0.0200)	(0.0345)	(0.0461)	(0.0200)
Number of phones	0.0402***	0.0250***	0.0633***	0.0574***	0.0403	0.0136
	(0.0070)	(0.0083)	(0.0088)	(0.0213)	(0.0282)	(0.0111)
Constant	7.0168***	6.4099***	5.6013***	5.4992***	3.0724***	3.5092***
	(0.1058)	(0.1199)	(0.1396)	(0.2973)	(0.4107)	(0.1561)
Observations	29986	29986	29986	22021	26588	29933
R^2	0.1577	0.0669	0.1861	0.1645	0.0116	0.0903
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors (in parentheses) are clustered at the village level. The symbols ***, ** and * denote p<0.01, p<0.05 and p<0.10, respectively.

Table 2.4.4: Impact of weather shocks on per capita consumption (ln) for farmers

	Total	Food	NonFood	Education	Health	Rent
	(1)	(2)	(3)	(4)	(5)	(6)
Rain shock	0.0001	0.0005	0.0042	0.0063	-0.0370**	0.0221***
	(0.0040)	(0.0046)	(0.0053)	(0.0120)	(0.0157)	(0.0067)
Household size	-0.0821***	-0.0802***	-0.1008***	-0.2511***	-0.0209	-0.1452***
	(0.0195)	(0.0230)	(0.0226)	(0.0470)	(0.0685)	(0.0212)
Number of children	-0.0333*	-0.0067	-0.0439*	0.1550***	-0.0283	-0.0350*
	(0.0194)	(0.0208)	(0.0227)	(0.0465)	(0.0493)	(0.0199)
Age	0.0100	0.0102	0.0127*	0.0027	0.0093	0.0003
	(0.0061)	(0.0082)	(0.0067)	(0.0170)	(0.0184)	(0.0070)
Marital status (Married=1)	0.0351	0.0489	0.0130	0.1433	0.0359	-0.0050
	(0.0389)	(0.0388)	(0.0469)	(0.1068)	(0.1152)	(0.0400)
Education (ref: Primary school)						
Education: Secondary school	0.0469	0.0824**	0.0042	-0.0104	-0.1435	-0.0499
	(0.0324)	(0.0339)	(0.0434)	(0.0775)	(0.1221)	(0.0418)
Education: High school	0.1289***	0.1593***	0.0899	-0.1838	0.2007	-0.1184*
	(0.0481)	(0.0486)	(0.0784)	(0.1387)	(0.1771)	(0.0688)
Education: Bachelor or higher	0.1481	0.1447	0.1134	-0.2357	0.7829**	-0.1564
	(0.1044)	(0.1089)	(0.1517)	(0.3354)	(0.3820)	(0.1383)
Member of socio-political organisation	0.0276*	0.0136	0.0762***	0.0610	-0.0607	0.0088
	(0.0163)	(0.0192)	(0.0229)	(0.0460)	(0.0687)	(0.0243)
Household crop area (1000m^2)	0.0092**	0.0137***	0.0065	0.0193	0.0831	-0.0200***
	(0.0044)	(0.0040)	(0.0056)	(0.0156)	(0.0617)	(0.0050)
Number of tractors	0.0783***	0.0648***	0.0840***	0.0876	0.0912	0.0652*
	(0.0200)	(0.0233)	(0.0293)	(0.0571)	(0.0690)	(0.0355)
Number of vehicles	0.1411***	0.0872***	0.2727***	0.0193	-0.0145	0.0235
	(0.0154)	(0.0164)	(0.0237)	(0.0465)	(0.0532)	(0.0228)
Number of phones	0.0517***	0.0342***	0.0772***	0.0598***	0.0706**	0.0129
	(0.0088)	(0.0107)	(0.0114)	(0.0226)	(0.0347)	(0.0140)
Constant	6.4311***	5.8347***	5.0056***	4.3014***	2.2730**	3.5198***
	(0.2610)	(0.3446)	(0.3121)	(0.6494)	(0.9204)	(0.3349)
Observations	10537	10537	10537	7073	9386	10513
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: Standard errors (in parentheses) are clustered at the village level. The symbols ***, ** and * denote p<0.01, p<0.05 and p<0.10, respectively.

Table 2.4.5: Heterogeneity tests

	Crop Revenue				
	(1)	(2)	(3)	(4)	(5)
Rain shock	-0.0353**	-0.0264**	0.0680***	0.0123	-0.1418***
	(0.0143)	(0.0108)	(0.0210)	(0.0288)	(0.0367)
Household crop area ($1000m^2$)	0.0645	0.0673	0.0663	0.0668	0.0626
	(0.0469)	(0.0474)	(0.0462)	(0.0473)	(0.0450)
Rain shock × Household crop area ($1000m^2$)	0.0138				
	(0.0190)				
Number of children	0.0465*	0.0487*	0.0518*	0.0511*	0.0537**
	(0.0280)	(0.0276)	(0.0274)	(0.0279)	(0.0272)
Marital status (Married=1)	0.0245	0.0210	0.0270	0.0254	0.0246
	(0.0369)	(0.0371)	(0.0368)	(0.0369)	(0.0373)
Education: Secondary school	0.0453	0.0376	0.0383	0.0442	0.0496
	(0.0434)	(0.0429)	(0.0431)	(0.0431)	(0.0432)
Education: High school	0.0090	-0.0004	0.0077	0.0055	0.0049
	(0.0519)	(0.0527)	(0.0519)	(0.0517)	(0.0513)
Education: Bachelor or higher	-0.1293	-0.1321*	-0.1225	-0.1292	-0.1162
	(0.0785)	(0.0786)	(0.0783)	(0.0785)	(0.0779)
Member of socio-political organisation	0.0072	0.0061	0.0121	0.0081	0.0074
	(0.0289)	(0.0289)	(0.0293)	(0.0290)	(0.0288)
Household crop area ($1000m^2$)	0.0000				
	(.)				
Number of tractors	0.1209***	0.1213***	0.1326***	0.1209***	0.1209***
	(0.0437)	(0.0438)	(0.0437)	(0.0439)	(0.0438)
Number of vehicles	0.0498	0.0505	0.0516	0.0492	0.0540
	(0.0369)	(0.0369)	(0.0368)	(0.0369)	(0.0372)
Number of phones	0.0282	0.0281	0.0294*	0.0273	0.0298*
	(0.0175)	(0.0175)	(0.0174)	(0.0173)	(0.0173)
Farmer=1 × Rain shock		-0.0103			
		(0.0083)			
Thua Thien Hue × Rain shock			-0.1131***		
			(0.0266)		
Dak Lak × Rain shock			-0.2391***		
			(0.0557)		
Rain shock × Household size				-0.0069	
				(0.0045)	
Ethnicity (Kinh=1)=1 × Rain shock					0.1233***
					(0.0373)
Observations	27507	27507	27507	27507	27507
Individual FE					
Year FE	Yes	Yes	Yes	Yes	Yes

Note: Standard errors (in parentheses) are clustered at the village level. The symbols ***, ** and * denote p<0.01, p<0.05 and p<0.10, respectively.

Three standard deviations away from the mean

The first alternative is to construct the weather shock variable as the number of times that monthly rainfall is three standard deviations away from the mean, as in Nguyen and Nguyen (2020). I refer to this variable as "Rain month"; the results of my analysis on the impact of "Rain month" on household welfare are presented in Tables 2.4.6 and 2.4.7.

Table 2.4.6: Impact of weather shocks on equivalised income (ln) – Rain month

	Total income	Crop income	Crop revenue	Livestock	Hunting	Off Farm	Remittance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Rain month	-0.1058***	0.0410	-0.1670***	-0.0297	-0.3710***	0.1585***	-0.3377***
	(0.0366)	(0.0546)	(0.0537)	(0.0658)	(0.0611)	(0.0598)	(0.0849)
Household size	-0.1318***	-0.1164**	-0.1346***	-0.2065***	0.0032	-0.0939	0.0142
	(0.0364)	(0.0588)	(0.0329)	(0.0703)	(0.0513)	(0.0769)	(0.0772)
Number of children	0.0367	0.1129**	0.0493*	0.1485**	0.0046	0.0811	-0.0317
	(0.0276)	(0.0434)	(0.0270)	(0.0636)	(0.0502)	(0.0522)	(0.0875)
Age	0.0000	-0.0043	0.0022	0.0142	0.0150**	0.0049	-0.0211
	(0.0040)	(0.0062)	(0.0042)	(0.0087)	(0.0059)	(0.0131)	(0.0155)
Marital status (Married=1)	0.0824**	-0.0486	0.0204	0.0584	0.1961***	-0.1554	0.1461
	(0.0395)	(0.0488)	(0.0371)	(0.0787)	(0.0683)	(0.1016)	(0.0901)
Education (ref: Primary school)							
Education: Secondary school	0.0698*	-0.0084	0.0486	0.0662	-0.0669	0.0143	0.0307
	(0.0393)	(0.0546)	(0.0433)	(0.0840)	(0.0615)	(0.0695)	(0.1039)
Education: High school	0.0576	-0.0638	0.0136	-0.0574	-0.0989	0.0526	-0.3122***
	(0.0382)	(0.0667)	(0.0512)	(0.0778)	(0.0785)	(0.0763)	(0.1072)
Education: Bachelor or higher	-0.0425	-0.2009**	-0.1297*	-0.1238	-0.1371	-0.2300**	-0.0661
	(0.0527)	(0.0837)	(0.0782)	(0.1186)	(0.1828)	(0.1097)	(0.1441)
Member of socio-political organisation	0.0072	-0.0044	0.0134	0.0092	0.0934**	0.0049	-0.2664***
	(0.0208)	(0.0380)	(0.0282)	(0.0570)	(0.0408)	(0.0476)	(0.0674)
Household crop area ($1000m^2$)	0.0207	0.0201	0.0646	-0.0698	0.0134	0.0066	0.0092
	(0.0147)	(0.0174)	(0.0460)	(0.0567)	(0.0117)	(0.0072)	(0.0203)
Number of tractors	0.0927***	0.0331	0.1167***	0.1350**	-0.0233	0.0467	-0.2260**
	(0.0332)	(0.0512)	(0.0437)	(0.0661)	(0.0783)	(0.0828)	(0.0968)
Number of vehicles	0.1623***	0.0756	0.0502	0.1426**	-0.0223	0.1095**	-0.0988
	(0.0257)	(0.0478)	(0.0366)	(0.0626)	(0.0551)	(0.0465)	(0.0622)
Number of phones	0.0417***	-0.0133	0.0279	0.0404	-0.0060	0.0477	0.0436
	(0.0138)	(0.0277)	(0.0169)	(0.0270)	(0.0322)	(0.0300)	(0.0430)
Constant	7.6409***	6.1457***	7.9121***	5.4004***	3.0342***	6.3955***	6.2143***
	(0.2451)	(0.4038)	(0.2535)	(0.5035)	(0.3268)	(0.5642)	(0.6408)
Observations	29994	24179	27507	20480	16831	17865	14262
R^2	0.0395	0.1121	0.4774	0.0209	0.0615	0.3349	0.0368
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: "Rain month" is measured by the number of times monthly rainfall is three standard deviation away from the mean. Standard errors (in parentheses) are clustered at the village level. The symbols ***, ** and * denote $p<0.01$, $p<0.05$ and $p<0.10$, respectively.

Table 2.4.7: Impact of weather shocks on per capita consumption (ln) – Rain month

	Total	Food	NonFood	Education	Health	Rent
	(1)	(2)	(3)	(4)	(5)	(6)
Rain month	-0.0475***	-0.0362**	-0.0597***	-0.0701*	-0.1582***	0.1313***
	(0.0145)	(0.0177)	(0.0191)	(0.0364)	(0.0493)	(0.0257)
Household size	-0.0871***	-0.0880***	-0.0797***	-0.2767***	-0.0619	-0.1456***
	(0.0146)	(0.0169)	(0.0181)	(0.0383)	(0.0499)	(0.0179)
Number of children	-0.0302*	0.0004	-0.0608***	0.1239***	-0.0270	-0.0236
	(0.0153)	(0.0166)	(0.0188)	(0.0349)	(0.0456)	(0.0171)
Age	0.0013	0.0010	0.0004	0.0076	0.0018	0.0015
	(0.0020)	(0.0022)	(0.0026)	(0.0057)	(0.0079)	(0.0031)
Marital status (Married=1)	-0.0048	0.0167	-0.0027	-0.1116**	-0.0120	-0.0002
	(0.0185)	(0.0180)	(0.0257)	(0.0549)	(0.0591)	(0.0235)
Education (ref: Primary school)						
Education: Secondary school	-0.0163	-0.0081	0.0180	-0.2388***	-0.0289	0.0121
	(0.0176)	(0.0188)	(0.0216)	(0.0430)	(0.0649)	(0.0244)
Education: High school	-0.0153	0.0161	0.0402*	-0.3845***	0.0494	-0.0191
	(0.0160)	(0.0183)	(0.0206)	(0.0522)	(0.0663)	(0.0225)
Education: Bachelor or higher	-0.0544**	-0.0068	-0.0197	-0.4027***	0.0688	0.0206
	(0.0245)	(0.0273)	(0.0369)	(0.0908)	(0.0885)	(0.0432)
Member of socio-political organisation	0.0232**	0.0037	0.0533***	0.0382	-0.0512	-0.0125
	(0.0095)	(0.0097)	(0.0131)	(0.0266)	(0.0389)	(0.0160)
Household crop area (1000m^2)	0.0065	0.0066**	0.0058	0.0230	0.0342	-0.0178***
	(0.0047)	(0.0032)	(0.0072)	(0.0162)	(0.0467)	(0.0057)
Number of tractors	0.0755***	0.0725***	0.0810***	0.0509	0.0514	0.0878***
	(0.0172)	(0.0181)	(0.0248)	(0.0446)	(0.0588)	(0.0281)
Number of vehicles	0.1474***	0.0973***	0.2626***	0.0228	-0.0034	0.0243
	(0.0133)	(0.0145)	(0.0199)	(0.0348)	(0.0461)	(0.0201)
Number of phones	0.0391***	0.0239***	0.0621***	0.0545**	0.0407	0.0137
	(0.0069)	(0.0083)	(0.0087)	(0.0214)	(0.0282)	(0.0111)
Constant	7.0267***	6.4267***	5.6113***	5.5571***	2.9956***	3.5568***
	(0.1042)	(0.1191)	(0.1384)	(0.2931)	(0.4066)	(0.1519)
Observations	29986	29986	29986	22021	26588	29933
R^2	0.1613	0.0685	0.1891	0.1656	0.0125	0.0959
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: "Rain month" is measured by the number of times monthly rainfall is three standard deviation away from the mean.
Standard errors (in parentheses) are clustered at the village level. The symbols ***, ** and * denote p<0.01, p<0.05 and p<0.10, respectively.

The results using the number of times that monthly rainfall is three standard deviations away from the mean tell the same story as the main results presented in Tables 2.4.1 and 2.4.3. The impacts of "Rain month" on annual equivalised income have the same sign and are of the same relative magnitude as the main results, with the percentage of hunting income most reduced, followed by remittance income and crop revenue. Spending on health is the type of consumption most reduced when households experience a month of heavy rain. In addition, an additional month with rainfall three standard deviations away from the mean would result in an increase in rent spending. These results are all consistent with the main results presented in Tables 2.4.3 and 2.4.4.

It should be noticed here that the absolute magnitudes of the impact using the Rain month variable are larger than those using the weather shocks variable constructed by counting the number of heavy rain days. This is plausible, because a month of heavy rain is likely to be more disastrous than a few days of heavy rain.

Rainfall deviation

Another way to construct the weather shock variable is to use "Rainfall deviation", defined by the natural log of the year's rainfall minus the natural log of mean annual rainfall in a given village, as used in Maccini and Yang (2009). The results presented in Tables 2.4.8 and 2.4.9 show the same direction of impacts of rainfall deviation on different income sources. However, the results should only be used to support the main findings because of the short span of the weather data, which cover only the period from 2007 to 2014. The deviation from the mean for such a short period of time might not be a good indicator of weather shocks.

Table 2.4.8: Impact of weather shocks on equivalised income (ln) – Rainfall deviation

	Total income	Crop income	Crop revenue	Livestock	Hunting	Off Farm	Remittance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Rainfall deviation	-0.0720	-0.4681**	-0.0138	0.2178	-0.8345**	-0.3559	-0.3882
	(0.1456)	(0.2194)	(0.2202)	(0.2603)	(0.3236)	(0.2527)	(0.3140)
Household size	-0.1354***	-0.1225**	-0.1363***	-0.2032***	-0.0019	-0.0936	0.0091
	(0.0359)	(0.0588)	(0.0325)	(0.0705)	(0.0514)	(0.0787)	(0.0775)
Number of children	0.0371	0.1203***	0.0472*	0.1436**	-0.0103	0.0906*	-0.0282
	(0.0276)	(0.0430)	(0.0275)	(0.0638)	(0.0512)	(0.0530)	(0.0891)
Age	0.0006	-0.0065	0.0028	0.0157*	0.0153**	-0.0029	-0.0202
	(0.0039)	(0.0063)	(0.0042)	(0.0088)	(0.0063)	(0.0136)	(0.0157)
Marital status (Married=1)	0.0839**	-0.0476	0.0236	0.0589	0.1990***	-0.1614	0.1345
	(0.0393)	(0.0489)	(0.0369)	(0.0785)	(0.0698)	(0.1053)	(0.0901)
Education (ref: Primary school)							
Education: Secondary school	0.0658*	-0.0030	0.0424	0.0640	-0.0730	0.0300	0.0405
	(0.0393)	(0.0546)	(0.0437)	(0.0836)	(0.0611)	(0.0686)	(0.1057)
Education: High school	0.0525	-0.0607	0.0058	-0.0580	-0.0992	0.0726	-0.3247***
	(0.0385)	(0.0657)	(0.0523)	(0.0779)	(0.0774)	(0.0770)	(0.1103)
Education: Bachelor or higher	-0.0498	-0.1733**	-0.1374*	-0.1347	-0.1168	-0.1925*	-0.1535
	(0.0527)	(0.0815)	(0.0779)	(0.1179)	(0.1872)	(0.1114)	(0.1469)
Member of socio-political organisation	0.0011	0.0024	0.0049	0.0046	0.0849**	0.0246	-0.2847***
	(0.0211)	(0.0383)	(0.0287)	(0.0580)	(0.0404)	(0.0480)	(0.0679)
Household crop area ($1000m^2$)	0.0231	0.0168	0.0693	-0.0657	0.0161	0.0019	0.0109
	(0.0163)	(0.0167)	(0.0490)	(0.0574)	(0.0098)	(0.0077)	(0.0234)
Number of tractors	0.0932***	0.0396	0.1177***	0.1314*	-0.0307	0.0602	-0.2257**
	(0.0336)	(0.0506)	(0.0440)	(0.0671)	(0.0819)	(0.0841)	(0.0990)
Number of vehicles	0.1633***	0.0732	0.0520	0.1439**	-0.0173	0.1036**	-0.1099*
	(0.0258)	(0.0481)	(0.0370)	(0.0628)	(0.0552)	(0.0464)	(0.0629)
Number of phones	0.0441***	-0.0167	0.0311*	0.0438	-0.0102	0.0370	0.0498
	(0.0139)	(0.0274)	(0.0173)	(0.0273)	(0.0335)	(0.0305)	(0.0437)
Constant	7.6195***	6.1266***	7.8920***	5.3828***	2.7833***	6.5413***	6.1137***
	(0.2445)	(0.4072)	(0.2575)	(0.5103)	(0.3538)	(0.5744)	(0.7302)
Observations	29994	24179	27507	20480	16831	17865	14262
R^2	0.0353	0.1139	0.4732	0.0211	0.0368	0.3323	0.0238
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: "Rainfall deviation" is measured by the natural log of the year rainfall minus the natural log of mean annual rainfall in a given village. Standard errors (in parentheses) are clustered at the village level. The symbols ***, ** and * denote p<0.01, p<0.05 and p<0.10, respectively.

Table 2.4.9: Impact of weather shocks on per capita consumption (ln) – Rainfall deviation

	Total	Food	NonFood	Education	Health	Rent
	(1)	(2)	(3)	(4)	(5)	(6)
Rainfall deviation	-0.1557**	-0.1579**	-0.0852	-0.4967***	0.1063	0.1586
	(0.0649)	(0.0795)	(0.0892)	(0.1585)	(0.2232)	(0.1162)
Household size	-0.0900***	-0.0908***	-0.0817***	-0.2815***	-0.0638	-0.1416***
	(0.0144)	(0.0168)	(0.0182)	(0.0379)	(0.0513)	(0.0182)
Number of children	-0.0281*	0.0027	-0.0600***	0.1283***	-0.0272	-0.0248
	(0.0154)	(0.0167)	(0.0187)	(0.0348)	(0.0460)	(0.0174)
Age	0.0009	0.0006	0.0005	0.0055	0.0033	0.0012
	(0.0020)	(0.0023)	(0.0026)	(0.0055)	(0.0081)	(0.0032)
Marital status (Married=1)	-0.0046	0.0168	-0.0021	-0.1096**	-0.0134	-0.0016
	(0.0184)	(0.0179)	(0.0258)	(0.0545)	(0.0595)	(0.0236)
Education (ref: Primary school)						
Education: Secondary school	-0.0173	-0.0086	0.0160	-0.2367***	-0.0341	0.0171
	(0.0178)	(0.0190)	(0.0216)	(0.0427)	(0.0657)	(0.0246)
Education: High school	-0.0174	0.0144	0.0376*	-0.3830***	0.0458	-0.0133
	(0.0160)	(0.0184)	(0.0206)	(0.0517)	(0.0660)	(0.0226)
Education: Bachelor or higher	-0.0515**	-0.0027	-0.0213	-0.3820***	0.0536	0.0256
	(0.0241)	(0.0266)	(0.0371)	(0.0898)	(0.0878)	(0.0441)
Member of socio-political organisation	0.0215**	0.0027	0.0503***	0.0380	-0.0623	-0.0056
	(0.0096)	(0.0098)	(0.0133)	(0.0266)	(0.0386)	(0.0162)
Household crop area ($1000m^2$)	0.0071	0.0068**	0.0072	0.0215	0.0535	-0.0210***
	(0.0054)	(0.0034)	(0.0082)	(0.0167)	(0.0485)	(0.0056)
Number of tractors	0.0775***	0.0746***	0.0820***	0.0580	0.0482	0.0859***
	(0.0173)	(0.0181)	(0.0251)	(0.0445)	(0.0587)	(0.0282)
Number of vehicles	0.1476***	0.0973***	0.2633***	0.0207	0.0004	0.0228
	(0.0134)	(0.0145)	(0.0200)	(0.0346)	(0.0461)	(0.0201)
Number of phones	0.0389***	0.0234***	0.0627***	0.0526**	0.0446	0.0121
	(0.0070)	(0.0084)	(0.0088)	(0.0213)	(0.0281)	(0.0112)
Constant	7.0134***	6.4155***	5.5976***	5.5094***	2.9733***	3.5880***
	(0.1049)	(0.1193)	(0.1400)	(0.2837)	(0.4142)	(0.1562)
Observations	29986	29986	29986	22021	26588	29933
R^2	0.1597	0.0684	0.1864	0.1680	0.0089	0.0826
Individual FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: "Rainfall deviation" is measured by the natural log of the year rainfall minus the natural log of mean annual rainfall in a given village.
Standard errors (in parentheses) are clustered at the village level. The symbols ***, ** and * denote p<0.01, p<0.05 and p<0.10, respectively.

Overall, the two alternative variables, Rain month and Rainfall deviation, provide similar results to my main findings. However, the newly constructed weather shock variable has the advantage of being an absolute measure that does not depend on the duration of weather data observation, which makes it more suitable for situations when only a short period of rainfall data is available.

2.4.3 Impact of Weather Shocks on Income Inequality

The total level of household income inequality, as measured by the Gini indexes, was around 0.5 for all three provinces in the period of analysis (Table 2.4.10), which represents a high level of inequality (The World Bank, 2021). Although these Gini indexes are of similar magnitude to that for the six North Central region provinces, measured at 0.554 in 2016 (Nguyen and Tran, 2018), they are higher than those computed for rural Vietnam from 2002 to 2018 (Kang and Imai, 2012; Benjamin et al., 2017; Le and Nguyen, 2020). In addition, they are higher than the Gini index for Vietnam as a whole, which remained below 0.4 during the period 2002–2018 (The World Bank, 2021). This indicates a higher level of inequality in the three focus provinces than in the country as a whole. Moreover, the Gini indexes of the three provinces are higher than the world average Gini index of 0.47 and that of the United States, standing at 0.41 in 2021 (World Population Review, 2021). It is also important to note that the general trend was for the Gini indexes to increase with time, indicating a growing level of income inequality. Using G_p as an indicator, Ha Tinh, Thua Thien Hue and Dak Lak experienced increases of 14.65, 15.32 and 2.06 percentage points in their indexes from 2008 to 2013.

I computed G, G_a, and G_p to better evaluate the differences between distributions when dealing with negative values, following suggestions from De Battisti et al. (2019), as described in Section 2.3.4. The original Gini coefficient G is always higher than the one computed when the nonpositive values are dropped for all three provinces in all three years, which makes intuitive sense because there are households with nonpositive incomes in all those provinces in each year. However, the difference between G and G_a is not very large for years in which a province has a small number of households with nonpositive total income, as shown in Table 2.3.1. For example, the reduction from G to G_a is between 2 to 6 percentage points for Ha Tinh. The reduction tends to be larger for a province with a larger number of households with nonpositive income. For example, the reduction from G to G_a for Dak Lak is about 16.18 percentage points in 2010 and 13.37 percentage points in 2013. It is important to point out that the reduction from G to G_a depends not only on the number of households with negative incomes, but on the degree to which they are negative.

Although there is a drop from G to G_a when the nonpositive values are excluded, the G_p obtained by normalising G is higher than G_a, but smaller than G. The relative magnitudes of the three indexes are consistent for all provinces in all three years, with G the largest, followed by G_p in the second position, and G_a ranking the third.

The relative magnitude of G, G_p, and G_a also holds true for crop income Gini indexes as shown in Table 2.4.11. As expected, G_a are smaller than G, meaning that positive crop income reduces income inequality among rural households. More importantly, there is a higher level of inequality in crop income than in total income for all three provinces, as shown by larger Gini

Table 2.4.10: Gini coefficient of total income (at the household level)

	2008	2010	2013
Ha Tinh			
G	0.477	0.502	0.552
G_a (Dropping the ≤ 0 values)	0.463	0.487	0.520
G_p (the Raffinetti et al. normalisation)	0.471	0.498	0.540
Thua Thien Hue			
G	0.545	0.520	0.577
G_a (Dropping the ≤ 0 values)	0.463	0.490	0.560
G_p (the Raffinetti et al. normalisation)	0.496	0.512	0.572
Dak Lak			
G	0.541	0.550	0.576
G_a (Dropping the ≤ 0 values)	0.517	0.461	0.499
G_p (the Raffinetti et al. normalisation)	0.533	0.511	0.544

indexes in Table 2.4.11. The original Gini coefficient G for Thua Thien Hue in 2010 is greater than one (1.358), signalling there was either a significant proportion of sampled households with negative crop income or their crop income had become really negative. This is consistent with the frequency of the weather shocks variable constructed in Table 2.3.2, which show that about 50% of sampled households experiencing weather shocks at level 6 and 37% of sampled households experiencing weather shocks at level 10. By comparing the G_p of three provinces in three years, we see that crop income inequality is consistently higher for Thua Thien Hue – the province that experienced more severe weather shocks.

In this dataset, weather shocks decrease household income by reducing income from crops, and the impacts are different for households with different characteristics, so the next question to answer is: What are the impacts on income inequality? I estimate the contribution of crop income to income inequality for the three provinces of interest, as shown in Table 2.4.12 – Owing to the small number of observations, I combine the three years of household data for each province. The top (bottom) half of the table presents the Gini decomposition when using the original Gini G (G_a).

Column 1 of Table 2.4.12 displays the income share from crops, showing that it plays an important role in rural Vietnamese household income. As described in Section 2.3.4, the income share is computed as the ratio of mean crop income to mean total household income. Therefore, when the means of crop income are often smaller when including negative values, the same applies to the contribution of crop income to total income when including negative values.

Table 2.4.11: Gini coefficient of crop income (at the household level)

	2008	2010	2013
Ha Tinh			
G	0.638	0.699	0.848
G_a (Dropping the ≤ 0 values)	0.596	0.601	0.621
G_p(the Raffinetti et al. normalisation)	0.629	0.665	0.729
Thua Thien Hue			
G	0.753	1.358	0.998
G_a (Dropping the ≤ 0 values)	0.609	0.631	0.666
G_p (the Raffinetti et al. normalisation)	0.713	0.809	0.796
Dak Lak			
G	0.694	0.870	1.010
G_a (Dropping the ≤ 0 values)	0.586	0.594	0.610
G_p (the Raffinetti et al. normalisation)	0.658	0.713	0.759

Crop income makes up about 18.26%, 13.09%, and 45.52% of total income for households in Ha Tinh, Thua Thien Hue, and Dak Lak, respectively, when excluding negative values.

There are two important points to note about the Gini coefficient of crop income and total income presented in column 2 (Table 2.4.12). First, there is more inequality in crop income than total income for each province. Second, Gini coefficients when dropping negative values are always smaller than those computed using all the values. This means that positive crop income and total income contribute to reducing the level of income inequality among rural households.

We see a high Gini correlation between crop income and total income, as indicated in column 3 (Table 2.4.12). The Gini correlation coefficients are generally higher when including all values than when excluding the negative values. When excluding negative values, this correlation ranges from 0.524 for Thua Thien Hue to 0.789 for Dak Lak. The high correlation coefficients reflect the importance of crop income to the overall Gini coefficient.

Column 4 (Table 2.4.12) shows the impact of a small change in crop income on the overall Gini coefficient. When including negative values, crop income contributes very little to reducing income inequality. An increase of one percentage point in crop income, holding other things constant, decreases the Gini coefficient by 0.008% for Ha Tinh and 0.01% for Thua Thien Hue. For Dak Lak, crop income even contributes to increasing income inequality when considering all negative values.

Table 2.4.12: Gini decomposition by income source (at the household level)

	Income Share	Gini coefficient	Gini correlation with total income	Percentage change in Gini coefficient
Using G				
Ha Tinh				
Crop income	0.134	0.983	0.737	-0.008
Total income		0.768		
Hue				
Crop income	0.102	0.947	0.712	-0.010
Total income		0.747		
Dak Lak				
Crop income	0.382	0.873	0.839	0.004
Total income		0.726		
Using G_a				
Ha Tinh				
Crop income	0.183	0.612	0.660	-0.033
Total income		0.491		
Hue				
Crop income	0.131	0.637	0.524	-0.047
Total income		0.518		
Dak Lak				
Crop income	0.455	0.605	0.789	-0.022
Total income		0.461		

Note: The top (bottom) half of the table presents the Gini decomposition using the original Gini G (G_a).

When considering only positive values, crop income contributes to reducing income inequality for all three provinces, and the reduction is larger than when negative values are included. An increase of one percentage point in crop income decreases the Gini coefficient by 0.033% for Ha Tinh and by 0.022% for Dak Lak. The magnitude is greatest for Thua Thien Hue at 0.047%. Because weather shocks reduce the crop income of rural households, they will contribute to increasing income inequality. The situation is worse in Thua Thien Hue, because this province experiences more severe weather shocks and crop income plays a more important role in reducing income inequality there.

2.5 Conclusion

In this chapter, I assessed the impact of weather shocks on households' welfare and income inequality using a panel household dataset collected in three provinces of Vietnam. The necessity of identifying at-risk communities when addressing the issue of income inequality arises because resources are scarce. This chapter provides governments and NGOs with the tools to identify such communities.

To evaluate the impact of weather shocks, I first proposed a new measure of weather shocks to overcome the limitations of existing measures. The weather shocks variable is defined as the total number of days with rainfall of at least 100 mm when there was also at least two such days in a row in the period that coincides with the household survey. This measure has the advantage of being an absolute measure that does not depend on the duration of weather samples observed. Then, I computed the Gini decomposition to identify the contribution of crop income to income inequality in the three provinces. I found that weather shocks have a significant negative impact on the income sources of rural households, particularly income from crops. In addition, weather shocks affect households with different characteristics differently. Farmers with larger crop areas are more severely affected, and ethnic minority groups and large households are also disproportionately impacted by weather shocks.

I found that crop income contributes to reducing income inequality in the three provinces by computing the Gini decomposition of income sources. Because weather shocks reduce income from crops, they contribute to increasing income inequality. The results from this chapter should assist governments and NGOs to identify at-risk communities that are more prone to weather shocks to provide necessary support. In addition, it is important for governments in countries that are agriculturally intensive to create and facilitate a market for weather derivatives to mitigate the effects of unfavourable weather patterns.

This study has has several limitations that can potentially be improved with further research. First, it is important to investigate the mechanisms for which weather shocks are translated into income inequality. Is it the case that farmers are afraid of weather shocks so they stop growing crops or transition to other occupations? Second, although the Gini decomposition computed using the original G and G_a provides valuable information about the contribution of crop income to reducing income inequality, it would be interesting to see how the results change if using G_p. Future researchers might explore how to compute the Gini decomposition when using G_p, which is a good complementary index for the original G and G_a when a distribution includes negative values. Furthermore, the contribution of other income sources, such as livestock income or hunting income, could be added in the Gini decomposition to obtain a more complete picture. However, the number of households with negative livestock income might be different from the number of households with negative hunting income, in which case the current Gini decomposition method would not be useful. Thus, another direction for future research is to develop a decomposition method that is suitable for dealing with different numbers of negative values for different income sources.

Chapter 3

Economic Disparity, Life Events and the Gender Mental Health Gap

3.1 Introduction

Despite a century of economic and political efforts to improve equality in wealth and human well-being across the globe, disparities – including in health outcomes – remain a persistent feature of modern society. There are strong correlations between health and socioeconomic variables, including income, wealth, education, labour market participation and professions (Grossman, 2017; Kim and Koh, 2021; Lenhart, 2019; O'Donnell et al., 2007; Pickett and Wilkinson, 2015; Kawachi and Kennedy, 1999; Li and Powdthavee, 2015; Kulhánová et al., 2014; Huguet et al., 2008; Lordan et al., 2012; Cutler and Lleras-Muney, 2006). These associations persist over a variety of health markers such as longevity, body mass index and subjective-wellbeing, and often reinforce the advantages of higher economic status (Pudrovska et al., 2014; Kim et al., 2020). It has been shown that income disparities among others are associated with individual well-being (Clark et al., 2008; Clark and D'Ambrosio, 2015). The associations between health outcomes and economic inequalities have been of significant interest to policymakers, as evidenced by the loci of social policies and interventions that compress economic outcomes can generate positive spillover effects to public health outcomes (Hashmi et al., 2020; Osmani and Sen, 2003).

In this paper, we examine sources of gender inequality in mental health in Australia. We are interested in determining if economic differentials (in terms of income, labour market participation, and education) between men and women are sufficient to explain observed outcome gaps in psychological wellbeing. We show that women have generally poorer mental health than men on average, and are strongly over-represented amongst individuals with very poor health (there is a larger percentage of women at the lower end of the mental health distribution).

Using regression-based decompositions based upon the Recentered Influence Function (RIF), we show that socioeconomic variables such as income and being employed do indeed play important roles in understanding the aggregate gender mental health inequality, accounting for the entire gender gap amongst individuals with very poor psychological wellbeing (the 10th quantile), while education partly closes that gap.

Apart from economic factors like income or education, we are also interested in understanding whether negative life events contribute to the mental health gap between men and women. At the individual level, we found that negative life events such as being victims of physical violence have dire consequences on mental health scores for both men and women, especially in the lower quantiles. Although the impacts of being exposed to violence are much stronger for women than men in all the main quantiles, the rare occurrence of such events to both genders means that those negative life events do not contribute significantly to the mental health inequality between men and women at the aggregate level. However, the scale and severity of domestic violence, and therefore its impact on mental health, could be significantly worse due to under-reporting. Many victims of domestic violence do not report it, for multiple reasons (see for example Davis et al., 2003; Drijber et al., 2013). Moreover, COVID-19 may have increased the prevalence of violence against women (Su et al., 2022). For example, 16 women were killed because of domestic violence between 23rd March and 12th April, 2020 in the UK compared to an average of five for the same period in the last decade (Grierson, 2020). Interventions tailored to domestic violence victims, such as shelters, are needed to improve the victims' mental health and close the research gap. We also found that unobserved heterogeneity, such as biological differences or differences in reporting and responding to mental health issues, contribute significantly to the gap between men and women.

Our research contributes to a body of literature on health inequality in several ways. First, we go beyond the mean and provide full distribution decomposition to determine the structure of the mental health inequality. This is particularly relevant in the case of mental health, because the social and economic cost of **very poor** mental health is often catastrophic; that is, people at the very low end of the mental health spectrum can be suicidal or incapable of engaging in productive activity, including education, work and social engagements, and they often require a much higher level of healthcare interventions (inclusive of medications, therapies and social support). We look at the two distributions (by gender group) and the low mental health tail and examine the associations within each distribution. This is often called subgroup analysis, an important part of intervention targeting and one of the ideas behind distributional cost effectiveness analysis. Second, we show that economic variables such as income and being employed do not only contribute to inequality within each group of men and women (intra-group disparity), but are strongly associated with the gender mental health gap (inter-group disparity). However, education as a measure of socioeconomic status reduces the inter-group disparity. While individuals who experience negative life events often have poor mental health, the rarity of such events in both men and women means they have a minor contribution to the mental health gap. Third, we contribute to the literature on so- cioeconomic inequalities of life shock exposure by showing that women with poor mental wellbeing are particularly vulnerable to those negative life events. Our findings can guide the design of more effective policies for mitigating gender inequalities, and especially the mental health disparities that women face.

The paper is structured as follows. Section 3.2 describes the dataset and variables of interest and why they are relevant to answering our research question. Section 3.3 describes regression models and why they are appropriate for analysing the gender mental health gap. Section 3.4 summarises the correlates of mental health inequality. Section 3.5 provides a discussion of factors that contribute substantially to the gender mental health gap, and outlines the limitations of the research. Section 3.6 concludes and suggests policies that might be more effective than current approaches in alleviating very poor mental health and closing the gender mental health gap.

3.2 Data and Study Variables

3.2.1 The HILDA Dataset

This paper studies gender inequality in mental health using data from Release 20 of the Household, Income and Labour Dynamics in Australia (HILDA) survey. HILDA is an ongoing longitudinal study that began in 2001 with a nationally representative sample of Australian households (Watson and Wooden, 2021). The HILDA longitudinal survey is comparable to the USA's Panel Study of Income Dynamics (PSID), the British Household Panel Survey (BHPS) or the German Socio-Economic Panel (SOEP) (Hashmi et al., 2020). The first wave (in 2001) comprised 13,969 participants from 7,682 households and 2,153 households were added in 2011. We pool 20 years of data (from 2001 to 2020) for this paper, because in studying inequality, we care about differences between individuals rather than the changes for individuals over time.[1]

The HILDA survey uses multiple instruments (questionnaires) to collect information about the composition of households, income, employment, family relationships and personal (individual) wellbeing. Individual person questionnaires are administered to every member of the household aged 15 years or older (with parental consent sought before interviewing persons aged under 18 years and living with their parents). All respondents completing an individual questionnaire are also asked to complete a separate self-completion questionnaire (SCQ) about general health and wellbeing, lifestyle and living situation, personal and household finances, job and the workplace,

[1]We note here that it is possibly useful to know if someone has "repeated experience" of traumatic events, and will come back to this point in the discussion section.

and parenting, data from which were used to derive some variables of interest for use in the current study.

3.2.2 Variables of Interest

We provide a description of the variables used for our analysis, starting with our dependent variable, the Mental Health Inventory (MHI-5) score, followed by other explanatory variables including standard socioeconomic variables such as marital status, education levels, household income, indicators capturing whether or not an individual experienced a negative life event in the past year, and residential area index of socioeconomic advantage. All monetary values used in this study are adjusted for inflation. We calculate the Consumer Price Index (CPI) for each state each year as the average of four quarters - all commodity groups, following the data from the Australian Bureau of Statistics with index reference period: 2011-12 = 100.0.

Mental health

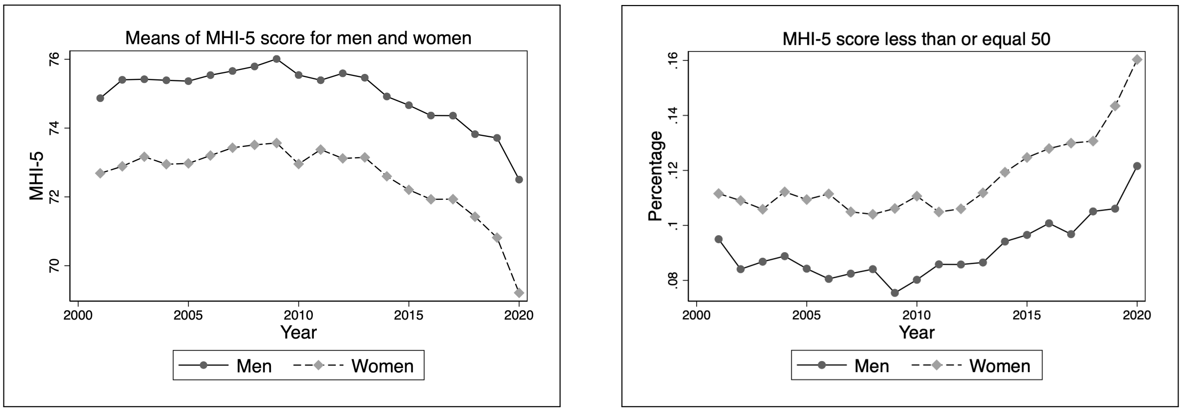
HILDA administers the 36-Item Short Form Survey (SF-36), of which the mental health is one of eight domains (alongside vitality, physical functioning, bodily pain, general health perceptions, physical role functioning, emotional role functioning and social role functioning). The SF-36 is considered one of the best multi-dimensional instruments to measure health outcomes, thanks to substantial international research evidence of its responsiveness, validity and reliability in a range of studies. The mental health component consists of questions in the general health and wellbeing section,[2] from which the MHI-5 score was constructed following the procedure described in Ware et al. (1994).[3] The MHI-5 values range from 0 to 100, with lower values indicating poorer mental health.

We observe that men have a significantly higher mean MHI-5 score than women in all years of the survey (from 2001 to 2020), as shown in the left panel of Figure 3.2.1. This observation is consistent with regression findings reported in Hashmi et al. (2020), which suggests that women, in general, have lower MHI-5 scores than men. Additionally, the percentage of women who have MHI-5 scores of less than or equal to 50 is consistently higher than that of men in each of those years, as shown in the right panel of Figure 3.2.1. This means that not only do women have lower MHI-5 scores on average, but also a higher proportion of women have lower MHI-5 scores than men. In other words, the left tail of the MHI-5 distribution for women is always thicker than that for men, as shown in Figure 3.2.2. Using a regression with gender as a dummy variable, we confirm that men have statistically significant higher MHI-5 score with lower variance on average than women. This is a signal that, compared to men, women are

[2]See Appendix B.1 for the actual questions.

[3]See Appendix B.2 for the construction of MHI-5, which includes five items: been a nervous person, felt so down in the dumps nothing could cheer you up, felt calm and peaceful, felt down, and been a happy person.

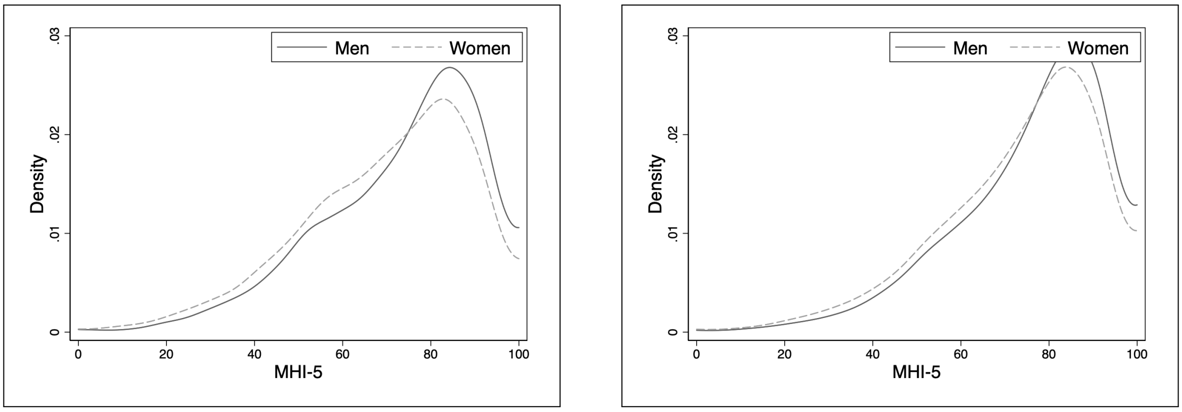
Figure 3.2.1: MHI-5 scores, HILDA, 2001-20



Notes: The left graph shows the means of MHI-5 scores for men and women across all years from 2001 to 2020. The right graph shows the percentages of men and women with MHI-5 scores less than or equal to 50 in all those years.

more likely to suffer from poor mental health. Additionally, there are significant differences between the mean MHI-5 scores men and women across sub-groups, as shown in Table 3.2.1.

Figure 3.2.2: Kernel density of MHI-5 scores, HILDA, 2001–20



Notes: The kernel densities of MHI-5 scores were estimated using the Gaussian kernel function, bandwidth 3.5 with boundary correction of 0 and 100. The left figure is the kernel density estimated for men and women in 2020. The right figure is the kernel density estimate for the pooled sample.

Demographic variables

We include demographic variables such as marital status, household size, and individual age in the analysis. These variables form the first group of variables in Table 3.2.1. Marital status has three categories: legally married or de facto, divorced/separated, and single; "marital status = single" is then used as the reference group. We hypothesise that people in the "Legally married or de facto" group are in a supportive relationship that has a positive impact on their mental health, thus their self-reported MHI-5 scores would be higher on average than those for members of the "single" group. Meanwhile, people in the "Divorced or separated" group have experienced a shock that is likely to have harmed their mental health. As can be seen from

Table 3.2.1: T-tests for men and women, HILDA data, 2001–21

	Men	Women	Diff. (Men - Women)
MHI-5 score	74.9702	72.5079	2.4623***
Demographic variables			
Marital Status: Legally married or de facto	0.6604	0.6153	0.0450***
Marital Status: Divorced/separated/widowed	0.0884	0.1715	-0.0832***
Marital Status: Single	0.2513	0.2131	0.0382***
Household size	2.9005	2.8624	0.0381***
Age	44.5748	44.9505	-0.3756***
Age square	2326.8461	2366.9377	-40.0916***
Living outside a major city	0.1296	0.1252	0.0044***
Young family	0.1358	0.1468	-0.0110***
Foreign born	0.2126	0.2068	0.0058***
Economic variables			
After-tax equivalised income	47703.8645	45440.3430	2263.5215***
Labour force status: Employed	0.6939	0.5876	0.1064***
Labour force status: Unemployed	0.0415	0.0336	0.0078***
Labour force status: Not in the labour force	0.2646	0.3788	-0.1142***
Education: Year 12 or below	0.4159	0.4967	-0.0809***
Education: Certificates and diploma	0.3623	0.2479	0.1144***
Education: Undergraduate	0.1720	0.2125	-0.0405***
Education: Postgraduate	0.0498	0.0429	0.0069***
Life event variables			
Death of spouse or child	0.0063	0.0100	-0.0037***
Death of close relative or family member	0.1107	0.1228	-0.0120***
Serious injury or illness to family member	0.1308	0.1695	-0.0387***
Natural disaster damaged/destroyed home	0.0150	0.0141	0.0009
Fired or made redundant	0.0401	0.0249	0.0152***
Victim of physical violence	0.0148	0.0158	-0.0011**
Close family member detained in jail	0.0123	0.0184	-0.0061***
Detained in jail	0.0040	0.0013	0.0026***
Residential area index of socioeconomic advantage			
1st quintile	0.1905	0.1943	-0.0037**
2nd quintile	0.1977	0.1978	-0.0002
3rd quintile	0.1963	0.1991	-0.0028*
4th quintile	0.2027	0.2029	-0.0002
5th quintile	0.2128	0.2059	0.0069***

Note: The list of variables remained in our analysis are divided into *demographic, economic*, *life event*, and *residential area index of socioeconomic advantage* groups in the first column. The second and the third columns show the means of those variables for the men and women group. The fourth column shows the difference between the means, with ***, **, and * denoting p<0.01, p<0.05, and p<0.10, respectively.

Table 3.2.1, men are more likely to be in a committed relationship, while women have higher probability of relationship breakdown.

Other demographic characteristics like living area, whether a person has small children or being born overseas are included as confounding factors. *Living outside a major city* is a dummy variable that equals 1 if the person resides in outer regional, remote or very remote Australia. *Young family* is a dummy variable indicating whether the individual has any children aged 0-4 years. *Foreign born* is a dummy variable indicating that an individual was born outside of Australia.

Economic variables

After-tax equivalised income, labour force participation status and education level are included as measures of economic status, which make up the middle group in Table 3.2.1. With regard to after-tax equivalised income, we use the "OECD-modified scale" to adjust for household size,[4] following the formula

$$\text{Equivalised income} = \frac{\text{Household disposable regular income}}{1 \times \text{first adult} + 0.5 \times \text{additional adults} + 0.3 \times \text{children}}. \tag{3.1}$$

Men earn significantly more than women, receiving almost 5% more after-tax equivalised annual income on average. Please note that this is undoubtedly an underestimate of the true income difference (Australia's national gender pay gap in 2021 was 13.8%),[5] because calculating after-tax equivalised income gives the same value for everyone in the household.

Labour force status has three categories: employed, unemployed and not in the labour force, with the last group treated as the reference group. As seen in Table 3.2.1, men are both more likely to be employed, and more likely to be unemployed than women, while women are more likely to report being "not in the labour force". Note that this group includes women who perform unpaid work in the home as a primary carer, either for children or for older adults in their family.

The highest education level achieved has four categories: certificates and diplomas, undergraduate, postgraduate, and grade 12 or below, with the last group treated as the reference group. Women are more likely to have an undergraduate degree, and make up a larger proportion of the lower education group (i.e. Grade 12 or below) than men, while men are more likely to obtain certificates, diplomas and postgraduate degrees, as shown in Table 3.2.1.

Life events

In the HILDA survey, a negative life event is defined as a major life event that has a negative impact on one's life. Negative life events and their descriptive statistics are listed in Table

[4] https://www.oecd.org/els/soc/OECD-Note-EquivalenceScales.pdf

[5] https://www.wgea.gov.au/publications/australias-gender-pay-gap-statistics

3.2.1. All of these life event variables are dummy variables, with a value of 1 indicating that an individual had experienced that negative event in the past year.

All major life events occur in the male and female groups at proportions that are significantly different, with the exception of experiencing a natural disaster that damaged or destroyed their home. While men are more likely to be sacked, made redundant or jailed, women reported more life events such as being a victim of physical violence, death of a spouse, child, close relative or family member, a family member experiencing serious injury or illness, or a close family member being jailed

Residential area index of socioeconomic advantage

We control for neighbourhood characteristics using five quintiles constructed from the socio-economic indexes for areas (Summerfield et al., 2020). The addition of neighbourhood characteristics also likes to the theme of economic inequality that we consider in this study.

3.3 Econometric Models

Distributional regression models have become popular in recent years due to their advantage of providing a more complete picture of distributional characteristics than conventional regression techniques that focus on explaining the mean. Empirical researchers are often interested in understanding how a relationship between explanatory variables and a dependent variable differs according to the location in the distribution, because the expectation of a variable is not necessarily representative of its entire distribution. This is especially true when studying phenomena such as in- equality or efficiency, which are intrinsically linked to distributions rather than means (Kneib et al., 2021).

There are many forms of distributional regression, most of which are reviewed comprehensively in Kneib et al. (2021). In this study, we are interested in the impact of specific (explanatory) variables on the distribution of the dependent variable (MHI-5 score). The focus, therefore, is on structured regression models rather than the prediction-oriented approaches. Distribution regression relates some distributional statistics $\nu(F)$ to multiple explanatory variables $\mathbf{X}$. For example, F could be a univariate mental health score distribution function and $\nu(F)$ is a generic functional such as quantile, inequality measure or poverty index. Two common approaches in recent literature that are most appropriate for our focus are unconditional quantile regression (Firpo et al., 2009) and distribution function modelling (Chernozhukov et al., 2013).

We use the unconditional quantile regression (UQR), a new technique developed by Firpo et al. (2009), to model the unconditional quantiles. This is done by transforming the variable using a recentered influence function (RIF) and regressing the results against explanatory factors. The

UQR allows us to evaluate the impact of changes in the distribution of the explanatory variables ($\mathbf{X}$) on quantiles of the unconditional distribution of an outcome variable (Y). The UQR differs from the conditional quantile regression (CQR) (Koenker and Hallock, 2001) because τth unconditional quantile is often not the same as the τth conditional quantile. Also, while we can usually switch from conditional to unconditional by applying the law of iterated expectations, this property is not available for quantiles, making the interpretation of conditional quantile estimates "uninteresting". Although the CQR has been used to study health-related questions (Kessels, 2020), the UQR has been applied mostly in studying gender wage gaps. UQR looks at the entire distribution of the dependent variable, and thus is relevant for understanding the key factors that affect different subgroups, from which policies and interventions can be developed to target (sub-)populations with different needs (Borah and Basu, 2013). For instance, it is important to know that, on average, women have poorer mental health than men. A policy designed to improve mental health for women, such as community-based counselling for all (targeting the mean), might show no effect at all in improving mental health for those at the absolute bottom end of the mental health spectrum, because they either do not (choose to) use the services or generic mental health counselling is only a part of multiple policies and social interventions designed to encourage access and utilisation by this subgroup. Counselling alone might in fact increase disparities between women across different socio-economic groups.

We first apply the UQR method to estimate the correlates of the mental health inequality in Australia for men and women. The regressions for men and women are:

$$\mathrm{RIF}(y_{Mit}, q_\tau, F_{Y_M}) = \alpha_M + \gamma_t + \mathbf{x}'_{Mit}\boldsymbol{\beta}_{M,\tau} + \epsilon_{Mit,\tau} \tag{3.2}$$

and

$$\mathrm{RIF}(y_{Wit}, q_\tau, F_{Y_W}) = \alpha_W + \gamma_t + \mathbf{x}'_{Wit}\boldsymbol{\beta}_{W,\tau} + \epsilon_{Wit,\tau} \tag{3.3}$$

where y_{Mit} and y_{Wit} are the i's observation of mental health in year t for men and women, respectively; α_M and α_W are constant terms for men and women; γ_t is the year fixed effects; $\mathbf{x}_{Mit}$ and $\mathbf{x}_{Wit}$ are the vectors of demographic, economic and life event variables for men and women i in year t; $\boldsymbol{\beta}_{M,\tau}$ and $\boldsymbol{\beta}_{W,\tau}$ are vectors of coefficients of the corresponding covariates at the τ quantile for men and women; $\epsilon_{Mit,\tau}$ and $\epsilon_{Wit,\tau}$ are the error terms.

Although HILDA collects panel data, we do not use standard panel data techniques like random effects or fixed effects because we are interested in inequality; using fixed effects will wipe out the dispersion that we care about, and there is no random effects version of our approach. Instead, we run the pooled model with year fixed-effects. We acknowledge that there is likely correlation between ϵ_{it} and $\epsilon_{i,t-1}$ that is not necessarily controlled for in the model (i.e., an individual responses can be similar over the years). However, we do not include a lag dependent variable, because if we did, what explains inequality in y_t will be in y_{t-1}, and what explains inequality in y_{t-1} will be in y_{t-2}. This recursion will not help us to get to the root source of inequality. As a result, instead of accounting for the correlations as coefficients in the model, we resolve this issue using clustered robust standard error at the individual level.

We then use the RIF regression results to decompose the predicted outcome gap to examine how each factor contributes to the gender-based mental health gap. There are many ways to perform decomposition in economics (e.g., Fortin et al., 2011). Here, we focus on the aggregate effect of each explanatory variable, so we compute the gap at the τth quantile following the formula:

$$\hat{\Delta}_{j,\tau} = \bar{x}_{jM}\hat{\beta}_{jM,\tau} - \bar{x}_{jW}\hat{\beta}_{jW,\tau} \tag{3.4}$$

where $\hat{\Delta}_{j,\tau}$ is the gap at quantile τth associated with variable j; $\bar{x}_{jM}$ and $\bar{x}_{jW}$ are the means of variable j for the men and women group; $\hat{\beta}_{jM,\tau}$ and $\hat{\beta}_{jW,\tau}$ are the estimates of variable j from regression (3.2) and (3.3) respectively.

Furthermore, we use Oaxaca-Blinder decomposition to decompose the difference in RIF between men and women into two components:

$$\hat{\Delta}_{j,\tau} = (\bar{x}_{jM} - \bar{x}_{jW})\hat{\beta}_{jW,\tau} + \bar{x}_{jM}(\hat{\beta}_{jM,\tau} - \hat{\beta}_{jW,\tau}) = \hat{\Delta}_{jE,\tau} + \hat{\Delta}_{jC,\tau} \tag{3.5}$$

where $\hat{\Delta}_{jE,\tau}$ is the endowment effect and $\hat{\Delta}_{jC,\tau}$ is the coefficient effect associated with variable j.

3.4 Results

3.4.1 Correlates of Mental Health Inequality

As discussed in Section 3.2, our study aims to understand the role of economic factors and major life events in explaining the gender-based disparity in MHI-5 scores. Results for five main quantiles (10th, 25th, 50th, 75th and 90th) for men and women are shown in Tables 3.4.1 and 3.4.2, respectively.[6] Most variables are statistically significant and have the expected signs, with the economic variables having positive impacts on mental health score while life event shocks have negative impacts on mental health score, and the effect magnitudes being disproportionately large at the lower quantiles.

[6]The full results are provided in Appendix B.7

Table 3.4.1: RIF Regression for Men – Quantiles, HILDA, 2001-20

	Q10	Q25	Q50	Q75	Q90
	(1)	(2)	(3)	(4)	(5)
Marital status (ref: Single)					
Marital Status: Legally married or de facto	3.8047***	4.9766***	2.8572***	1.5560***	0.9726***
Marital Status: Divorced/separated/widowed	-1.4989	-1.1898	-0.1629	-0.2391	-0.3037
Household size	0.5505***	0.5949***	0.2391***	0.2279***	0.1679***
Age	-0.7049***	-0.8411***	-0.3565***	-0.1880***	-0.1621***
Age square	0.0091***	0.0103***	0.0050***	0.0033***	0.0028***
Living outside a major city	2.0662***	1.4209**	0.9259**	0.7297**	0.7969***
Young family	-1.0371**	-1.4837***	-0.6037*	-0.5062*	-0.3002
Foreign born	-0.1559	-0.9857*	-1.0345***	-0.5595*	-0.2137
Log of after-tax equivalised income	2.5925***	3.4689***	2.4390***	1.5285***	0.9330***
Labour force status (ref: Not in the labour force)					
Labour force status: Employed	9.5647***	8.7401***	3.5972***	1.8173***	1.1235***
Labour force status: Unemployed	0.6572	-0.9490	-0.0283	-0.0060	-0.1953
Education (ref: Year 12 or below)					
Education: Certificates and diploma	-0.0185	0.8665	0.6038*	0.2260	0.1781
Education: Undergraduate	1.0112*	2.2119***	0.1304	-0.5425	-0.7760***
Education: Postgraduate	0.9662	2.7461***	0.6920	-0.3455	-0.7807*
Death of spouse or child	-8.6758***	-8.9625***	-4.9791***	-2.8283***	-2.0984***
Death of close relative or family member	-1.6002***	-1.7416***	-0.6258***	-0.4953***	-0.3902**
Serious injury or illness to family member	-3.2562***	-3.7891***	-2.3313***	-1.7381***	-1.5022***
Natural disaster damaged/destroyed home	-3.7757***	-4.3044***	-1.8353***	-1.3340***	-0.8970**
Fired or made redundant	-3.6356***	-3.3375***	-1.9318***	-1.5969***	-1.2075***
Victim of physical violence	-17.4942***	-12.3104***	-5.5307***	-2.8199***	-1.6649***
Close family member detained in jail	-2.6129*	-3.7574***	-1.8020***	-0.8495*	-0.6998*
Detained in jail	-16.0369***	-10.6467***	-3.1745***	-0.8848	-0.6286
Residential area index of socioeconomic advantage (ref: 1st quintile)					
2nd quintile	1.4266**	2.1901***	1.0504***	0.5361*	0.3798
3rd quintile	2.2640***	3.5959***	1.4417***	0.5505*	0.1826
4th quintile	2.8677***	4.3004***	1.5757***	0.7256**	0.5677*
5th quintile	3.4326***	4.8127***	2.0495***	0.5273	0.1349
Constant	25.2133***	31.4837***	54.1135***	71.2021***	84.0031***
Average RIF	52.6727	66.8980	80.7696	89.5043	94.8236
Observations	82975	82975	82975	82975	82975
R^2	0.0447	0.0557	0.0467	0.0383	0.0270
WE	Yes	Yes	Yes	Yes	Yes

Q10, Q25, Q50, Q75, and Q90 are the 10th, 25th, 50th, 75th and 90th quantiles. Clustered robust standard errors at the individual level. The symbols ***, **, and * denote that p<0.01, p<0.05, and p<0.10, respectively.

Table 3.4.2: RIF Regression for Women – Quantiles, HILDA, 2001-20

	Q10	Q25	Q50	Q75	Q90
	(1)	(2)	(3)	(4)	(5)
Marital status (ref: Single)					
Marital Status: Legally married or de facto	4.9545***	4.2649***	2.2664***	0.9987***	-0.1077
Marital Status: Divorced/separated/widowed	-0.3745	-0.6511	-0.7690	-0.8817**	-1.1286***
Household size	0.2737	0.0296	-0.0286	-0.0209	-0.0421
Age	-0.4264***	-0.3678***	-0.0930**	0.0609*	0.0581*
Age square	0.0073***	0.0063***	0.0028***	0.0011***	0.0009***
Living outside a major city	1.5188**	1.0120*	1.0443**	0.7613**	0.8469**
Young family	3.8730***	3.2544***	1.9309***	0.8988***	0.3934*
Foreign born	0.6676	-0.8427*	-0.9383**	-0.6334**	-0.4255
Log of after-tax equivalised income	1.9719***	2.8623***	2.5637***	1.6515***	1.0638***
Labour force status (ref: Not in the labour force)					
Labour force status: Employed	8.6687***	6.7399***	3.0664***	1.3093***	0.2621
Labour force status: Unemployed	-1.4798	-2.2026***	-1.2901***	-0.6314*	-0.4006
Education (ref: Year 12 or below)					
Education: Certificates and diploma	1.2768**	1.3045***	1.0079***	0.6144**	0.2557
Education: Undergraduate	2.4511***	3.1718***	1.5967***	0.2224	-0.6528**
Education: Postgraduate	2.8352***	2.4625***	0.5771	-0.8050	-1.5130***
Death of spouse or child	-10.2787***	-10.2712***	-5.7620***	-3.8009***	-3.0206***
Death of close relative or family member	-2.5401***	-2.2527***	-1.4745***	-1.0125***	-0.7926***
Serious injury or illness to family member	-4.5806***	-4.0107***	-2.8514***	-2.2031***	-1.8147***
Natural disaster damaged/destroyed home	-4.2828***	-3.5856***	-2.2729***	-1.9417***	-0.9131**
Fired or made redundant	-7.9633***	-5.9125***	-4.4727***	-2.8502***	-1.5674***
Victim of physical violence	-25.8405***	-16.8173***	-7.8593***	-4.0886***	-1.8104***
Close family member detained in jail	-4.6872***	-3.5495***	-2.4150***	-1.2543**	-0.6944
Detained in jail	-3.5027	-2.5117	-3.2023*	-1.1419	0.3133
Residential area index of socioeconomic advantage (ref: 1st quintile)					
2nd quintile	2.1634***	2.5254***	2.1729***	1.1692***	0.4301
3rd quintile	3.0514***	3.6412***	2.5524***	1.3348***	0.4614
4th quintile	3.0194***	3.5146***	2.5589***	1.2331***	0.3324
5th quintile	4.7782***	5.0572***	3.7012***	1.6157***	0.6462*
Constant	20.9046***	27.3819***	45.1824***	64.3291***	78.8634***
Average RIF	48.8132	63.4395	78.0359	87.7448	93.9473
Observations	94337	94337	94337	94337	94337
R^2	0.0495	0.0617	0.0520	0.0453	0.0304
Year FE	Yes	Yes	Yes	Yes	Yes

Q10, Q25, Q50, Q75, and Q90 are the 10th, 25th, 50th, 75th and 90th quantiles. Clustered robust standard errors at the individual level. The symbols ***, **, and * denote that p<0.01, p<0.05, and p<0.10, respectively.

Demographic Factors

Being in a committed relationship (married or de-facto) has a positive impact on MHI-5 score, while experiencing a relationship breakdown (divorced or separated) has a negative impact on MHI-5 score. This suggests a protective effect of a strong relationship on mental health, as expected and discussed widely in the literature (Coombs, 1991; Glenn, 1975; Diener et al., 2000; Kahneman et al., 1999; Stack and Eshleman, 1998). This result holds for both men and women at all five quantiles, even though the magnitude of the impact differs between men and women across quantiles. Relationship status appears to have a stronger impact on mental health score for people at the low end of the MHI-5 distribution, as evidenced by larger absolute values of the coefficients at lower quantiles. For example, if the proportion of legally-married or de-factor men in the population increased by 10 percentage points, from 66.04% as currently observed to 76.04%, the MHI-5 10th quantile for men would increase by 0.36 points or $(3.8047/52.6727) \times 0.1 \approx 0.72\%$ in relative terms, while the MHI-5 90th quantile for men would only increase by 0.1 points or $(0.9726/94.8236) \times 0.1 \approx 0.1\%$ in relative terms.

Interestingly, being in a committed relationship tends to benefit women more while experiencing a relationship breakdown tends to hurt women less, but these results only hold at the low quantile of the mental health distribution (e.g., the 10th quantile). For example, if the proportion of legally married or de factor women in the population increased by 10 percentage points, from 61.53% as currently observed to 71.53%, the 10th quantile of MHI-5 for women would increase by 0.46 points or $(4.9545/48.8132) \times 0.1 \approx 1.02\%$ (compared to 0.72% for men) in relative terms. This is probably due to the fact that most divorce filers are women, who feel relieved after ending a relationship that was not working for them (Brinig and Allen, 2000; Parker et al., 2022). However, this psychological benefit is not necessarily enjoyed by all women (Saleh and Luppicini, 2017).

Other demographic variables have the same direction of effects on men and women, with the exception of having small children. More specifically, large household size and living outside a major city tend to have positive impacts on men's and women's mental health scores, especially at the lower quantiles, while being born overseas is associated with lower mental health score for both men and women. It is noteworthy that the impact sizes here are small. However, having small children aged 0–4 years has a negative impact on mental health scores for men but a positive impact on mental health scores for women. It is undeniable that having small children changes couples' lives significantly; maternal instinct probably compensates for the stress that women face when having young children.

Economic factors

Economic variables like after-tax equivalised income, being employed or having a higher level of education were correlated positively with MHI-5 scores, especially in the lower quantiles. This might reflect increasing returns to scale in mental health: that is, for someone whose mental state is low, improvements in other aspects of life might have much stronger returns – in terms of improving mental health – than for those whose mental health scores were already high. Income is more strongly correlated with MHI-5 scores for men at the 10th and 25th quantiles than women, but the impact's magnitude reduces in higher quantiles. At the 50th, 75th and 90th quantiles, increasing income appears to benefit the mental health of women more than men. For example, a 1% increase in income would increase the average 10th quantile of mental health for men and women by 0.03 and 0.02, but would increase the average 90th quantile of mental health for men and women by 0.0093 and 0.011, respectively.

Being employed improves MHI-5 score significantly for both men and women, and the impacts are amplified at lower quantiles. This result is consistent with findings in the literature that suggest supported employment benefits people with severe mental illness (Burns et al., 2007; Affleck et al., 2018). Interestingly, being unemployed is not statistically correlated with lower MHI-5 score for men, while the opposite is true for women. This is probably due to a tendency that women are more worried about their ability to secure a job compared to their counterparts while it has been shown that greater job insecurity is associated with reduced self-reported health (Lepinteur, 2021).

Another important observation is that higher level of education is not always associated with higher mental health scores. Having a higher level of education is associated with higher mental health scores for those at lower quantiles but lower mental health scores of those at higher quantiles. An explanation for this could be that higher education is not necessarily translated into higher paid jobs. For example, in order to work as a social worker in Australia, one needs a bachelor or master of social work degree. However, social workers are often paid poorly, which is likely to reduce their mental health.

Major life events

All negative life events reported in HILDA had negative impacts on mental health score for both men and women.[7] Notably, the impact magnitudes were always larger for the lower quantiles. Among the eight major life shocks in our analysis, the top three that significantly reduce the mental health of men in the 10th quantile are being a victim of physical violence, being detained in jail or experiencing the death of a spouse or child. While it is common knowledge that women who experienced physical assaults suffer from poor mental health, our findings remind us that being a victim of physical violence also harms men's mental health. This side of the physical violence story has attracted much less media and research attention than the women's side. If we are to talk about equality, it is essential to emphasise the fairness for all genders.

[7] Hashmi et al. (2020) did not study the mental health of men and women separately, but they found that exposure to negative life events was most harmful for the mental health of people in the most disadvantaged socioeconomic groups.

However, being a victim of physical violence does reduce women's mental health score more than men's, and this holds true at all five quantiles across the distribution. For example, if the proportion of victims of physical violence in the population was to increase by 10 percentage points, the average 10th quantile of MHI-5 for women would drop by 5.29 points, while that of men would drop by 3.32 points. Experiencing the death of a spouse or child and being sacked or made redundant have stronger negative consequences for the mental health of women than men, ranking as the second and the third major predictors for poor mental health in women. The negative impact of these negative events might have implication on intergenerational mobility as mothers' mental health is found as a channel that translates financial problems on the noncognitive outcomes of their children (Clark et al., 2020).

Most importantly, we notice that economic outcomes like household income or education are less important contributors to low MHI-5 score **at the individual level** than negative life events such as being a victim of violence, especially for those at the bottom of the MHI-5 distribution. With larger effect sizes (especially in the lower quantiles) and higher prevalence of exposure to negative life events among women, this conclusions aligns with the hypothesis that negative life events are strongly correlated with worse mental health in women who already are in the low mental health score range

3.4.2 The Gender Mental Health Gap: Contributing Factors

Based on the RIF regression results presented in Section 3.4.1, we decompose the predicted outcome gap using Equation (3.4) to understand the contribution of each explanatory variable to the gender-based mental health gap. The results presented in Table 3.4.3 indicate that a factor could contribute to increasing or decreasing the gender-based mental health gap depending on which position of the distribution we examine. At the low end, inequality in income would be enough to account for the disparity in mental health between men and women. However, at the top end, it works in the opposite direction – inequality in income narrows the gender mental health gap.

Table 3.4.3: RIF Difference – Quantiles, HILDA, 2001-20

	Q10	Q25	Q50	Q75	Q90
	(1)	(2)	(3)	(4)	(5)
Marital status (ref: Single)					
Marital Status: Legally married or de facto	-0.5362	0.6620	0.4922	0.4130	0.7085
Marital Status: Divorced/separated/widowed	-0.0682	0.0066	0.1175	0.1301	0.1668
Household size	0.8133	1.6408	0.7754	0.7209	0.6075
Age	-12.2539	-20.9591	-11.7105	-11.1175	-9.8372
Age square	3.8957	9.0548	5.0068	5.0750	4.3849
Living outside a major city	0.0777	0.0575	-0.0107	-0.0007	-0.0027
Young family	-0.7093	-0.6792	-0.3654	-0.2007	-0.0985
Foreign born	-0.1712	-0.0353	-0.0259	0.0120	0.0426
Log of after-tax equivalised income	6.6868	6.5860	-1.1850	-1.2158	-1.3301
Labour force status (ref: Not in the labour force)					
Labour force status: Employed	1.5438	2.1049	0.6945	0.4918	0.6256
Labour force status: Unemployed	0.0770	0.0348	0.0422	0.0210	0.0054
Education (ref: Year 12 or below)					
Education: Certificates and diploma	-0.3232	-0.0094	-0.0311	-0.0704	0.0011
Education: Undergraduate	-0.3470	-0.2936	-0.3169	-0.1406	0.0053
Education: Postgraduate	-0.0734	0.0312	0.0097	0.0173	0.0260
Death of spouse or child	0.0480	0.0462	0.0262	0.0202	0.0170
Death of close relative or family member	0.1346	0.0836	0.1117	0.0694	0.0541
Serious injury or illness to family member	0.3504	0.1842	0.1783	0.1461	0.1111
Natural disaster damaged/destroyed home	0.0039	-0.0139	0.0046	0.0074	-0.0005
Fired or made redundant	0.0525	0.0134	0.0339	0.0069	-0.0094
Victim of physical violence	0.1506	0.0844	0.0427	0.0231	0.0041
Close family member detained in jail	0.0540	0.0190	0.0222	0.0126	0.0042
Detained in jail	-0.0588	-0.0388	-0.0083	-0.0020	-0.0029
Residential area index of socioeconomic advantage (ref: 1st quintile)					
2nd quintile	-0.1460	-0.0667	-0.2222	-0.1253	-0.0100
3rd quintile	-0.1632	-0.0192	-0.2252	-0.1577	-0.0560
4th quintile	-0.0313	0.1587	-0.1998	-0.1031	0.0476
5th quintile	-0.2534	-0.0171	-0.3259	-0.2205	-0.1043
Constant	4.3086	4.1017	8.9310	6.8729	5.1396
Average RIF (Men - Women)	3.8595	3.4585	2.7337	1.7595	0.8763

Q10, Q25, Q50, Q75, and Q90 are the 10th, 25th, 50th, 75th and 90th quantiles. Numbers in the table are computed using Equation (3.4).

For people with very low mental health score (in the 10th quantile), demographic attributes such as marital status, age, having small children or being born overseas contribute to reducing the gender-based mental health gap. Among economic factors, income and being employed significantly widen the gender-based mental health gap, while education helps to close it. More specifically, we could close the mental health gap in the 10th quantile if we could equalise income, all else equal.

In the higher quantiles, marital status contributes to increasing the gender-based mental health gap. This result is consistent with the literature on marital status and personal wellbeing, which shows that marriage is more advantageous to men than women. Women tend to provide an emotional safety net and other support to the men with whom they are in a relationship (Coombs, 1991).

A larger household size appears to be associated with a larger gender mental health gap between men and women and this result is persistent at all quantiles. One possible explanation is that women are more likely to contribute to housework and as home production is an increasing function of family size, women in large households carry out a greater share of household tasks. This result adds to the literature on the relationship between family size and the housework gender gap which shows that growing up in a large family makes girls more likely to contribute to housework while having no impact on boys, resulting in a wider housework gender gap as adults (Menta and Lepinteur, 2021).

All the negative life events are associated with increasing gender-based mental health gap, and this result holds across quantiles, even though their contribution to the gap decreases as we move to higher quantiles. Among negative life events, serious injury or illness to a family member contributes most to the mental health gap between men and women. This is probably due to the fact that women often take on the role of carers after such events, regardless of whether their own or their spouse's family members are affected. If the women facing such situations also work and have egalitarian ideologies, then they are likely to perceive the additional responsibility as an unfair division of labour, resulting in reduced life satisfaction (Flèche et al., 2020). However, the contribution of all these negative life events to the gender mental health gap is relatively small compared to the contribution of economic factors such as income. Although the impacts of some life shocks are extremely severe at the individual level, these types of events are infrequent for both men and women, meaning they contribute only marginally to the gender-based mental health gap.

We find that apart from income, differences in the age distribution plays an important role in the gender mental health gap, possibly because men and women have different paths through life and their trajectories through life are different. The contemporary literature on mental health inequality includes similar findings (e.g., Hashmi et al., 2020; Veisani and Delpisheh,

2015) and attributes this fact to the ageing or retirement effect. It is hypothesised that as people get older, they become more psychologically stable, thus men's and women's mental health states tend to converge. The constant in Table 3.4.3 accounts for common underlying risk factors that are unrelated to other variables examined in our studies, such as biological differences,[8] or differences in reporting or responding to mental illness. The decomposition results using Equation (3.5), which can be provided upon request, tell the same story. Overall, the gender mental health gap decreases as one moves from the lower quantiles to the higher quantiles of the distribution, as shown in the last row of Table 3.4.3.

3.5 Discussion

This study examines the gender-based mental health gap in Australia using a regression-based decomposition technique based on the RIF. We show that at the individual level, negative life events such as being a victim of physical violence have severe consequences on mental health scores for both men and women. Although the magnitude of the effect differs significantly between men and women, the rareness of negative life events for both genders translates into a relatively small contribution to gender inequality at the aggregate level. Among negative life events, serious injury or illness to family members is the one that contributes most to the mental health gap between men and women, possibly due to the fact that women often take on the carer's role. Inequalities in economic outcomes remain the major factors in driving the gender-based mental health gap at the low end.

Our study has several limitations. First, we do not aim to establish causal effects but focus on understanding the statistical correlations between mental health scores and explanatory factors. As such, our results are vulnerable to reverse-causal effects (i.e., endogeneity); for example, people with very poor mental health are unable to engage in productive work and thus, have lower income. Despite the descriptive nature of the method, we contribute to the understanding of inequality by identifying factors that are associated with women being on the low end of the mental health score distribution and factors that drive the gender-based mental health gap. If women who experience physical violence also have very poor mental health, policies and social interventions can be initiated to address these issues holistically.

Because few participants experienced major life events such as physical violence, focusing on causality at the population level using self-reported surveys can lead to erroneous conclusions and encourage ineffective policies. For example, a series of major life events can be very damaging for one's mental health, sending someone with relatively good and stable mental health toward the bottom end of the distribution; on the other hand, people in chronically poor mental health might be more likely to encounter negative life events (e.g., physical violence,

[8]HILDA does not collect variables such as blood group, heart rate or blood pressure.

being with partners with poor mental health and challenging lifestyles, having negative impacts on their children's cognitive and non-cognitive skills (Menta et al., 2021)). Interventions for the former situation are likely to be different from the latter. Healthcare interventions such as medication, counselling and psychotherapies are likely to be more effective for people who normally enjoy good mental health because the situation is acute; social interventions combined with ongoing healthcare programs that provide regular support and empowerment are likely to be more effective for people with chronically poor mental health.

As mentioned in Section 3.2, we acknowledge that it would be useful to know whether an individual suffers from "repeated" traumatic experiences, but are unable to incorporate that information into the current analysis. Repeated traumatic events such as physical violence requires further exploration, and often more intervention. Poor mental health might be the result of a traumatic shock, or it might be a consequence of repeated smaller events over times. The interventions for these two scenarios are different. There is little we can do to prevent a traumatic shock. In contrast, repeated experiences can be prevented from happening, especially when they stem from a systematic family problem.

Another issue, which is intrinsic to all self-report surveys like the HILDA, is that men and women might interpret and report mental health differently. Existing studies note that although surveys indicate men have a significantly lower prevalence of mental disorders such as anxiety and depression than women, probably due to men being much less likely to report their problems for fear of stigma or to do with dominant notions of masculinity (Affleck et al., 2018). The discrepancy between men's lower rates of depression and higher rates of suicide than women might signal "masked depression". Nonetheless, the MHI-5 score remains an effective instrument for detecting mental health issues in the general population (Rumpf et al., 2001; Thorsen et al., 2013; Hoeymans et al., 2004), and our results using the MHI-5 remains valid.

Another issue could be that a victim of violence does not report the incident because of shame or the psychology of being abused. Some victims think that they deserve to be treated that way and do not consider themselves victims. Some fear of even worse violence from their partners, fear for their children's safety or fear that they are unable to survive financially without the relationship. These issues are worthy of consideration for future research.

3.6 Conclusion

The research presented in this chapter analysed gender inequality in mental health scores in Australia using data from the HILDA longitudinal survey. We observed that men have significantly better mean mental health scores, and lower variances, such that the left tail of the combined mental health distribution is disproportionately female. We asked whether economic disparity or negative life events are sufficient to explain the observed gap. Because we care about people with poor mental health, the conventional regression techniques that focus on explaining the mean would not be meaningful. Therefore, we used the UQR technique and decomposition based on RIF to better understand the extent our results vary along the mental health distribution, with a special focus on the low end of the distribution.

We found that disparities in income account for most of the gender gap amongst individuals with very poor psychological wellbeing. However, education, as an indicator of socioeconomic status, partly closes the gap. Other unobserved factors, such as differences in biology, may be significant predictors of the gender mental health gap.

We also examined the mental health effects of various negative life events, such as the death of a family member, serious injury or illness to family members, and being a victim of violence. At the individual level, these variables have large effect sizes but are not correlated strongly enough with gender to explain the gender mental health disparities at the aggregate level. In spite of the small role of negative life events in the gender mental health gap, our results suggest that social policy that supports individuals undergoing negative life events would be helpful in improving the lives of people with very poor psychological wellbeing.

Chapter 4

The Lexicographic Serial Rule

4.1 Introduction

The resource allocation problem of indivisible objects is one of the fundamental issues in economics. Various real-life problems require distributing indivisible objects to a set of individuals in situations in which monetary transfers are not allowed. Examples includes the assignment of teaching responsibilities to faculty members, the provision of social housing, and the allocation of doctors to hospitals, just to name a few.

In general, we are concerned with assigning a set of objects A to a set of agents I: objects do not care to whom they are allocated, but we want the allocation to be responsive to the agents' preferences. An assignment problem can be deterministic or random. A *deterministic assignment* is a one-to-one mapping from I into A, which can be represented by a permutation matrix, whereas a *random assignment* is the probability distribution over deterministic assignments. Because the classical Birkhoff-von Neumann theorem implies that any random assignment can be induced by a probabilistic distribution over deterministic assignments (Budish et al., 2011), we focus exclusively on random assignments.

A random assignment can be represented by a matrix $p = (p_{ia})$ with rows p_i standing for agents, columns p_a standing for objects, and entries in $[0, 1]$ such that $\sum_{a \in A} p_{ia} \leq 1$ for all $i \in I$ and $\sum_{i \in I} p_{ia} \leq 1$ for all $a \in A$. Here p_{ia} is the probability that agent $i \in I$ receives object $a \in A$. This formal structure can also be interpreted as the deterministic allocation of infinitely divisible goods when each agent is entitled to a total of at most one unit (Bogomolnaia, 2015).

An agent may have any strict preference over objects and not receiving an object. In this chapter, we focus on *lexicographic preferences* in which an agent prefers any amount of one good (e.g., x) to any amount of another good (e.g., y). Given some bundles of goods, an agent with lexicographic preferences will choose the bundle that offers the most x regardless of the amount of y; only when the amounts of x are equal, would the agent prefer the bundle with more y.[1] An *assignment rule* is a function mapping each profile of preferences to a random assignment.

The oldest assignment rule used throughout recorded history is *random priority*, in which an equiprobable lottery chooses a first agent, who receives her favourite object, if there is one she prefers to receiving no object. After that, another equiprobable lottery chooses a second agent, who receives her favourite among the objects not chosen by the first agent if one of them is preferred to receiving no object, and so forth. For example, consider a situation where we have three agents 1, 2, 3 and three objects a, b, c that need to be allocated to agents such that each agent receives one object. Suppose the preference profile is

$$a \succ_1 b \succ_1 c, \quad a \succ_2 c \succ_2 b, \quad b \succ_3 c \succ_3 a.$$

In this case, if agent 1 is chosen first (which happens with probability 1/3), then it does not matter who is chosen second; agent 1 will select object a, agent 2 will select object c, and agent 3 will select object b (i.e., $\frac{1}{3}(a \to 1, c \to 2, b \to 3)$). If agent 1 chooses last, then it does not matter who chooses first; object a will be allocated to agent 2, object b will be allocated to agent 3 and object a will be left for agent 1 (i.e., $\frac{1}{3}(c \to 1, a \to 2, b \to 3)$). If agent 2 is chosen first, followed by agent 1 second, then we will have $\frac{1}{6}(b \to 1, a \to 2, c \to 3)$. And if agent 3 is chosen first, followed by agent 1 second, then we will have $\frac{1}{6}(a \to 1, c \to 2, b \to 3)$. Random priority procedure gives

$$\frac{1}{3}(a \to 1, c \to 2, b \to 3) + \frac{1}{3}(c \to 1, a \to 2, b \to 3)$$

$$+\frac{1}{6}(b \to 1, a \to 2, c \to 3) + \frac{1}{6}(a \to 1, c \to 2, b \to 3).$$

The matrix of assignment probabilities will be:

$$\begin{array}{cccc} & a & b & c \\ 1 & \frac{1}{2} & \frac{1}{6} & \frac{1}{3} \\ 2 & \frac{1}{2} & 0 & \frac{1}{2} \\ 3 & 0 & \frac{5}{6} & \frac{1}{6} \end{array}$$

Bogomolnaia and Moulin (2001) point out that random priority may be inefficient in the sense that there can be another random assignment that stochastically dominates the one produced by random priority for all agents, and strictly for some. For example, consider four agents and four objects with the following preference profile:

$$1 \text{ and } 2 : a \succ b \succ c \succ d, \quad 3 \text{ and } 4 : b \succ a \succ d \succ c.$$

[1]See Appendix C.1 for a brief discussion of lexicographic preferences.

Random priority generates the assignment probability matrix:

$$\begin{array}{ccccc} & a & b & c & d \\ 1 & \frac{5}{12} & \frac{1}{12} & \frac{5}{12} & \frac{1}{12} \\ 2 & \frac{5}{12} & \frac{1}{12} & \frac{5}{12} & \frac{1}{12} \\ 3 & \frac{1}{12} & \frac{5}{12} & \frac{1}{12} & \frac{5}{12} \\ 4 & \frac{1}{12} & \frac{5}{12} & \frac{1}{12} & \frac{5}{12} \end{array}$$

However,

$$\begin{array}{ccccc} & a & b & c & d \\ 1 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 2 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 3 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 4 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{array}$$

is strictly better for everyone.

They then propose a different assignment rule called the *(probabilistic) serial rule* that uses a *simultaneous eating* algorithm. The objects are thought of as "cakes" of probability of unit size. Over the unit interval of time, the agents "eat" cake at unit speed, at each point in time consuming the favourite object whose cake has not yet been exhausted if there is one that is preferred to receiving no object. For example, returning to the preference profile

$$a \succ_1 b \succ_1 c, \quad a \succ_2 c \succ_2 b, \quad b \succ_3 c \succ_3 a.$$

then agent 1 and 2 will begin eating cake a and agent 3 begins eating cake b. At the time $\frac{1}{2}$ when cake a is exhausted, agent 1 switches to eating cake b (which is her second favourite) together with agent 3, while agent 2 switches to eating cake c (which is the second favourite to agent 2). After time $\frac{3}{4}$ when cake b is exhausted, all agents eat cake c. Hence, the probabilistic serial rule gives the assignment matrix:

$$\begin{array}{cccc} & a & b & c \\ 1 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 2 & \frac{1}{2} & 0 & \frac{1}{2} \\ 3 & 0 & \frac{3}{4} & \frac{1}{4} \end{array}$$

When Bogomolnaia and Moulin (2001) was published, the serial rule was the only alternative to random priority that depends on the agents' ordinal preferences. The well known market mechanism proposed by Hylland and Zeckhauser (1979) depends on agents' cardinal preferences, and even for that domain it is not, strictly speaking, an allocation rule, because it is possible that the market has multiple equilibria. It is difficult to elicit vNM utilities, first of all because untrained agents often do not know what they are. Zhou (1990) shows that there is no efficient and equitable (in the weak sense of equal treatment of equals) assignment rule based on cardinal preferences that is strategy-proof. Finally, it can be quite difficult for agents to understand how their declaration of a vNM utility affects what they receive. Thus, ordinal mechanisms have important practical advantages.

The purpose of this chapter is to propose another assignment rule that depends only on ordinal preferences. The *lexicographic (serial) rule* is tailored for situations in which each agent's primary concern is to maximise the probability of receiving her favourite object; she cares much more about the difference between her second and third favourite than the difference between her third and fourth favourite, and so forth. Like the serial rule, it can be described as an eating procedure, which is divided into rounds. In the first round each agent eats her favourite cake until all such cakes have been exhausted. In the k^{th} round, for $k = 2, 3, \ldots$, the agents who are not yet sated, and whose k^{th} favourite objects have not yet been fully allocated, eat the corresponding cakes. For each such object, among those agents for whom it is the k^{th} favourite, at first only the agents who received the least total probability in previous rounds are allowed to eat. Other agents for whom this object is the k^{th} favourite join in when the total probability assigned to those currently eating equals their total prior probability, so all agents who eat a positive amount of this cake in the k^{th} round end the k^{th} round with the same total probability, and all agents for whom this object is the k^{th} favourite who do not eat it in round k have greater total prior probability. Recall the preference profile:

$$a \succ_1 b \succ_1 c, \quad a \succ_2 c \succ_2 b, \quad b \succ_3 c \succ_3 a.$$

With the lexicographic serial rule, agent 1 and 2 begin eating cake a and agent 3 begins eating cake b until these objects are gone. In the second round, agent 3 is sated, agent 2 eats cake c until he is sated, while the second favourite of agent 1 (which is b) is no longer available, so agent 1 eats nothing. In the third round, agent 1 eats the rest of c. Hence, the lexicographic serial rule gives the assignment matrix:

$$\begin{array}{cccc} & a & b & c \\ 1 & \frac{1}{2} & 0 & \frac{1}{2} \\ 2 & \frac{1}{2} & 0 & \frac{1}{2} \\ 3 & 0 & 1 & 0 \end{array}$$

We now discuss efficiency. For a given strict preference $\succ_i$ over the objects and receiving no object, a probability measure p_i' on this set is *strictly lex preferred* to another probability measure p_i if, for some k, the two measures give the same probability of receiving the ℓ^{th} best object for all $\ell < k$ and p_i' gives a higher probability of the k^{th} object than p_i. We say that p_i' is *(weakly) lex preferred* to p_i if it is strictly lex preferred or $p_i = p_i'$. A random allocation p is *lex efficient* if there does not exist another random allocation p' such that for all i, p_i' is lex preferred to p_i, with strict lex preference for some i.

We say that the *rank* of an object for an agent is k if the object is her k^{th} favourite, and we say that i ranks the object at least as highly as j (strictly higher than j) if its rank for i is not greater than (less than) its rank for j. An allocation is *lex fair* if, for all agents i and j and objects a that both i and j prefer to receiving no object, if $p_{ia} > 0$, then: a) if j ranks a more highly than i, or if the ranks of a for i and j are the same and $p_{ja} = 0$, then the probability that j receives an object that is better than a is one; b) if the two agents rank the object equally and $p_{ja} > 0$, then the two agents have the same total probability assigned to their k top objects. Our main result is that for each profile there is a unique random allocation that is lex efficient and lex fair, which is the random allocation given by the lex serial rule.

For a given preference $\succ_i$, p'_i is *strictly sd preferred* to p_i if, for all k, p'_i gives at least as much probability as p_i to the first k favourite objects, with strict inequality for some k.[2] We say that p'_i is *(weakly) sd preferred* to p_i if it is strictly sd preferred or $p_i = p'_i$. A random allocation p is *sd efficient* if there does not exist another random allocation p' such that for all i, p'_i is sd preferred to p_i, with strict sd preference for some i. The serial rule gives sd efficient random allocations.

If p'_i is (strictly) sd preferred to p_i, then it is (strictly) lex preferred to p_i, so if p is lex efficient, then it is also sd efficient. McLennan (2002) shows that an allocation is sd efficient if and only if it maximises the sum of utilities for some vNM utility functions consistent with the ordinal preferences. Thus each outcome produced by the serial rule maximises some Bergson-Samuelson social welfare function consistent with the ordinal preferences, but the utility function of an agent may vary from one profile to the next even if that agent's ordinal preferences do not change. A random allocation is lex efficient if and only if, for sufficiently small $\varepsilon > 0$, it maximises the sum over agents and k of ε^k times the probability that the agent receives her k^{th} favourite, so the lex rule has a Bergson-Samuelson welfarist interpretation that is quite clear, though also quite narrow insofar as it is restricted to domains in which the sorts of preferences for which it is appropriate can be (at least approximately) expected. Bogomolnaia (2015) shows that the serial rule is the unique allocation rule that always maximises the minimum probability that any agent receives her favourite, that maximises (within the space of random allocations satisfying this condition) the minimum probability that any agent receives one of her top two objects, and so forth. She views this as a Rawlsian welfarist interpretation of the serial rule.

In the Boston mechanism for school choice (e.g., Kojima and Ünver (2014)) each student submits her ordinal ranking of schools, and the schools have priorities that are, in effect, weak preference orderings of the students, which include the possibility that a student may not be qualified to attend the school. Each school first considers the students for which the school is the first choice, accepting those with the highest priorities up to the point where all qualified

[2]Here "sd" is an abbreviation of "first order stochastic dominance". For more details, please see Appendix C.2

students have been accepted or all seats at the school have been allocated. Each student who was not accepted at her favourite school is then considered at her second favourite school, which again accepts those with the highest priorities up to the point where all qualified students have been accepted or all of the school's seats have been filled, and so forth. If, at any stage, the school is indifferent between a number of students that is greater than the number of remaining seats, then the accepted students are determined randomly.

The problem considered here can be viewed as the special case of the Boston mechanism in which the schools are completely indifferent between all students. The lex serial rule points to a possibility that does not seem to have been considered in the related school choice literature: in randomising over the students of minimum priority for the remaining seats at a school, one may take into account the students' probabilities of being accepted at a school they prefer. For example, suppose that i and j have both been rejected at their respective favourite schools, and are now competing for a seat at their common second favourite. If i had a high probability of acceptance at her favourite, but the probability of acceptance at j's favourite was quite low, is it really "fair" that, conditional on both being rejected by their favourite, they should have equal probability of acceptance at the second favourite?

Lexicographic preferences are in a certain sense approximated by the cardinal utility function in which the most preferred object yields utility 1, the second most preferred object yields utility ϵ, the third most preferred object yields utility ϵ^2, and so forth. A comparison between the lexicographic serial rule and the limits as $\epsilon \to 0$ of the allocations produced by the Hylland-Zeckhauser procedure can be found in Appendix C.3.

The chapter is organised as follows. Section 4.2 presents the basic model and defines the lexicographic rule. Section 4.3 defines the axioms that the lexicographic rule satisfy and proves them. It also discusses other axioms which have been defined in the literature to see if the lexicographic rule satisfies them or not. Section 4.5 presents motivating examples and compares the lexicographic rule with the serial rule. The last section concludes.

4.2 The Lexicographic Serial Rule

In this section we study a form of the problem that is more general than is customary in this literature, insofar as the agents have appetites and the objects have quotas, which need not be integers. Since we restrict attention to strict preferences, allowing the objects to have integer quotas is required by applications such as school choice. Allowing the agents to have different appetites is in line with the alternative interpretation suggested by Bogomolnaia (2015), in which the objects are perfectly divisible. Integer quotas are studied in Hashimoto et al. (2014), and integer quotas and appetites are studied in Heo (2014). We will see that this additional generality entails no additional complexity, and in fact we have not found any arguments in either of these papers that are not correct, without any modification, if the assumption that the appetites and quotas are integers is dropped. At the same time, while there are natural applications such as course allocation for integer appetites and capacities, applications with non-integer appetites and quotas are not evident, so the additional generality seems to be primarily a comment on the underlying mathematics.

An *economy* is a tuple (I, A, α, q) where I is a finite set of *agents* with $|I| = n$,[3] A is a finite set of *objects* with $|A| = m$, $\alpha \in \Re^I_{++}$ is a vector of positive *appetites*, and $q \in \Re^A_{++}$ is a vector of positive *quotas*. We always assume that $\sum_a q_a \geq \sum_i \alpha_i$. For problems with total appetite greater than total quota, or with agents who prefer receiving no object to some of the objects, we can add an additional artificial object with infinite capacity that represents not receiving an object.

Let a strict ordering $\succ_i$ over A be individual i's (strict) preference. As usual, we write $a \succsim_i a'$ to indicate that either $a \succ_i a'$ or $a \sim_i a'$. Let $\mathcal{P}$ be the set of preferences. For $\succ_i \in \mathcal{P}$ and $a \in A$ let $U(\succ_i, a) = \{\, b \in A : b \succsim_i a \,\}$ be the weak upper contour set of a, i.e., set of all objects that are at least as good as a to individual i. For $\succ_i \in \mathcal{P}$, $a \in A$, and $\pi_i \in \Re^A_+$ let

$$F(\succ_i, a, \pi_i) = \sum_{b \in U(\succ_i, a)} \pi_{ib}$$

be i's consumption of objects that are at least as good as a for $\succ_i$. The *rank* $r_a(\succ_i)$ of a in $\succ_i$ is $|U(\succ_i, a)|$. For $k \in A$, let $f_k(\succ_i)$ be the k^{th} favourite of i; this may be defined implicitly by requiring that $r_{f_k(\succ_i)}(\succ_i) = k$.

A *preference profile* is an n-tuple $\succ = (\succ_i)_{i \in I} \in \mathcal{P}^I$. For the remainder of this section we work with a fixed preference profile $\succ$. For an object a and an integer $k = 1, \ldots, m$ let $I(a, k, \succ) = \{\, i \in I : f_k(\succ_i) = k \,\}$ be the set of agents for whom a is the k^{th} favourite.

For a matrix $r = (r_{ia})_{i \in I, a \in A}$ with nonnegative entries let $c_i(r) = \sum_a r_{ia}$ be agent i's total consumption at r, and let $d_a(r) = \sum_i r_{ia}$ be the total consumption of a. We say that r is a *partial allocation* if $c_i(r) \leq \alpha_i$ for all i and $d_a(r) \leq q_a$ for all $a \in A$. We say that agent i is *sated* (*unsated*) at r if $c_i(r) = \alpha_i$ $(c_i(r) < \alpha_i)$, and we say that object a is *exhausted* (*available*) at r if $d_a(r) = q_a (d_a(r) < q_a)$. A partial allocation is an *allocation* if every agent is sated. For a partial allocation r, r_i denotes the i-row of r.

For a partial allocation r and an integer $k = 1, \ldots, m$ we define a new partial allocation $\rho(r, k, \succ)$ that is obtained from r by allocating remaining amounts of k^{th} favourites. Consider an object a. If a is exhausted at r, or if every element of $I(a, k, \succ)$ is sated at r, let $t(r, a, k, \succ) = 0$. Otherwise we define $t(r, a, k, \succ)$ implicitly by letting $t(r, a, k, \succ)$ be the smallest number such

[3] For an arbitrary finite set X, $|X|$ denotes its cardinality, i.e., number of elements in the set X.

that

$$\sum_{i\in I(a,k,\succ)} \max\big\{\min\{t(r,a,k,\succ),\alpha_i\}-c_i(r),0\big\} = \min\Big\{\underbrace{\sum_{i\in I(a,k,\succ)} \alpha_i - c_i(r)}_{\text{total remaining appetite}},\ \underbrace{q_a - d_a(r)}_{\text{remaining of } a}\Big\}.$$

The left hand side is a continuous increasing function of $t(r,a,k,\succ)$; it is 0 when $t(r,a,k,\succ)=0$, and it is $\sum_{i\in I(a,k,\succ)}\alpha_i - c_i(r)$ when $t(r,a,k,\succ)\geq \max_{i\in I(a,k,\succ)}\alpha_i$, so $t(r,a,k,\succ)$ is well defined and positive. Intuitively, $t(r,a,k,\succ)$ is the number such that allocating remaining k so as to bring the total consumption of each $i\in I(a,k,\succ)$ up to $\max\big\{\min\{t(r,a,k,\succ),\alpha_i\},c_i(r)\big\}$ either exhausts the remaining supply of a or sates everyone in $I(a,k,\succ)$. We define the partial allocation $\rho(r,k,\succ)$ by setting

$$\rho_{ia}(r,k,\succ) = \begin{cases} r_{ia}, & i\notin I(a,k,\succ),\\ r_{ia} + \max\big\{\min\{t(r,a,k,\succ),\alpha_i\}-c_i(r),0\big\}, & i\in I(a,k,\succ),\end{cases}$$

for all $i\in I$ and $a\in A$. Note that $\rho_{ia}(r,k,\succ)=r_{ia}$ if a is unavailable at r or i is sated at r.

We can now describe the lexicographic serial rule inductively. Let $\varrho^0(\succ)$ be the $n\times m$ matrix whose entries are all zero. For $k=1,\ldots,m$, let $\varrho^k(\succ)=\rho(\varrho^{k-1}(\succ),k,\succ)$. Let $\lambda(\succ)=\varrho^m(\succ)$. The *lexicographic serial rule* (lex rule) is the allocation rule that assigns $\lambda(\succ)$ to each $\succ\in\mathcal{P}^I$.

4.3 Properties of the Lexicographic Serial Rule

In this section, we discuss whether the lexicographic serial rule satisfies some important properties. We start with some new properties related to lexicographic preferences. Then, we check whether the lexicographic serial rule satisfies some other properties that have been discussed in the literature. First, we note a few observations related to the lexicographic serial rule that will be important in proving the properties that follow.

Lemma 1. *For each $i\in I$ and $k=1,\ldots,m$, if $f_k(\succ_i)=a$, then*

$$\varrho^\ell_{ia}(\succ) = \begin{cases} 0, & \ell=0,\ldots,k-1,\\ \lambda_{ia}(\succ), & \ell=k,\ldots,m,\end{cases}$$

Proof. From the definition, if $\ell<k$, then $i\notin I(a,k,\succ)$. Thus, $\varrho^\ell_{ia}(\succ)=\varrho^{\ell-1}_{ia}(\succ)=\cdots=0$. If $\ell\geq k$, then $i\in I(a,k,\succ)$. Thus, $\varrho^\ell_{ia}(\succ)=\varrho^{\ell-1}_{ia}(\succ)+\max\big\{\min\{t(\varrho^{\ell-1}_{ia}(\succ),a,k,\succ),\alpha_i\}-c_i(\varrho^{\ell-1}_{ia}(\succ)),0\big\}=\lambda_{ia}(\succ)$ □

Lemma 2. *For each $i\in I$ and $k=1,\ldots,m$, if $F(\succ_i,a,\lambda_i(\succ))<\alpha_i$, then $d_a(\varrho^k(\succ))=q_a$.*

Proof. When $F(\succ_i, a, \lambda_i(\succ)) < \alpha_i$, the consumption of objects that are at least as good as a for individual i is still less than i's appetite. Therefore, object a must have been exhausted at the k^{th} favourite. $\square$

Lemma 3. *For each $i \in I$ and $k = 1, \ldots, m$, if $F(\succ_i, a, \lambda_i(\succ)) < \alpha_i$ and $\lambda_{ia}(\succ) = 0$, then $d_a(\varrho^{k-1}(\succ)) = q_a$.*

Proof. When $F(\succ_i, a, \lambda_i(\succ)) < \alpha_i$, again, the consumption of objects that are at least as good as a for individual i is still less than i's appetite. However, at the k^{th} round, there is no object a left to be consumed by individual i. Thus, object a must have been consumed entirely in the $(k-1)^{\text{th}}$ round. $\square$

4.3.1 Lex Efficient

For $\succ_i \in \mathcal{P}$ let $\succ_i^{lex}$ denote the induced lexicographic ordering of $\Re_+^A$. For $\pi_i, \pi_i' \in \Re_+^A$, $\pi_i' \succ^{lex} \pi_i$ if, for some $k = 1, \ldots, m$, $\pi'_{if_\ell(\succ_i)} = \pi_{if_\ell(\succ_i)}$ for all $\ell < k$ and $\pi'_{if_k(\succ_i)} > \pi_{if_k(\succ_i)}$. We write $\pi_i' \succeq^{lex} \pi_i$ to indicate either that $\pi_i' \succ^{lex} \pi_i$ or that there is some $k = 1, \ldots, m$ such that $\pi'_{if_\ell(\succ_i)} = \pi_{if_\ell(\succ_i)}$ for all $\ell < k$, $\pi'_{if_k(\succ_i)} \leq \pi_{if_k(\succ_i)}$, and $F(\succ_i, f_k(\succ_i), \pi_i') = \alpha_i$. If $\sum_a \pi'_{ia} = \sum_a \pi_{ia}$, then these conditions are simpler and more intuitive: $\pi_i' \succ^{lex} \pi_i$ if, for some $k = 1, \ldots, m-1$, $\pi'_{if_\ell(\succ_i)} = \pi_{if_\ell(\succ_i)}$ for all $\ell < k$ and $\pi'_{if_k(\succ_i)} > \pi_{if_k(\succ_i)}$, and $\pi_i' \succeq^{lex} \pi_i$ if $\pi_i' \succ^{lex} \pi_i$ or $\pi_i' = \pi_i$.

Definition 1. *An allocation p'* (weakly) lex Pareto dominates *an allocation p for $\succ$ if $p_i' \succeq_i^{lex} p_i$ for every i and $p_i' \succ_i^{lex} p_i$ for some i. An allocation p is* lex efficient *for $\succ$ if it is not weakly lex Pareto dominated for $\succ$.*

Lemma 4. *$\lambda(\succ)$ is lex efficient for $\succ$.*

Proof. Let $p = \lambda(\succ)$. By way of contradiction, suppose that p is lex Pareto dominated for $\succ$ by p'. Let k be the smallest integer such that $p_{if_k(\succ_i)} < p'_{if_k(\succ_i)}$ for some i. Fixing such an i, let $a = f_k(\succ_i)$. For all $j \in I$ and $\ell < k$ the definition of k gives $p'_{jf_\ell(\succ_j)} \leq p_{jf_\ell(\succ_j)}$, and $p'_{jf_\ell(\succ_j)} \geq p_{jf_\ell(\succ_j)}$ because p' lex Pareto dominates p, so $p'_{jf_\ell(\succ_j)} = p_{jf_\ell(\succ_j)}$. In particular $F(\succ_i, a, p_i) < F(\succ_i, a, p_i') \leq \alpha_i$, so $d_a(p) = d_a(\varrho^k(\succ)) = q_a$. Therefore there is some $j \in I(a, k, \succ)$ such that $p'_{ja} < p_{ja}$ and thus $p_j \succ_j^{lex} p_j'$, contrary to assumption. $\square$

Definition 2. *An allocation p is* lex unbiased *for $\succ$ if, for all agents i, j and objects a such that $r_a(\succ_i) > r_a(\succ_j)$, if $p_{ia} > 0$, then $F(\succ_j, a, p_j) = \alpha_j$.*

Lemma 5. *$\lambda(\succ)$ is lex unbiased.*

Proof. Consider i, j, and a such that $k = r_a(\succ_i) > \ell = r_a(\succ_j)$. If $\lambda_{ia}(\succ) > 0$, then $d_a(\varrho^\ell(\succ)) < q_a$, and consequently $F(\succ_j, a, \lambda_j(\succ)) = \alpha_j$. $\square$

4.3.2 Lex Fair

Definition 3. *A allocation p is* lex fair *for $\succ$ if, for all agents i, j and objects a, if $r_a(\succ_i) = r_a(\succ_j)$ and $p_{ia} > 0$, then $F(\succ_j, a, p_j) \geq \min\{F(\succ_i, a, p_i), \alpha_j\}$.*

A random allocation is fair if at any time, for any object that does not make any agent full, whenever an agent is assigned this object with positive probability, his total probability shares of all objects that are weakly preferred to this object is no less than that of any other agent. If an agent becomes sated by eating part of an object and leaving part of it for other agents to consume in later rounds, then the above notion of fairness should hold when excluding this object.

Lemma 6. *$\lambda(\succ)$ is lex fair.*

Proof. Let $p = \lambda(\succ)$, and fix i, j, and a such that $r_a(\succ_i) = r_a(\succ_j) = k$ and $p_{ia} > 0$. If $p_{ja} = 0$, then either j is sated in $\varrho^{k-1}(\succ)$ and $F(\succ_j, a, p_j) = \alpha_j$, or $c_j(\varrho^{k-1}(\succ)) \geq c_i(\varrho^k(\succ))$ and thus $F(\succ_j, a, p_j) \geq F(\succ_j, a, p_i)$. If $p_{j,a} > 0$, then

$$F(\succ_j, a, p_j) = c_j(\varrho^k(\succ)) \geq \min\{c_i(\varrho^k(\succ)), \alpha_j\} = \min\{F(\succ_i, a, p_i), \alpha_j\}.$$

□

4.3.3 Lex Envy-free

The basic idea of envy freeness is that each agent prefers her own bundle to anyone else's bundle.

Definition 4. *A allocation p is* lex envy-free *for $\succ$ if $p_i \succeq_i^{lex} p_j$ for all $i, j \in I$.*

Lemma 7. *$\lambda(\succ)$ is lex envy-free.*

Proof. Fixing distinct $i, j \in I$, we will show that $\rho_i^k(\succ) \succeq_i^{lex} \rho_j^k(\succ)$ for all $k = 0, \ldots, m$. This is trivially true for $k = 0$, so by induction it suffices to show for a given k that $\rho_i^{k-1}(\succ) \succeq_i^{lex} \rho_j^{k-1}(\succ)$ implies $\rho_i^k(\succ) \succeq_i^{lex} \rho_j^k(\succ)$. Clearly $\rho_i^{k-1}(\succ) \succ_i^{lex} \rho_j^{k-1}(\succ)$ implies $\rho_i^k(\succ) \succ_i^{lex} \rho_j^k(\succ)$, so we may suppose that $\rho_i^{k-1}(\succ) = \rho_j^{k-1}(\succ)$. Let $a = f_k(\succ_i)$. Note that $\rho_{ja}^{k-1}(\succ) = 0$ because $\rho_{ia}^{k-1}(\succ) = 0$ (Lemma 1).

If $r_a(\succ_j) < k$, then $d_a(\rho^{r_a(\succ_j)-1}(\succ)) = q_a$ (Lemma 3) so $\rho_{ia}^k(\succ) = \rho_{ja}^k(\succ) = 0$ and thus $\rho_i^k(\succ) = \rho_j^k(\succ)$.

If $r_a(\succ_j) > k$, then $0 = \rho_{ja}^k(\succ) \leq \rho_{ia}^k(\succ)$ (Lemma 1) so $\rho_i^k(\succ) \succeq_i^{lex} \rho_j^k(\succ)$.

Suppose that $r_a(\succ_j) = k$. Let $F = F(\succ_i, a, \varrho_i^{k-1}(\succ)) = F(\succ_j, a, \varrho_j^{k-1}(\succ))$. If $F = \alpha_i$, then $\varrho_i^k(\succ) = \varrho_i^{k-1}(\succ) \succeq_i^{lex} \varrho_j^k$ holds automatically. If $F(\succ_i, a, \varrho_i^k(\succ)) = \alpha_i$, then $\varrho_i^k(\succ) \succeq_i^{lex} \varrho_j^k(\succ)$

holds regardless of whether $\varrho_{ja}^k(\succ)$ is larger of smaller than $\varrho_{ia}^k(\succ)$. Otherwise

$$\varrho_{ia}^k(\succ) = t(\varrho^{k-1}(\succ), a, k, \succ)) - F \geq \min\{t(\varrho^{k-1}(\succ), a, k, \succ)), \alpha_j\} - F = \rho_{ja}^k(\succ),$$

and again $\rho_i^k(\succ) \succeq_i^{lex} \rho_j^k(\succ)$. □

4.3.4 Lex Strategy Proofness

Definition 5. *An allocation rule p is* lex strategy-proof *if $p_i(\succ) \succeq_i^{lex} p_i(\succ_i', \succ_{-i})$ for all profiles $\succ$, all agents i, and all preferences $\succ_i'$.*

Lemma 8. *$\lambda(\succ)$ is not lex strategy-proof.*

Proof. Since $\succeq_i^{lex}$ is transitive it would suffice to show that $\lambda(\succ_i) \succeq_i^{lex} \lambda(\succ_i', \succ_{-i})$ when $\succ_i'$ differs from $\succ_i$ only insofar as it swaps the order of an adjacent pair a and b. Suppose that $a \succ_i b$, and that a is all gone by the time we get to round $r_a(\succ_i)$, but there is still some b available. By pretending to prefer b to a the agent gets some of it. An example of this could be a situation in which we have four agents 1, 2, 3 and 4, and four objects a, b, c, and d. The preference ordering of each agent is as follows:

Agent 1: $a \succ_1 b \succ_1 c \succ_1 d$

Agent 2: $b \succ_2 a \succ_2 c \succ_2 d$

Agent 3: $b \succ_3 a \succ_3 c \succ_3 d$

Agent 4: $b \succ_4 a \succ_4 c \succ_4 d$

If agent 2 reports truthfully, the resulting allocation from the lexicographic rule is:

	a	b	c	d
1	1	0	0	0
2	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
3	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
4	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

However, if agent 2 pretends to have a different preference ordering, the false preference profile is:

Agent 1: $a \succ_1 b \succ_1 c \succ_1 d$

Agent 2: $b \succ_2 c \succ_2 a \succ_2 d$

Agent 3: $b \succ_3 a \succ_3 c \succ_3 d$

Agent 4: $b \succ_4 a \succ_4 c \succ_4 d$

Then, the resulting allocation from the lexicographic rule is

	a	b	c	d
1	1	0	0	0
2	0	$\frac{1}{3}$	$\frac{2}{3}$	0
3	0	$\frac{1}{3}$	0	$\frac{1}{2}$
4	0	$\frac{1}{3}$	0	$\frac{1}{2}$

This new allocation is lex preferred to agent 2 compared to the allocation received by reporting truthfully. □

4.3.5 Individual Rationality

Usually individual rationality refers to an outside option that is always available to an agent, and it requires that she does not receive a positive share of any object that is worse than that. In our framework there is no designated outside option, but there may be one or more objects with inexhaustible supplies.

Definition 6. *An allocation rule p is individually rational if for all $\succ \in \mathcal{P}^I$, $i \in I$, and $a, b \in A$, if $q_a \geq \sum_i \alpha_i$ and $a \succ_i b$, then $p_{ib}(\succ) = 0$.*

Lemma 9. *The lexicographic rule $\lambda(\succ)$ is individually rational.*

Proof. The lexicographic rule allows each agent to consume the most favourite object available to them at each round. When $q_a \geq \sum_i \alpha_i$, it means that the supply of object a is sufficient to satisfy all agents' appetites. Therefore, agent i will not consume object b which is considered inferior to a for i. □

4.3.6 Sd-Efficiency

An allocation is sd efficient if there is no other allocation that weakly dominates the given one for all agents (Bogomolnaia and Heo, 2012). We say p_i stochastically dominates $p_i^{'}$, which we write $p_i \succ_i^{sd} p_i^{'}$, if for each $a \in A$, $\sum_{b \in U(\succ_i, a)} p_{ib} \geq \sum_{b \in U(\succ_i, a)} p_{ib}^{'}$.

Definition 7 (sd-efficient)**.** *(See Bogomolnaia and Moulin, 2001) An allocation p is sd-efficient if there does not exist another feasible assignment matrix $p^{'}$ such that for each $i \in I$, $p_i^{'} \succ_i^{sd} p_i$.*

Lemma 10. *The lexicographic procedure $\lambda(\succ)$ is sd-efficient.*

Proof. The lexicographic serial rule is lex efficient and lex efficiency implies sd efficiency. □

4.3.7 Invariance

Definition 8 (Bounded invariance). *(See Bogomolnaia and Moulin, 2001) For each* $\succ, \succ' \in \mathcal{P}^I$, *each* $i \in I$, *and each* $a \in A$, *if* $\succ_i |_{U(\succ_i, a)} = \succ'_i |_{U(\succ'_i, a)}$, *then for each* $j \in I$, $p_{ja}(\succ) = p_{ja}(\succ'_i, \succ_{-i})$.

Lemma 11. *The lexicographic rule* $\lambda(\succ)$ *satisfies bounded invariance.*

Proof. By the construction of the lexicographic procedure, this is direct. □

Definition 9 (Weakly invariant). *An allocation* p *is weakly invariant if for all* $\succ \in \mathcal{P}^I, i \in I, a \in A, p_{ia}(\succ) = p_{ia}(\succ'_i, \succ_{-i})$ *whenever* $U(\succ'_i, a) = U(\succ_i, a)$ *and* $\succ'_i |_{U(\succ'_i, a)} = \succ_i |_{U(\succ'_i, a)}$.

Lemma 12. *The lexicographic rule* $\lambda(\succ)$ *is weakly invariant*

Proof. The lexicographic serial rule is weakly invariant because it is bounded invariant. □

4.3.8 Weakly Truncation Robust

A preference relation $\succ'_i$ is called a *truncation* of $\succ_i$ if $U(\succ'_i, \emptyset) \subseteq U(\succ_i, \emptyset)$ and $\succ_i |_{U(\succ'_i, \emptyset)} = \succ'_i |_{U(\succ'_i, \emptyset)}$. Applying to the lexicographic serial rule, it is the same as preferences restricted to the set of remaining objects.

Definition 10 (Weakly truncation robust). *(See Noda, 2020) An allocation* p *is weakly truncation robust if for all* $\succ \in \mathcal{P}^I, i \in I$ *and* $a \in A, p_{ia}(\succ) = p_{ia}(\succ'_i, \succ_{-i})$ *whenever* $a \succ'_i \emptyset$ *and* $\succ'_i$ *is a truncation of* $\succ_i$.

Lemma 13. $\lambda(\succ)$ *is weakly truncation robust.*

Proof. When $\succ'_i$ is a truncation of $\succ_i$, $U(\succ'_i, \emptyset) \subseteq U(\succ_i, \emptyset)$. Due to the construction of the lexicographic serial rule that agents consume their more favourite objects first, the allocation restricted to the set of more favourite objects would be the same. □

4.4 Egalitarian Lex Welfare Maximisation

In this section, we study the relationship between the lexicographic serial rule and welfare maximisation.

Definition 11. *An allocation* p *is an* egalitarian lexicographic welfare maximiser *(ELWM) for* $\succ$ *if, for all sufficiently small* $\varepsilon > 0$, *there does not exist an allocation* p' *such that*

$$\sum_{i \in I} \sum_{k=1}^{m} \varepsilon^k p'_{i, f_k(\succ_i)} > \sum_{i \in I} \sum_{k=1}^{m} \varepsilon^k p_{i, f_k(\succ_i)}. \quad (*)$$

An allocation p is a weak egalitarian lex welfare maximiser *(WELWM) for $\succ$ if there does not exist an allocation p' such that (∗) holds for all sufficiently small $\varepsilon > 0$.*

Lemma 14. *If an allocation p is an ELWM for $\succ$ then it is lex unbiased for $\succ$.*

Proof. Suppose that p is an ELWM for $\succ$. Consider agents i and j and an object a such that $r_a(\succ_i) > r_a(\succ_j)$ and $p_{ia} > 0$. If $F(\succ_j, a, p_j) < \alpha_j$, then we can create an allocation p' by tranferring some a from i to j in exchange for an equal amount of any object that j likes less than a, and again (∗) will hold when ε is sufficiently small. Thus p is lex unbiased. □

Lemma 15. *The lex serial rule needs not give an ELWM.*

Proof. The example below shows that the lex serial rule need not give an ELWM. Let $I = \{1, 2, 3\}$ and $A = \{a, b, c\}$. Consider the following preference profile where each agent has an appetite of 1:

$$a \succ_1 b \succ_1 c \qquad a \succ_2 c \succ_2 b \qquad c \succ_3 a \succ_3 b.$$

The lexicographic serial rule give the allocation:

	a	b	c
1	$\frac{1}{2}$	$\frac{1}{2}$	0
2	$\frac{1}{2}$	$\frac{1}{2}$	0
3	0	0	1

However, the allocation:

	a	b	c
1	$\frac{1}{3}$	$\frac{2}{3}$	0
2	$\frac{2}{3}$	$\frac{1}{3}$	0
3	0	0	1

has the same total allocation of first favourites and a greater allocation of second favourites. □

4.5 Comparing the Lexicographic Serial Rule and the Probabilistic Serial Rule

In this section, we compare the lexicographic rule with the serial rule in terms of some important properties.

4.5.1 Lex Efficiency

Lemma 16. *The lexicographic serial rule is lex efficient but the probabilistic serial rule is not.*

Proof. We proved that the lexicographic serial rule is lex efficient in Section 4.3.1. To show that the probabilistic serial rule is not lex efficient, we return to an example discussed in Section 4.1. Assuming that we have the preference profile:

$$a \succ_1 b \succ_1 c, \quad a \succ_2 c \succ_2 b, \quad b \succ_3 c \succ_3 a.$$

The probabilistic serial rule gives the assignment matrix p:

$$\begin{array}{cccc} & a & b & c \\ 1 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 2 & \frac{1}{2} & 0 & \frac{1}{2} \\ 3 & 0 & \frac{3}{4} & \frac{1}{4} \end{array}$$

We will show that this allocation under probabilistic serial rule is weakly lex Pareto dominated by the allocation under the lexicographic serial rule, which is $p^{'}$:

$$\begin{array}{cccc} & a & b & c \\ 1 & \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 2 & \frac{1}{2} & 0 & \frac{1}{2} \\ 3 & 0 & \frac{3}{4} & \frac{1}{4} \end{array}$$

Let choose $k = 1$, we have $p^{'}_{3f_{1(\succ_3)}} = 1 > \frac{3}{4} = p_{3f_{1(\succ_3)}}$ for agent 3. Meanwhile, for agents 1 and 2, i.e., for $i = \{1, 2\}$, we have $p^{'}_{if_{1(\succ_i)}} = p_{if_{1(\succ_i)}}$. Thus, by Definition 1, p is not lex efficient. □

4.5.2 Ordinal Fairness

Definition 12 (Ordinally fair)**.** *(See Bogomolnaia and Moulin, 2001) An assignment is* ordinally fair *if for all agents $i, j \in I$ and $a \in A$ with $p_{j,a} > 0$, $F(\succ_i, a, \lambda_i(\succ)) \geq F(\succ_j, a, \lambda_j(\succ))$.*

Lemma 17. *The probabilistic serial rule is ordinally fair but the lexicographic rule is not.*

Proof. For the proof that the probabilistic serial rule is ordinally fair, see Theorem 1 in Bogomolnaia and Heo (2012). To see that the lexicographic rule is not ordinally fair, consider the following preference profile:

Agent 1: $a \succ_1 c \succ_1 d \succ_1 b \succ_1 e$

Agent 2: $a \succ_2 c \succ_2 d \succ_2 b \succ_2 e$

Agent 3: $a \succ_3 c \succ_3 d \succ_3 b \succ_3 e$

Agent 4: $b \succ_4 d \succ_4 c \succ_4 a \succ_4 e$

Agent 5: $b \succ_5 c \succ_5 e \succ_5 a \succ_5 d$

The allocation resulting from the lexicographic rule is L:

	a	b	c	d	e
1	$\frac{1}{3}$	0	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{4}$
2	$\frac{1}{3}$	0	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{4}$
3	$\frac{1}{3}$	0	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{4}$
4	0	$\frac{1}{2}$	0	$\frac{1}{2}$	0
5	0	$\frac{1}{2}$	$\frac{1}{4}$	0	$\frac{1}{4}$

In this example, consider the upper contour set at d, agent 4's total consumption of objects that are at least as good as d is $F(\succ_4, d, \lambda_4(\succ)) = 1$, while each agent 1, 2, and 3 only has 3/4 total consumption of objects that are at least as good as d. Thus, it is not ordinally fair. $\square$

4.5.3 Sd Envy-Free

Definition 13 (sd envy-free)**.** *(See Bogomolnaia and Moulin, 2001) An allocation rule p is sd envy-free if for each pair $i, j \in I$, $p_i \succ_i^{sd} p_j$.*

Lemma 18. *The probabilistic serial rule is sd envy-free, while the lexicographic serial rule is not.*

Proof. For the proof that the probabilistic serial rule is sd envy-free, see Theorem 2 in Bogomolnaia and Heo (2012). The following example shows that the lexicographic rule is not sd envy-free, while the serial rule is. Consider a situation in which there are three agents with preferences as follow:

$$\text{Agent 1: } a \succ_1 b \succ_1 c$$
$$\text{Agent 2: } b \succ_2 a \succ_2 c$$
$$\text{Agent 3: } b \succ_3 a \succ_3 c.$$

The resulting allocation under the lexicographic rule is:

	a	b	c
1	1	0	0
2	0	$\frac{1}{2}$	$\frac{1}{2}$
3	0	$\frac{1}{2}$	$\frac{1}{2}$

Because $p_{1a} = 1 > p_{2b}$ while $a, b \in U(\succ_1, c) \cap U(\succ_2, c)$, p is not sd envy-free. On the other hand, the probabilistic serial rule obviously satisfies sd envy-free, because the allocation resulting from it is

	a	b	c
1	$\frac{2}{3}$	0	$\frac{1}{3}$
2	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$
3	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{1}{3}$

□

4.6 The Lexicographic Serial Rule Without Catch Up Phase

This section motivates the existence of a catch-up phase in the lexicographic serial rule.

Consider the following preference profile $\succ$:

$$\begin{aligned}
&\text{Agent 1: } a \succ_1 c \succ_1 d \succ_1 e \succ_1 b\\
&\text{Agent 2: } a \succ_2 c \succ_2 d \succ_2 e \succ_2 b\\
&\text{Agent 3: } a \succ_3 c \succ_3 d \succ_3 e \succ_3 b\\
&\text{Agent 4: } b \succ_4 c \succ_4 d \succ_4 e \succ_4 a\\
&\text{Agent 5: } b \succ_5 c \succ_5 d \succ_5 e \succ_5 a
\end{aligned}$$

The allocation that results from the lexicographic rule without a catch-up phase is λ':

	a	b	c	d	e
1	$\frac{1}{3}$	0	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{4}{15}$
2	$\frac{1}{3}$	0	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{4}{15}$
3	$\frac{1}{3}$	0	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{4}{15}$
4	0	$\frac{1}{2}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{10}$
5	0	$\frac{1}{2}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{10}$

The allocation resulting from the lexicographic rule is λ:

	a	b	c	d	e
1	$\frac{1}{3}$	0	$\frac{1}{6}+\frac{1}{10}$	$\frac{1}{5}$	$\frac{1}{5}$
2	$\frac{1}{3}$	0	$\frac{1}{6}+\frac{1}{10}$	$\frac{1}{5}$	$\frac{1}{5}$
3	$\frac{1}{3}$	0	$\frac{1}{6}+\frac{1}{10}$	$\frac{1}{5}$	$\frac{1}{5}$
4	0	$\frac{1}{2}$	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{1}{5}$
5	0	$\frac{1}{2}$	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{1}{5}$

This example shows that the lexicographic rule without a catch-up phase does not satisfy fairness. Looking at the allocation of object c, because we have:

$$F(\succ_4, c, \lambda'_4(\succ)) = F(\succ_5, c, \lambda'_5(\succ)) = \frac{7}{10} > \frac{8}{15} = F(\succ_1, c, \lambda'_1(\succ)) = F(\succ_2, c, \lambda'_2(\succ)) = F(\succ_3, c, \lambda'_3(\succ))$$

although all agents obtain strictly positive probabilities for objects that they prefer less than object c. This example illustrates how we would fail to characterise the lexicographic rule if it lacks a catch-up phase.

4.7 Conclusion

In this chapter, we propose a new allocation rule that is appropriate for situations in which each agent cares primarily about getting her favourite, cares much more about the difference between her second and third favourite than the difference between her third and fourth, and so on. Although it has a narrow domain of application, within that domain it has a clear welfarist interpretation. Bogomolnaia and Moulin (2001) stress that, at the time they proposed the probabilistic serial rule, there was no alternative to random priority that depended only on ordinal preferences. We provide such an alternative.

Assignment rules that depend only on ordinal preferences have important practical advantages for several reasons. First, it is often difficult to elicit vNM utilities. Second, there is no efficient and equitable assignment rule based on cardinal preferences that is strategy-proof. Finally, it can be difficult for agents to understand how their declaration of a vNM utility affects what they receive.

Our lexicographic serial rule can be described as an eating procedure that is divided into rounds: a round of favourites, a round of second favourites, and so forth. In the round of favourites, each object that is someone's favourite is divided equally among all agents for whom it is the favourite. Suppose, after rounds $1, 2, \cdots, k-1$ we reach the round of k^{th} favourites. Agents who are not yet sated, and whose k^{th} favourite has not yet been exhausted, consume this object until they are all sated or it is exhausted. We propose that agents with least total previous probability "eat" first, and each other agent joins in when the total probability of the eating agents reaches their total probability. This is a novel variant of the Boston school choice mechanism that offers an improvement in terms of fairness.

Our main contribution to the literature is to propose an allocation rule that depends only on ordinal preferences and show that this rule is lex efficient, lex fair, and lex envy-free. A potential direction for future research would be to develop an alternative to the serial rule that is appropriate when agents are primarily concerned with avoiding their worst objects. Dividing "bads" is typically harder than dividing "goods", because agents no longer have the option of throwing away objects at no cost, and contemporary literature under linear preferences has not found a rule that normatively dominates the other (Chaudhury et al., 2020; Bogomolnaia et al., 2019). Thus, it might be promising to study the chore division problem under lexicographic preferences, when the difference between the third and fourth worst object is insignificant in comparison with the difference between the second and third worst object, and so on.

Chapter 5

Conclusion

> Far and away the best prize that life has to offer is the chance to work hard at work worth doing
>
> *Theodore Roosevelt*

This thesis makes empirical and theoretical contributions that can translate into practical action to improve people's lives. It does so by showing how to allocate scarce resources efficiently and equitably through studies that highlight the types and sources of inequality and potential change in resource allocations that can improve the welfare of individuals and society as a whole. In many contexts, welfare improvement requires a collaborative effort, both from individuals making decisions that maximise their own utilities and from governments and NGOs that provide targeted help for groups of vulnerable individuals.

In Chapter 2, we developed a model that can assist governments and NGOs to identify communities in Vietnam that are most affected by weather shocks and enable efficient provision of much-needed support. We started Chapter 2 asking how weather shocks would affect household welfare (measured as equivalised income and per capita consumption) and whether weather shocks exacerbate income inequality. While Vietnam was chosen as a convenient example in this research, due to data availability and our context expertise, climate change and weather-related disasters have become increasingly relevant in all jurisdictions. The frequency and intensity of these events will inevitably require more research to identify vulnerable communities and measure the implications of natural disasters for economic livelihoods, social welfare and inequality. The method demonstrated in this chapter will be useful in this research effort and can be replicated elsewhere.

The first step in answering the question posed above was to estimate the impacts of weather shocks on equivalised income and per capita inequality. We first recognised the limitations of existing measures, which are often dependent on the duration of weather samples observed. We constructed a new measure of shock, defined as the total number of days with rainfall of at least 100 mm when there were also at least two such days in a row in the period coinciding with the household survey. Using the newly constructed measure of weather shocks, we found that they have a significant negative impact on the incomes of rural households, and their health consumption, and force them to increase spending on rent. Robustness checks of the results were performed against other measures used in the literature.

In establishing the link between weather shocks and income inequality, we checked that the impacts of weather shocks vary depending on individual characteristics. In addition, we used the Gini index, one of the most well-known inequality measures, to capture both relative and absolute inequality (Kakwani and Son, 2016). On computing the Gini decomposition of income source, we found that crop income contributes to reducing income inequality. Since weather shocks reduce income from crops, they contribute to increasing income inequality.

Our findings add to the literature on geographical location and inequality that recognises that urban residents have more job opportunities than rural residents, who are often engaged in agricultural activities (Kakwani and Son, 2016). Even within rural regions, climate change can have unequal impacts and exacerbate inequality in areas that are hit harder by extreme weather patterns.

The findings from Chapter 2 should assist governments and NGOs to identify communities that are more prone to weather shocks so they can provide the necessary support. We also raised the issue that farmers in Vietnam are unable to insure themselves against weather shocks because of the underdeveloped market for agricultural insurance and weather derivatives, partly due to the typically small size of individual (family) farms and partly due to the persistent reliance on the informal economy rather than official banking and stock market systems. This situation is not unique to Vietnam – it is the case in most low to middle-income countries. It is essential for governments, especially in agriculturally intensive countries, to create and facilitate agricultural insurance and weather derivatives markets to mitigate the effects of unfavourable weather patterns.

We used Vietnam as a case study in Chapter 2; however, weather shocks and natural disasters occur worldwide. The new measure of weather shocks that we developed in Chapter 2 can be extended to measure weather shocks in other jurisdictions, especially those with large agriculture sectors, such as Australia and South American and South Asian nations. For example, Brazil, one of the world's leading grain producers and exporters, has a climate that varies greatly from the equatorial north to the temperate south. Heavy rainfall can curtail agricultural production in Brazil significantly in the short term, threatening the country's food security and export income (Pereda and Alves, 2018).

Moreover, even farmers in high-income countries such as Australia are not entirely protected from weather shocks. Although Australian farmers typically have much larger farms than their Vietnamese counterparts, they face similar difficulties in insuring their farms against natural disasters. It is also worth noting that our new measure can easily be adjusted to determine weather shocks in dry jurisdictions such as India, for example, by counting the number of days when rainfall is continuously below a certain threshold.

Although income inequality tends to dominate the inequality debate, non-income disparities exist and often have strong interactions with income inequality. In Chapter 3, we sought to explain the gender mental health gap in Australia, with a special focus on people with poor psychological wellbeing. Our main research question was whether economic disparities would be sufficient to explain the gender mental health gap and to what extent the negative life events contribute to the gap.

To answer this research question, we used decomposition based on regression analysis of the relationships between mental health and its correlates, which allowed us to see how much of the gender mental health inequality can be explained by disparities in socio-economic factors or negative life events. We found that disparities in income play a very substantial role, and account for most of the gender gap amongst individuals with very poor psychological wellbeing. However, education, as an indicator of socioeconomic status, narrows the gap. Negative life events have a large effect at the individual level, but they are not correlated with gender strongly enough to explain the observed mental health disparities.

In Chapter 4, we stepped back and considered the resource allocation problem in a more general setting. In a variety of contexts, it is considered morally and ethically unfair for people to be able to use their funds to achieve their desire outcomes at others' expense. For example, when assigning patients to receive treatment, a nurse or physician considers the relative urgency and severity of the conditions at hand, and the likely efficacy of treatment for individual patients, rather than their relative (financial) wealth. Similarly, for a country that experiences frequent natural disasters that affect communities disproportionally, governments and international donors might consider prioritising the need of more vulnerable (often agricultural) communities, rather than those less affected by weather shocks due to reliance on more stable and wealthy industrial or service sec- tors. Keeping these issues in mind, the new allocation rule we proposed in Chapter 4 focuses on situations in which money transfer is prohibited. The lexicographic serial rule is relatively simple to implement and maximises the probability of receiving the most favourite "objects" (e.g., healthcare treatments, priority loans or relief) for each agent (e.g., individual patients with serious life events, rural/remote communities with poor healthcare access, natural-disaster prone areas with large agricultural or tourism sectors), hence improving the efficiency and equity of resource allocation.

Putting the lexicographic serial rule to work in the context of Chapter 2, we can think of a situation in which agents are volunteers who help people suffering from natural disasters, performing activities such as cleaning or rebuilding homes after a flood. Each volunteer has preferences about where to do the work, depending on factors like the locations' distance from their homes or the nature of the jobs. The set of objects in this situation would be villages that need volunteers to support them in cleaning and rebuilding homes after a flood. Applying the lexicographic serial rule would ensure that each volunteer is most likely to be assigned to work in the village they prefer most. The strong alignment between "preference" and "job" means that each volunteer's skills are likely to be employed in the best possible way (resulting in an efficacy gain), higher fidelity (i.e., volunteers do not quit halfway, resulting in an effectiveness gain) and most villages receive support (resulting in an equity gain).

Similarly, in the context of Chapter 3, we might need to allocate psychologists or psychiatrists to different practices or hospitals to provide support for people who experience mental illnesses. Psychologists and psychiatrists are considered agents in the context of allocation problems; each has their own preferences for the location of work. Vacant positions at practices or hospitals (objects) need to be allocated to these professionals. If we apply the lexicographic serial rule, we can maximise the probability that each practitioner is assigned to the position that they prefer, while ensuring that all the locations are filled.

In conclusion, the three essays in this thesis relate to allocating scarce resources that meet certain criteria with a focus on equity. Allocations of resources can be unequal between individuals, and the way that resources are divided has implications for the distribution of resources. Climate change can have unequal impacts on geographical regions, potentially exacerbating inequality when some areas are hit harder by extreme weather patterns. Richer societies might have greater ability to adapt to and mitigate the effects of extreme weather. Similarly, the burden of poor mental health tends to fall disproportionately on people with the fewest resources. Applying the lexicographic serial rule in allocating human resources to alleviate the impacts of weather shocks or mental illness will ensure that each individual is likely to be allocated to the job they prefer most, thus improving the overall welfare of society.

Bibliography

Acemoglu, D. (1998). Why do new technologies complement skills? directed technical change and wage inequality. *The Quarterly Journal of Economics*, 113(4):1055–1089.

Affleck, W., Carmichael, V., and Whitley, R. (2018). Men's mental health: Social determinants and implications for services. *The Canadian Journal of Psychiatry*, 63(9):581–589.

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*, 2(3):244–263.

Australian Bureau of Statistics (2007). National survey of mental health and wellbeing: Summary of results. https://www.abs.gov.au/statistics/health/mental-health/national-survey-mental-health-and-wellbeing-summary-results/latest-release.

Backhouse, R. E., Baujard, A., and Nishizawa, T., editors (2021). *Welfare Theory, Public Action, and Ethical Values: Revisiting the History of Welfare Economics*. Cambridge University Press, first edition.

Banerjee, A. V. and Duflo, E. (2007). The economic lives of the poor. *Journal of Economic Perspectives*, 21(1):141–167.

Benjamin, D., Brandt, L., and McCaig, B. (2017). Growth with equity: income inequality in Vietnam, 2002–14. *The Journal of Economic Inequality*, 15(1):25–46.

Bogomolnaia, A. (2015). Random assignment: Redefining the serial rule. *Journal of Economic Theory*, 158:308–318.

Bogomolnaia, A. and Heo, E. J. (2012). Probabilistic assignment of objects: Characterizing the serial rule. *Journal of Economic Theory*, 147(5):2072–2082.

Bogomolnaia, A. and Moulin, H. (2001). A new solution to the random assignment problem. *Journal of Economic Theory*, 100(2):295–328.

Bogomolnaia, A., Moulin, H., Sandomirskiy, F., and Yanovskaia, E. (2019). Dividing bads under additive utilities. *Social Choice and Welfare*, 52(3):395–417.

Borah, B. J. and Basu, A. (2013). Highlighting differences between conditional and unconditional quantile regression approaches through an application to assess medication adherence. *Health Economics*, 22(9):1052–1070.

Brinig, M. and Allen, D. (2000). 'These boots are made for walking': why most divorce filers are women. *American Law and Economics Review*, 2(1):126–169.

Budish, E., Che, Y.-K., Kojima, F., Milgrom, P., Kamenica, S., Lahaie, M., Manea, H., Moulin, T., Sönmez, A., Roth, U., and Ünver (2011). Implementing random assignments: A generalization of the Birkhoff-von Neumann theorem.

Burns, T., Catty, J., Becker, T., Drake, R. E., Fioritti, A., Knapp, M., Lauber, C., Rössler, W., Tomov, T., van Busschbach, J., White, S., and Wiersma, D. (2007). The effectiveness of supported employment for people with severe mental illness: a randomised controlled trial. *The Lancet*, 370(9593):1146–1152.

Butterworth, P., Rodgers, B., and Windsor, T. D. (2009). Financial hardship, socio-economic position and depression: Results from the path through life survey. *Social Science & Medicine*, 69(2):229–237.

Card, D. and DiNardo, J. E. (2002). Skill-biased technological change and rising wage inequality: Some problems and puzzles. *Journal of Labor Economics*, 20(4).

Chaudhury, B. R., Garg, J., McGlaughlin, P., and Mehta, R. (2020). Dividing bads is harder than dividing goods: On the complexity of fair and efficient division of chores.

Chernozhukov, V., Fernández-Val, I., and Melly, B. (2013). Inference on counterfactual distributions. *Econometrica*, 81(6):2205–2268.

Chisholm, D., Sweeny, K., Sheehan, P., Rasmussen, B., Smit, F., Cuijpers, P., and Saxena, S. (2016). Scaling-up treatment of depression and anxiety: a global return on investment analysis. *The Lancet Psychiatry*, 3(5):415–424.

Clark, A. E. and D'Ambrosio, C. (2015). Attitudes to income inequality. In *Handbook of Income Distribution*, volume 2, pages 1147–1208. Elsevier.

Clark, A. E., D'Ambrosio, C., and Barazzetta, M. (2020). Childhood circumstances and young adulthood outcomes: The role of mothers' financial problems. *Health Economics*, 30(2):342–357.

Clark, A. E., Frijters, P., and Shields, M. A. (2008). Relative income, happiness, and utility: An explanation for the easterlin paradox and other puzzles. *Journal of Economic Literature*, 46(1):95–144.

Coombs, R. H. (1991). Marital status and personal well-being: A literature review. *Family Relations*, 40(1):97–102.

Corno, L., Hildebrandt, N., and Voena, A. (2020). Age of marriage, weather shocks, and the direction of marriage payments. *Econometrica*, 88(3):879–915.

Cutler, D. and Lleras-Muney, A. (2006). Education and health: Evaluating theories and evidence. Technical Report w12352, National Bureau of Economic Research, Cambridge, MA.

Dabla-Norris, M. E., Kochhar, M. K., Suphaphiphat, M. N., Ricka, M. F., and Tsounta, M. E. (2015). *Causes and Consequences of Income Inequality: A Global Perspective*. International Monetary Fund.

Davis, J. W., Parks, S. N., Kaups, K. L., Bennink, L. D., and Bilello, J. F. (2003). Victims of domestic violence on the trauma service: Unrecognized and underreported. *Journal of Trauma and Acute Care Surgery*, 54(2):352–355.

De Battisti, F., Porro, F., and Vernizzi, A. (2019). The gini coefficient and the case of negative values.

Decanq, K. and Neumann, D. (2016). Does the choice of well-being measure matter empirically?

Diener, E., Gohm, C. L., Suh, E., and Oishi, S. (2000). Similarity of the relations between marital status and subjective well-being across cultures. *Journal of Cross-Cultural Psychology*, 31(4):419–436.

Do, T. L., Nguyen, T. T., and Grote, U. (2021). Livestock production and income inequality in rural Vietnam. *Empirical Economics*.

Drijber, B. C., Reijnders, U. J. L., and Ceelen, M. (2013). Male victims of domestic violence. *Journal of Family Violence*, 28(2):173–178.

Dynan, K. E., Skinner, J., and Zeldes, S. P. (2004). Do the rich save more? *Journal of Political Economy*, 112(2):48.

Feenstra, R. C. and Hanson, G. H. (1996). Globalization, outsourcing, and wage inequality. *The American Economic Review*, 86(2):240–245.

Feenstra, R. C. and Hanson, G. H. (2001). Global production sharing and rising inequality: A survey of trade and wages. *National Bureau of Economic Research Cambridge, Mass., USA*.

Figini, P. and Görg, H. (2011). Does foreign direct investment affect wage inequality? an empirical investigation. *The World Economy*, 34(9):1455–1475.

Firpo, S., Fortin, N. M., and Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, 77(3):953–973.

Flèche, S., Lepinteur, A., and Powdthavee, N. (2020). Gender norms, fairness and relative working hours within households. *Labour Economics*, 65:101866.

Fleurbaey, M. (2012). *Fairness, Responsibility, and Welfare*. Oxford University Press, Oxford.

Fleurbaey, M. and Maniquet, F. (2011). *A theory of fairness and social welfare*. Econometric society monographs. Cambridge University Press, Cambridge ; New York.

Fortin, N., Lemieux, T., and Firpo, S. (2011). Decomposition methods in economics. In *Handbook of Labor Economics*, volume 4, pages 1–102. Elsevier.

Glenn, N. D. (1975). The contribution of marriage to the psychological well-being of males and females. *Journal of Marriage and Family*, 37(3):594–600.

Grabrucker, K. and Grimm, M. (2021). Is there a rainbow after the rain? how do agricultural shocks affect non-farm enterprises? evidence from thailand. *American Journal of Agricultural Economics*, 103(5):1612–1636.

Grierson, J. (2020). Domestic abuse killings 'more than double' amid COVID-19 lockdown. *The Guardian*. https://www.theguardian.com/society/2020/apr/15/domestic-abuse-killings-more-than-double-amid-covid-19-lockdown.

Grossman, M. (2017). *Determinants of Health: An Economic Perspective*. Columbia University Press.

Guiteras, R., Jina, A., and Mobarak, A. M. (2015). Satellites, self-reports, and submersion: Exposure to floods in bangladesh. *American Economic Review*, 105(5):232–236.

Hampel, F. R. (1974). The influence curve and its role in robust estimation. *Journal of the American Statistical Association*, 69(346):383–393.

Hashimoto, T., Hirata, D., Kesten, O., Kurino, M., and Ünver, M. U. (2014). Two axiomatic approaches to the probabilistic serial mechanism. *Theoretical Economics*, 9(1):253–277.

Hashmi, R., Alam, K., and Gow, J. (2020). Socioeconomic inequalities in mental health in Australia: Explaining life shock exposure. *Health Policy*, 124(1):97–105.

Heo, E. J. (2014). Probabilistic assignment problem with multi-unit demands: A generalization of the serial rule and its characterization. *Journal of Mathematical Economics*, 54:40–47.

Hoddinott, J. (2006). Shocks and their consequences across and within households in rural zimbabwe. *Journal of Development Studies*, 42(2):301–321.

Hoeymans, N., Garssen, A. A., Westert, G. P., and Verhaak, P. F. (2004). Measuring mental health of the Dutch population: a comparison of the GHQ-12 and the MHI-5. *Health and Quality of Life Outcomes*, page 6.

Huguet, N., Kaplan, M. S., and Feeny, D. (2008). Socioeconomic status and health-related quality of life among elderly people: Results from the joint Canada/United States survey of health. *Social Science & Medicine*, 66(4):803–810.

Hylland, A. and Zeckhauser, R. (1979). The efficient allocation of individuals to positions. *Journal of Political Economy*, 87(2):293–314.

Imbert, C., Seror, M., Zhang, Y., and Zylberberg, Y. (2018). Migrants and firms: Evidence from china. *SSRN Electronic Journal*.

Jayachandran, S. (2006). Selling labor low: Wage responses to productivity shocks in developing countries. *Journal of Political Economy*, 114(3):538–575.

Kahneman, D., Diener, E., and Schwarz, N. (1999). *Well-Being: Foundations of Hedonic Psychology*. Russell Sage Foundation.

Kakwani, N. and Son, H. H. (2016). *Social Welfare Functions and Development*. Palgrave Macmillan UK, London.

Kang, W. and Imai, K. S. (2012). Pro-poor growth, poverty and inequality in rural Vietnam. *Journal of Asian Economics*, 23(5):527–539.

Kaur, S. (2019). Nominal wage rigidity in village labor markets. *American Economic Review*, 109(10):3585–3616.

Kawachi, I. and Kennedy, B. P. (1999). Income inequality and health: pathways and mechanisms. *Health Services Research*, 34(1 Pt 2):215–227. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1088996/.

Kessels, R. (2020). A distributional regression approach to income-related inequality of health in Australia. *International Journal for Equity in Health*, page 19.

Kim, S. and Koh, K. (2021). The effects of income on health: Evidence from lottery wins in singapore. *Journal of Health Economics*, 76:102414.

Kim, Y., Lee, H., and Park, A. (2020). Adverse childhood experiences, economic hardship, and obesity: Differences by gender. *Children and Youth Services Review*, 116:105214.

King, M. and Singh, A. P. (2020). Understanding farmers' valuation of agricultural insurance: Evidence from Vietnam. *Food Policy*, 94:101861.

Kneib, T., Silbersdorff, A., and Säfken, B. (2021). Rage against the mean – a review of distributional regression approaches. *Econometrics and Statistics*.

Koenker, R. and Hallock, K. F. (2001). Quantile regression. *Journal of Economic Perspectives*, 15(4):143–156.

Kojima, F. and Ünver, M. U. (2014). The "Boston" school-choice mechanism: an axiomatic approach. *Economic Theory*, 55(3):515–544.

Kulhánová, I., Hoffmann, R., Judge, K., Looman, C. W. N., Eikemo, T. A., Bopp, M., Deboosere, P., Leinsalu, M., Martikainen, P., Rychtaříková, J., Wojtyniak, B., Menvielle, G., and Mackenbach, J. P. (2014). Assessing the potential impact of increased participation in higher education on mortality: Evidence from 21 European populations. *Social Science & Medicine*, 117:142–149.

Le, Q. H. and Nguyen, B. N. (2020). The impact of credit on income inequality in Vietnam. *The Journal of Asian Finance, Economics and Business*, 7(5):111–118.

Le, T. N. T. (2020). Floods and household welfare: Evidence from southeast asia. *Economics of Disasters and Climate Change*, 4(1):145–170.

Lenhart, O. (2019). The effects of income on health: new evidence from the earned income tax credit. *Review of Economics of the Household*, 17(2):377–410.

Lepinteur, A. (2021). The asymmetric experience of gains and losses in job security on health. *Health Economics*, 30(9):2217–2229.

Lerman, R. I. and Yitzhaki, S. (1985). Income inequality effects by income source: A new approach and applications to the united states. *The Review of Economics and Statistics*, 67(1):151–156.

Li, J. and Powdthavee, N. (2015). Does more education lead to better health habits? evidence from the school reforms in Australia. *Social Science & Medicine*, 127:83–91.

Lohmann, S. and Lechtenfeld, T. (2015). The effect of drought on health outcomes and health expenditures in rural Vietnam. *World Development*, 72:432–448.

Lordan, G., Soto, E. J., Brown, R. P. C., and Correa-Valez, I. (2012). Socioeconomic status and health outcomes in a developing country. *Health Economics*, 21(2):178–186.

Maccini, S. and Yang, D. (2009). Under the weather: Health, schooling, and economic consequences of early-life rainfall. *American Economic Review*, 99(3):1006–1026.

Marx, N. d. L.-L. (2018). Weather shocks, income and inequality. http://piketty.pse.ens.fr/files/Laubier2018.pdf.

Mas-Colell, A., Whinston, M. D., and Green, J. R. (1995). *Microeconomic Theory*. Oxford University Press.

McLennan, A. (2002). Ordinal efficiency and the polyhedral separating hyperplane theorem. *Journal of Economic Theory*, 105(2):435–449.

Menta, G. and Lepinteur, A. (2021). Boys don't cry (or do the dishes): Family size and the housework gender gap. *Journal of Economic Behavior & Organization*, 186:164–188.

Menta, G., Lepinteur, A., Clark, A. E., Ghislandi, S., and D'Ambrosio, C. (2021). Maternal depression and child human capital: A genetic instrumental-variable approach. https://dondena.unibocconi.eu/working-paper-series/wp_143.

Morasae, E. K., Forouzan, A. S., Majdzadeh, R., Asadi-Lari, M., Noorbala, A. A., and Hosseinpoor, A. R. (2012). Understanding determinants of socioeconomic inequality in mental health in iran's capital, tehran: a concentration index decomposition approach. *International Journal for Equity in Health*, 11(1):18.

Mullainathan, S. and Shafir, E. (2013). *Scarcity: Why Having Too Little Means So Much*. Times Books.

Myerson, R. B. (1981). Utilitarianism, egalitarianism, and the timing effect in social choice problems. *Econometrica*, 49(4):883–897.

Neckerman, K. M. and Torche, F. (2007). Inequality: Causes and consequences. *Annual Review of Sociology*, 33(1):335–357.

Nguyen, G. and Nguyen, T. T. (2020). Exposure to weather shocks: A comparison between self-reported record and extreme weather data. *Economic Analysis and Policy*, 65:117–138.

Nguyen, T.-T., Nguyen, T. T., and Grote, U. (2020a). Credit and ethnic consumption inequality in the central highlands of Vietnam. *Social Indicators Research*, 148(1):143–172.

Nguyen, T. T., Nguyen, T.-T., Le, V.-H., Managi, S., and Grote, U. (2020b). Reported weather shocks and rural household welfare: Evidence from panel data in Northeast Thailand and Central Vietnam. *Weather and Climate Extremes*, 30:100286.

Nguyen, T. V. and Tran, T. Q. (2018). Forestland and rural household livelihoods in the North Central Provinces, Vietnam. *Land Use Policy*, 79:10–19.

Noda, S. (2020). Size versus truncation robustness in the assignment problem. *Journal of Mathematical Economics*, 87:1–5.

O'Donnell, O., van Doorslaer, E., Wagstaff, A., and Lindelow, M. (2007). *Analyzing Health Equity Using Household Survey Data: A Guide to Techniques and their Implementation*. The World Bank.

Oishi, S., Kesebir, S., and Diener, E. (2011). Income inequality and happiness. *Psychological Science*, 22(9):1095–1100.

Osmani, S. and Sen, A. (2003). The hidden penalties of gender inequality: fetal origins of ill-health. *Economics & Human Biology*, 1(1):105–121.

Parker, G., Durante, K. M., Hill, S. E., and Haselton, M. G. (2022). Why women choose divorce: An evolutionary perspective. *Current Opinion in Psychology*, 43:300–306.

Pereda, P. and Alves, D. (2018). Climate and weather impacts on agriculture: the case of brazil — economia aplicada. *Economia Aplicada*, 22(3):5–26. https://www.revistas.usp.br/ecoa/article/view/119584.

Pickett, K. E. and Wilkinson, R. G. (2010). Inequality: an underacknowledged source of mental illness and distress. *The British Journal of Psychiatry*, 197(6):426–428.

Pickett, K. E. and Wilkinson, R. G. (2015). Income inequality and health: A causal review. *Social Science & Medicine*, 128:316–326.

Pudrovska, T., Reither, E. N., Logan, E. S., and Sherman-Wilkins, K. J. (2014). Gender and reinforcing associations between socioeconomic disadvantage and body mass over the life course. *Journal of Health and Social Behavior*, 55(3):283–301.

Quiñones, E. J., Liebenehm, S., and Sharma, R. (2021). Left home high and dry-reduced migration in response to repeated droughts in thailand and Vietnam. *Population and Environment*, 42(4):579–621.

Quyen, N. H. (2019). Reducing rural poverty in Vietnam: Issues, policies, challenges. Technical report, Mekong Development Research Institute. https://www.un.org/development/desa/dspd/wp-content/uploads/sites/22/2019/03/pdfresizer.com-pdf-resize-4.pdf.

Rios Avila, F. (2019). Recentered influence functions in stata: Methods for analyzing the determinants of poverty and inequality. *SSRN Electronic Journal*.

Roemer, J. E. (1998). *Equality Of Opportunity*. Harvard University Press.

Rumpf, H.-J., Meyer, C., Hapke, U., and John, U. (2001). Screening for mental health: Validity of the MHI-5 using DSM-IV Axis I psychiatric disorders as gold standard. *Psychiatry Research*, 105(3):243–253.

Saleh, R. H. and Luppicini, R. (2017). Exploring the challenges of divorce on Saudi women. *Journal of Family History*, 42(2):184–198.

Sen, A. (1984). *Collective choice and social welfare*. Number v. 11 in Advanced textbooks in economics. North-Holland ; Sole distributors for the U.S.A. and Canada, Elsevier Science Pub. Co, Amsterdam ; New York : New York.

Sen, A. (2002). Why health equity? *Health Economics*, 11(8):659–666.

Shah, M. and Steinberg, B. M. (2017). Drought of opportunities: Contemporaneous and long-term impacts of rainfall shocks on human capital. *journal of political economy*, 125:35.

Shorrocks, A. F. (1982). Inequality decomposition by factor components. *Econometrica*, 50(1):193–211.

Stack, S. and Eshleman, J. R. (1998). Marital status and happiness: A 17-nation study. *Journal of Marriage and Family*, 60(2):527–536.

Stark, O., Taylor, J. E., and Yitzhaki, S. (1986). Remittances and inequality. *The Economic Journal*, 96(383):722–740.

Stiglitz, J. E. (2012). *The price of inequality: how today's divided society endangers our future*. New York, W.W. Norton & Co.

Su, Z., Cheshmehzangi, A., McDonnell, D., Šegalo, S., Ahmad, J., and Bennett, B. (2022). Gender inequality and health disparity amid covid-19. *Nursing Outlook*, 70(1):89–95.

Summerfield, M., Bevitt, A., Fok, Y. K., Hahn, M., La, N., Macalalad, N., O'Shea, M., Watson, N., Wilkins, R., and Wooden, M. (2020). Hilda user manual release 20.

The World Bank (2021). Gini index (world bank estimate). https://data.worldbank.org/indicator/SI.POV.GINI?end=2019&start=2019&view=map&year=2013.

Thorsen, S. V., Rugulies, R., Hjarsbech, P. U., and Bjorner, J. B. (2013). The predictive value of mental health for long-term sickness absence: The major depression inventory (MDI) and the mental health inventory (MHI-5) compared. *BMC Medical Research Methodology*, 13(1):115.

Tran, D. T. H. and Otake, T. (2020). Temperature-based options for Vietnam: An option pricing and policy insight. *Journal of Economics and International Finance*, 12(1):29–38.

Tran, T. Q., Nguyen, C. V., and Van Vu, H. (2018). Does economic inequality affect the quality of life of older people in rural Vietnam? *Journal of Happiness Studies*, 19(3):781–799.

Tran, V. Q. (2015). Household's coping strategies and recoveries from shocks in Vietnam. *The Quarterly Review of Economics and Finance*, 56:15–29.

Veisani, Y. and Delpisheh, A. (2015). Decomposing of socioeconomic inequality in mental health: A cross-sectional study into female-headed households. *Journal of Research in Health Sciences*, 15(4):218–222.

Vu, L. H. (2015). How we measure poverty in Vietnam. https://blogs.worldbank.org/eastasiapacific/how-we-measure-poverty-{V}ietnam.

Ware, J. E., Kosinski, M., and Keller, S. D. (1994). SF-36 physical and mental health summary scales: a user's manual.

Watson, N. and Wooden, M. (2021). The household, income and labour dynamics in Australia (HILDA) survey. *Jahrbücher für Nationalökonomie und Statistik*, 241(1):131–141.

World Population Review (2021). Gini coefficient by country 2021. https://worldpopulationreview.com/country-rankings/gini-coefficient-by-country.

Zhou, L. (1990). On a conjecture by gale about one-sided matching problems. *Journal of Economic Theory*, 52:123–135.

Appendix A

Appendix for Chapter 2

A.1 Descriptive statistics

Table A.1.1: Descriptive Statistics by Provinces in 2008 (all values are annual means)

	All	Ha Tinh	Hue	Dak Lak
Demographic variables				
Household size	5.7150	5.2744	5.9328	5.8916
Number of children	1.5422	1.1821	1.6674	1.7357
Gender (Female=1)	0.5082	0.5169	0.5078	0.5008
Age	29.3044	32.4281	29.2894	26.6452
Ethnicity (Kinh=1)	0.7737	0.9988	0.7563	0.5931
Marital status (Married=1)	0.3860	0.4035	0.3648	0.3907
Education: Primary school	0.6525	0.5105	0.7259	0.7059
Education: Secondary school	0.2322	0.3283	0.1815	0.1968
Education: High school	0.1063	0.1531	0.0848	0.0865
Education: Bachelor or higher	0.0090	0.0082	0.0078	0.0109
Farmer	0.3372	0.3455	0.2814	0.3824
Member of socio-political organisation	0.5110	0.7176	0.5419	0.3053
Percentage of socio-political women	0.0535	0.0747	0.0575	0.0315
Income variables (2005 PPP USD)				
Equivalised total income	2450.3438	2256.3640	1672.5100	3352.0620
Equivalised crop income	861.1995	539.7121	322.1815	1652.9473
Equivalised crop revenue	1368.4900	746.3568	511.6226	2723.9795
Equivalised livestock income	231.8785	384.3237	45.0325	276.3623
Equivalised hunting income	74.7231	55.3639	120.4305	37.3654
Equivalised off-farm employment income	282.3772	262.5458	287.2668	296.5757
Equivalised remittance: family/friends	181.9482	291.7136	102.4960	161.0609
Consumption variables (2005 PPP USD)				
Per capita total consumption	940.9489	875.1804	857.0137	1079.8651
Per capita food consumption	483.3882	426.5717	460.3430	556.1908
Per capita nonfood consumption	345.8316	316.2999	314.3599	401.6792
Per capita education consumption	53.5194	76.6693	33.8980	52.0696
Per capita health consumption	34.0296	34.2197	24.0510	43.4252
Per capita rent	24.3067	21.4197	24.7415	26.5002
Household asset variables				
Household crop area ($1000m^2$)	0.7429	0.4021	0.4701	1.2646
Number of tractors	1.0466	1.0156	1.0977	1.0714
Number of vehicles	1.3122	1.1737	1.4051	1.3437
Number of phones	1.6312	1.6790	1.6272	1.5932

Table A.1.2: Difference of a province compared to the rest in 2008

	Ha Tinh - rest	Hue - rest	Dak Lak - rest
Demographic variables			
Household size	-0.6310***	0.3306***	0.2776***
Number of children	-0.5150***	0.1919***	0.3036***
Gender (Female=1)	0.0128	-0.0004	-0.0113
Age	4.4289***	-0.0968	-4.1859***
Ethnicity (Kinh=1)	0.3266***	-0.0243***	-0.2775***
Marital status (Married=1)	0.0252**	-0.0317***	0.0073
Education: Primary school	-0.2030***	0.1116***	0.0842***
Education: Secondary school	0.1378***	-0.0767***	-0.0555***
Education: High school	0.0665***	-0.0331***	-0.0316***
Education: Bachelor or higher	-0.0013	-0.0019	0.0029
Farmer	0.0123	-0.0833***	0.0702***
Member of socio-political organisation	0.2937***	0.0424***	-0.3222***
Percentage of socio-political women	0.0303***	0.0057**	-0.0343***
Income variables (2005 PPP USD)			
Equivalised total income	-282.5260***	-1166.1709***	1391.9664***
Equivalised crop income	-471.9492***	-814.0819***	1216.8655***
Equivalised crop revenue	-910.5426***	-1294.1271***	2084.2165***
Equivalised livestock income	216.9706***	-282.4600***	65.9466*
Equivalised hunting income	-9.2627	86.3482***	-40.3159***
Equivalised off-farm employment income	-31.1010**	4.9294	19.6384
Equivalised remittance: family/friends	158.1308***	-119.1396***	-32.5643***
Consumption variables (2005 PPP USD)			
Per capita total consumption	-99.3506***	-129.8374***	210.6707***
Per capita food consumption	-84.6382***	-37.0169***	110.1112***
Per capita nonfood consumption	-43.4753***	-47.9063***	85.5611***
Per capita education consumption	32.9147***	-29.8319***	-2.7049
Per capita health consumption	-0.0296	-15.2211***	14.2447***
Per capita rent	-4.3061***	0.5162	3.2616***
Household asset variables			
Household crop area (1000m^2)	-0.5079***	-0.3895***	0.8301***
Number of tractors	-0.0607***	0.0565***	0.0424***
Number of vehicles	-0.1957***	0.1301***	0.0531***
Number of phones	0.0696***	-0.0080	-0.0631**

Table A.1.3: Descriptive Statistics by Provinces in 2010 (all values are annual means)

	All	Ha Tinh	Hue	Dak Lak
Demographic variables				
Household size	5.9966	5.5299	6.2042	6.2070
Number of children	1.3722	1.0850	1.4327	1.5653
Gender (Female=1)	0.5081	0.5174	0.5071	0.5009
Age	30.4916	33.5427	30.4741	27.8487
Ethnicity (Kinh=1)	0.7744	0.9988	0.7580	0.5942
Marital status (Married=1)	0.4483	0.4663	0.4243	0.4553
Education: Primary school	0.5772	0.4458	0.6511	0.6218
Education: Secondary school	0.2651	0.3494	0.2193	0.2350
Education: High school	0.1442	0.1899	0.1191	0.1281
Education: Bachelor or higher	0.0135	0.0149	0.0105	0.0150
Farmer	0.3422	0.3456	0.2662	0.4110
Member of socio-political organisation	0.5946	0.7948	0.5459	0.4663
Percentage of socio-political women	0.0548	0.0759	0.0537	0.0376
Income variables (2005 PPP USD)				
Equivalised total income	2031.7635	2120.4835	1902.3493	2076.7590
Equivalised crop income	408.4732	304.2740	146.9917	746.4647
Equivalised crop revenue	6059.0217	3365.6418	2619.9789	11657.4852
Equivalised livestock income	177.1366	306.7841	131.6158	107.1570
Equivalised hunting income	87.5304	33.9747	180.5435	46.2918
Equivalised off-farm employment income	458.2606	391.4401	506.8421	470.5836
Equivalised remittance: family/friends	235.2709	378.1352	251.9602	94.9665
Consumption variables (2005 PPP USD)				
Per capita total consumption	859.5877	761.8707	746.2567	1051.8897
Per capita food consumption	445.2317	391.2846	415.6603	520.2076
Per capita nonfood consumption	305.6685	265.4460	253.0957	390.4231
Per capita education consumption	48.0005	44.3811	37.0315	61.5240
Per capita health consumption	39.4567	38.8109	17.9216	60.3756
Per capita rent	21.5135	21.9481	23.1876	19.5522
Household asset variables				
Household crop area ($1000m^2$)	0.7371	0.3849	0.4870	1.2531
Number of tractors	1.0536	1.0410	1.0608	1.0674
Number of vehicles	1.4274	1.2805	1.4819	1.4933
Number of phones	1.9659	1.7590	2.0421	2.0943

Table A.1.4: Difference of a province compared to the rest in 2010

	Ha Tinh - rest	Hue - rest	Dak Lak - rest
Demographic variables			
Household size	-0.6757***	0.3125***	0.3263***
Number of children	-0.4158***	0.0910***	0.2994***
Gender (Female=1)	0.0135	-0.0014	-0.0112
Age	4.4182***	-0.0263	-4.0975***
Ethnicity (Kinh=1)	0.3250***	-0.0247***	-0.2793***
Marital status (Married=1)	0.0260**	-0.0361***	0.0109
Education: Primary school	-0.1903***	0.1113***	0.0692***
Education: Secondary school	0.1220***	-0.0690***	-0.0467***
Education: High school	0.0662***	-0.0378***	-0.0250***
Education: Bachelor or higher	0.0021	-0.0045*	0.0024
Farmer	0.0050	-0.1144***	0.1067***
Member of socio-political organisation	0.2898***	-0.0733***	-0.1991***
Percentage of socio-political women	0.0305***	-0.0017	-0.0268***
Income variables (2005 PPP USD)			
Equivalised total income	128.4749**	-194.7729***	69.7604
Equivalised crop income	-150.8903***	-393.5385***	524.0172***
Equivalised crop revenue	-3900.2708***	-5175.8764***	8679.7762***
Equivalised livestock income	187.7419***	-68.5103***	-108.4954***
Equivalised hunting income	-77.5537***	139.9879***	-63.9356***
Equivalised off-farm employment income	-96.7625***	73.1168***	19.1053
Equivalised remittance: family/friends	206.8811***	25.1180*	-217.5258***
Consumption variables (2005 PPP USD)			
Per capita total consumption	-141.5035***	-170.5670***	298.1422***
Per capita food consumption	-78.1206***	-44.5060***	116.2414***
Per capita nonfood consumption	-58.2459***	-79.1237***	131.4023***
Per capita education consumption	-5.2413***	-16.5088***	20.9665***
Per capita health consumption	-0.9352	-32.4110***	32.4325***
Per capita rent	0.6294	2.5196***	-3.0407***
Household asset variables			
Household crop area ($1000m^2$)	-0.5269***	-0.3563***	0.8199***
Number of tractors	-0.0250***	0.0081	0.0228***
Number of vehicles	-0.2077***	0.0799***	0.1081***
Number of phones	-0.3104***	0.1118***	0.1971***

A.2 Results Using Household Level Data

This section presents results when using household level data, controlling for household heads' characteristics. The information about age, gender, ethnicity, and education level is for the household head. Gender equals 0 if the household head is male, and 1 if female. Compared to statistics at individual level such as 2.2.2, we can see that household heads are predominantly males. Ethnicity is 1 if the household head belongs to an ethnic majority and 0 if from an ethnic minority. The poverty status is determined using the threshold value of income per capita of 1.90 USD/day adjusted using purchasing power parity (PPP) rates Nguyen et al. (2020b). The education dummy equals 1 if the household head has a university degree or higher. We observe a higher percentages of household heads holding a university degree or higher in year 2013 compared to previous years, though these numbers are still very low.

A.2.1 Summary statistics at the household level

Table A.2.1: Descriptive Statistics by Provinces in 2008 (all values are annual means)

	All	Ha Tinh	Hue	Dak Lak
Household demographics				
Number of females	2.53	2.38	2.59	2.61
Number of female children	0.73	0.56	0.76	0.85
Number of male children	0.73	0.56	0.78	0.84
Age	48.97	52.40	49.46	45.24
Gender	0.16	0.17	0.18	0.15
Ethnicity	0.79	1.00	0.76	0.63
Poverty	0.38	0.34	0.47	0.32
Education dummy	0.01	0.01	0.01	0.02
Household income (2005 PPP USD)				
Crop income	2315.76	1396.15	884.23	4576.02
Livestock income	628.39	996.63	128.36	749.63
Hunting income	200.60	142.37	326.80	104.45
Remittance from family members	342.05	544.03	117.69	358.25
Remittance from friends	181.14	245.52	192.88	108.55
Dwelling income	879.37	687.55	758.58	1184.62
Off-farm employment income	792.68	691.90	831.51	857.77
Total income	6714.53	5938.82	4700.86	9375.12
Household consumption (2005 PPP USD)				
Food consumption	2555.94	2125.88	2529.02	2986.24
Non-food consumption	1828.55	1576.16	1727.02	2156.66
Education consumption	282.99	382.06	186.23	279.57
Health consumption	180.19	171.31	132.13	233.15
Rent	128.53	106.77	135.92	142.28
Total consumption	4975.54	4362.17	4708.24	5797.90

Table A.2.2: Descriptive Statistics by Provinces in 2010 (all values are annual means)

	All	Ha Tinh	Hue	Dak Lak
Number of females	2.68	2.52	2.76	2.75
Number of female children	0.66	0.52	0.68	0.76
Number of male children	0.68	0.50	0.75	0.78
Age	50.73	54.15	50.95	47.19
Gender	0.17	0.18	0.19	0.15
Ethnicity	0.79	1.00	0.76	0.63
Poverty	0.43	0.42	0.47	0.41
Education dummy	0.01	0.01	0.01	0.02
Household income (2005 PPP USD)				
Crop income	1184.59	875.01	429.37	2209.51
Livestock income	518.36	851.35	387.68	316.72
Hunting income	249.43	94.82	527.67	135.22
Remittance from family members	324.83	630.93	253.24	93.10
Remittance from friends	390.96	445.13	535.94	199.36
Dwelling income	119.46	115.63	133.60	109.70
Off-farm employment income	1335.12	1065.37	1524.57	1418.61
Total income	5934.14	5891.37	5690.81	6208.50
Household consumption (2005 PPP USD)				
Food consumption	2463.81	2038.61	2399.58	2922.54
Non-food consumption	1691.50	1382.99	1461.11	2193.41
Education consumption	265.62	231.23	213.78	345.64
Health consumption	218.34	202.21	103.46	339.19
Rent	119.05	114.35	133.86	109.84
Total consumption	4756.76	3969.38	4308.09	5909.54

Table A.2.3: Descriptive Statistics by Provinces in 2013 (all values are annual means)

	All	Ha Tinh	Hue	Dak Lak
Household demographics				
Number of females	2.46	2.25	2.55	2.57
Number of female children	0.52	0.40	0.53	0.61
Number of male children	0.54	0.40	0.58	0.64
Age	53.49	56.28	54.22	50.20
Gender	0.19	0.20	0.20	0.18
Ethnicity	0.79	0.99	0.75	0.62
Poverty	0.43	0.46	0.42	0.43
Education dummy	0.05	0.06	0.04	0.05
Household income (2005 PPP USD)				
Crop income	1270.81	675.73	651.63	2399.37
Livestock income	830.01	1223.67	634.94	640.79
Hunting income	467.51	47.21	1285.56	107.44
Remittance from family members	580.03	1096.98	487.00	181.20
Remittance from friends	412.65	358.03	604.16	287.32
Dwelling income	117.63	113.14	110.55	128.36
Off-farm employment income	1839.37	1557.84	2189.31	1780.71
Total income	7102.03	6161.53	8031.68	7126.75
Household consumption (2005 PPP USD)				
Food consumption	3031.29	2413.78	3228.27	3428.59
Non-food consumption	2264.80	1924.01	2286.53	2564.23
Education consumption	408.00	355.73	383.72	479.38
Health consumption	252.11	259.95	208.41	285.04
Rent	117.63	113.14	110.55	128.36
Total consumption	6071.00	5064.59	6210.76	6885.60

A.2.2 Impact of weather shocks on rural household income

Table A.2.4: Impact of weather shocks on rural household per capita income (ln)

	Total income	Crop income	Crop revenue	Livestock	Hunting	Off Farm	Remittance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Rain shock	-0.0188**	-0.0225	-0.0258**	-0.0065	-0.0869***	-0.0058	-0.0350*
	(0.0086)	(0.0164)	(0.0115)	(0.0136)	(0.0170)	(0.0144)	(0.0207)
Crop area	0.1006***	0.1253**	0.3081***	-0.0388	-0.0522	0.0241	0.1558
	(0.0308)	(0.0487)	(0.0440)	(0.0681)	(0.0549)	(0.0519)	(0.1176)
Age × Age	-0.0010***	-0.0007***	-0.0009***	-0.0003	0.0001	-0.0019***	-0.0006
	(0.0002)	(0.0003)	(0.0002)	(0.0005)	(0.0004)	(0.0005)	(0.0007)
Age	0.1095***	0.0772**	0.0887***	0.0197	-0.0336	0.1999***	0.0375
	(0.0272)	(0.0298)	(0.0261)	(0.0490)	(0.0491)	(0.0600)	(0.0706)
Gender	-0.4095***	-0.1673	-0.1587	-0.4855*	-0.4409	-0.0645	-0.0440
	(0.1376)	(0.2001)	(0.1911)	(0.2588)	(0.2748)	(0.2814)	(0.3441)
Ethnicity	-0.1673	-0.1261	0.2817	0.0274	-0.7212	0.2163	0.9011
	(0.2863)	(0.3664)	(0.4787)	(0.6234)	(0.6000)	(0.3664)	(1.4856)
Household size	-0.0976***	-0.1263***	-0.1461***	-0.1299***	-0.0404	-0.0478	-0.0717
	(0.0292)	(0.0477)	(0.0318)	(0.0491)	(0.0460)	(0.0635)	(0.0734)
Education dummy	0.0043	0.2110	-0.0597	0.0581	0.1350	0.0039	0.5989
	(0.1237)	(0.1869)	(0.1834)	(0.2505)	(0.5517)	(0.2021)	(0.4013)
Constant	4.8082***	4.3526***	4.2743***	5.4161***	5.5495***	-0.4129	3.8529*
	(0.8436)	(0.9686)	(0.8376)	(1.4256)	(1.6180)	(1.7805)	(2.3085)
Observations	4594	4045	4605	3371	2517	2655	2200
Household FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wave FE							

Standard errors (in parentheses) are clustered at the village level. The symbols ***, ** and * denote p<0.01, p<0.05 and p<0.10, respectively.

Table A.2.5: Impact of weather shocks on rural household per capita income (ln) for farmers

	Total income	Crop income	Crop revenue	Livestock	Hunting	Off Farm	Remittance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Rain shock	-0.0157	-0.0350*	-0.0415***	0.0069	-0.0754***	-0.0316*	-0.0204
	(0.0111)	(0.0191)	(0.0126)	(0.0176)	(0.0216)	(0.0190)	(0.0262)
Crop area	0.0944**	0.0907*	0.2996***	-0.0925	-0.0778	0.0366	0.1052
	(0.0373)	(0.0533)	(0.0480)	(0.0773)	(0.0555)	(0.0763)	(0.0956)
Age × Age	-0.0009**	-0.0012***	-0.0010***	0.0000	0.0000	-0.0035***	-0.0005
	(0.0004)	(0.0004)	(0.0003)	(0.0007)	(0.0006)	(0.0008)	(0.0009)
Age	0.0686*	0.0927**	0.0696**	-0.0186	-0.0847	0.2566***	-0.0024
	(0.0372)	(0.0410)	(0.0335)	(0.0636)	(0.0736)	(0.0897)	(0.0864)
Gender	-0.3480*	0.0146	-0.0822	-0.5424	-0.4963	-0.4749	0.1670
	(0.1936)	(0.2058)	(0.2119)	(0.3298)	(0.3254)	(0.3999)	(0.4164)
Ethnicity	-0.3191	-0.2669	0.0343	0.1836	-0.7700	0.3785	0.9532
	(0.3092)	(0.3360)	(0.4942)	(0.6654)	(0.8224)	(0.4767)	(1.4734)
Household size	-0.1177***	-0.1392**	-0.1702***	-0.0811	-0.0674	-0.0608	0.0241
	(0.0411)	(0.0645)	(0.0330)	(0.0582)	(0.0566)	(0.0961)	(0.0830)
Education dummy	-0.2097	0.2075	-0.5000*	0.0605	0.8109***	-1.6040***	0.9609**
	(0.1981)	(0.2280)	(0.2963)	(0.4756)	(0.2304)	(0.5823)	(0.4842)
Constant	6.6682***	4.9638***	5.8588***	6.0528***	8.1482***	0.0598	4.7972**
	(1.1461)	(1.1665)	(0.9666)	(1.7163)	(2.5698)	(2.6607)	(2.4114)
Observations	3390	3040	3455	2566	1981	1882	1596
Household FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wave FE							

Standard errors (in parentheses) are clustered at the village level. The symbols ***, ** and * denote p<0.01, p<0.05 and p<0.10, respectively.

A.2.3 Impact of weather shocks on rural household consumption

Table A.2.6: Impact of weather shocks on rural household per capita consumption (ln)

	Total	Food	NonFood	Education	Health
	(1)	(2)	(3)	(4)	(5)
Rain shock	-0.0044	-0.0030	-0.0003	0.0078	-0.0300*
	(0.0038)	(0.0047)	(0.0047)	(0.0100)	(0.0156)
Crop area	0.0274*	0.0172	-0.0039	0.0805**	0.0015
	(0.0146)	(0.0154)	(0.0167)	(0.0324)	(0.0487)
Age × Age	-0.0002*	-0.0000	-0.0003***	-0.0009**	0.0004
	(0.0001)	(0.0001)	(0.0001)	(0.0003)	(0.0004)
Age	0.0199*	0.0018	0.0316***	0.0855**	-0.0445
	(0.0114)	(0.0125)	(0.0120)	(0.0363)	(0.0433)
Gender	-0.0702	0.0030	-0.1104	0.0726	0.0581
	(0.0691)	(0.0695)	(0.0737)	(0.1877)	(0.2154)
Ethnicity	-0.0724	-0.0662	-0.1942	0.4446***	-0.2207
	(0.0796)	(0.1571)	(0.1303)	(0.1659)	(0.4429)
Household size	-0.1123***	-0.0995***	-0.1241***	-0.2588***	-0.1240***
	(0.0136)	(0.0157)	(0.0145)	(0.0426)	(0.0438)
Education dummy	0.0712	0.0498	-0.0200	-0.0585	0.2090
	(0.0764)	(0.0848)	(0.0749)	(0.1768)	(0.2108)
Constant	7.0598***	6.8038***	5.2974***	2.9087***	4.9213***
	(0.3381)	(0.4057)	(0.3469)	(0.9985)	(1.2450)
Observations	4597	4597	4597	3174	4127
Household FE	Yes	Yes	Yes	Yes	Yes
Wave FE					

Note: Standard errors (in parentheses) are clustered at the village level. The symbols ***, ** and * denote p<0.01, p<0.05 and p<0.10, respectively.

Table A.2.7: Impact of weather shocks on rural household per capita consumption (ln) for farmers

	Total	Food	NonFood	Education	Health
	(1)	(2)	(3)	(4)	(5)
Rain shock	-0.0045	-0.0025	0.0051	0.0031	-0.0417**
	(0.0048)	(0.0056)	(0.0062)	(0.0137)	(0.0190)
Crop area	0.0204	0.0224	-0.0008	0.0801**	-0.0203
	(0.0166)	(0.0175)	(0.0199)	(0.0372)	(0.0638)
Age × Age	-0.0001	0.0002	-0.0001	-0.0012*	0.0004
	(0.0001)	(0.0002)	(0.0002)	(0.0006)	(0.0005)
Age	0.0166	-0.0154	0.0231	0.0951*	-0.0107
	(0.0157)	(0.0166)	(0.0186)	(0.0543)	(0.0538)
Gender	-0.1437	-0.0806	-0.1253	-0.3064	0.1584
	(0.0959)	(0.0868)	(0.1005)	(0.2071)	(0.2876)
Ethnicity	-0.1590**	-0.2121	-0.2322	0.6257***	-0.1680
	(0.0700)	(0.1488)	(0.1445)	(0.1735)	(0.5475)
Household size	-0.1181***	-0.1006***	-0.1362***	-0.2321***	-0.1331**
	(0.0174)	(0.0200)	(0.0179)	(0.0498)	(0.0519)
Education dummy	0.2398*	0.1530	0.1686	0.0456	0.5077
	(0.1254)	(0.1140)	(0.1224)	(0.3260)	(0.3626)
Constant	7.0875***	7.1164***	5.3622***	2.6782**	3.3350**
	(0.4509)	(0.5189)	(0.5075)	(1.3093)	(1.6650)
Observations	3409	3409	3409	2392	3062
Household FE	Yes	Yes	Yes	Yes	Yes
Wave FE					

Note: Standard errors (in parentheses) are clustered at the village level. The symbols ***, ** and * denote p<0.01, p<0.05 and p<0.10, respectively.

Table A.2.8: Heterogeneity test for crop area, farmers, provinces, household size, and ethnicity

	Crop Revenue					Total Income				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Rain shock	-0.0505***	-0.0444***	0.0538**	-0.0375	-0.2011***	-0.0176**	-0.0126	0.0122	0.0031	-0.1102***
	(0.0098)	(0.0149)	(0.0222)	(0.0236)	(0.0253)	(0.0082)	(0.0086)	(0.0182)	(0.0163)	(0.0199)
Crop area	0.5761***					0.2199***				
	(0.0251)					(0.0200)				
Rain shock × Crop area	-0.0292**					-0.0241**				
	(0.0132)					(0.0113)				
Farmer=1 × Rain shock		-0.0471***					-0.0334***			
		(0.0171)					(0.0110)			
Thua Thien Hue × Rain shock			-0.0917***					-0.0337*		
			(0.0249)					(0.0201)		
Dak Lak × Rain shock			-0.2312***					-0.1870***		
			(0.0831)					(0.0683)		
Rain shock × Household size				-0.0066					-0.0064**	
				(0.0040)					(0.0029)	
Ethnicity=1 × Rain shock					0.1400***					0.0853***
					(0.0263)					(0.0205)
Observations	4605	4798	4798	4798	4798	4594	5288	5288	5288	5288
Wave FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors (in parentheses) are clustered at the village level. The symbols ***, ** and * denote p<0.01, p<0.05 and p<0.10, respectively.

A.2.4 Robustness Check

Three standard deviation away from the mean

Table A.2.9 and Table A.2.10 present the results when weather shock is defined as the number of times monthly rainfall is three standard deviation away from the mean.

Rainfall deviation

Table A.2.11 and Table A.2.12 present the results using rainfall deviation which equals the natural log of the year rainfall minus the natural log of mean annual rainfall in a given village.

Table A.2.9: Impact of weather shocks on rural household per capita income (ln) - Rain month

	Total income	Crop income	Crop revenue	Livestock	Hunting	Off Farm	Remittance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Rain month	-0.1449***	0.0397	-0.1802***	-0.0675	-0.4011***	0.0903	-0.3808***
	(0.0406)	(0.0586)	(0.0546)	(0.0673)	(0.0799)	(0.0624)	(0.1035)
Crop area	0.0863***	0.1341***	0.2911***	-0.0442	-0.0717	0.0383	0.1151
	(0.0302)	(0.0495)	(0.0415)	(0.0669)	(0.0533)	(0.0508)	(0.1141)
Age × Age	-0.0011***	-0.0007***	-0.0010***	-0.0004	-0.0000	-0.0019***	-0.0008
	(0.0002)	(0.0003)	(0.0002)	(0.0005)	(0.0004)	(0.0005)	(0.0007)
Age	0.1175***	0.0749**	0.0996***	0.0225	-0.0208	0.1950***	0.0520
	(0.0276)	(0.0301)	(0.0260)	(0.0493)	(0.0488)	(0.0596)	(0.0704)
Gender	-0.3942***	-0.1842	-0.1375	-0.4798*	-0.3518	-0.0975	-0.0432
	(0.1365)	(0.2009)	(0.1922)	(0.2596)	(0.2604)	(0.2796)	(0.3350)
Ethnicity	-0.1343	-0.1358	0.3162	0.0378	-0.5677	0.1663	0.8910
	(0.2952)	(0.3540)	(0.4786)	(0.6227)	(0.6275)	(0.3434)	(1.5181)
Household size	-0.0966***	-0.1265***	-0.1455***	-0.1290***	-0.0413	-0.0475	-0.0680
	(0.0291)	(0.0477)	(0.0320)	(0.0492)	(0.0473)	(0.0632)	(0.0704)
Education dummy	0.0177	0.1929	-0.0485	0.0613	0.1459	-0.0237	0.6822*
	(0.1236)	(0.1842)	(0.1809)	(0.2510)	(0.5494)	(0.2025)	(0.4067)
Constant	4.6673***	4.3240***	4.0684***	5.3772***	5.3143***	-0.3733	3.7348
	(0.8542)	(0.9618)	(0.8336)	(1.4328)	(1.6357)	(1.7662)	(2.3107)
Observations	4594	4045	4605	3371	2517	2655	2200
Household FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wave FE							

Note: "Rain month" is measured by the number of times monthly rainfall is three standard deviation away from the mean. Standard errors (in parentheses) are clustered at the village level. The symbols ***, ** and * denote p<0.01, p<0.05 and p<0.10, respectively.

Table A.2.10: Impact of weather shocks on rural household per capita consumption (ln) - Rain month

	Total	Food	NonFood	Education	Health
	(1)	(2)	(3)	(4)	(5)
Rain month	-0.0926***	-0.0885***	-0.0465**	-0.0738	-0.2217***
	(0.0174)	(0.0208)	(0.0212)	(0.0484)	(0.0630)
Crop area	0.0162	0.0063	-0.0099	0.0700**	-0.0252
	(0.0139)	(0.0153)	(0.0160)	(0.0315)	(0.0486)
Age × Age	-0.0002**	-0.0001	-0.0003***	-0.0010***	0.0002
	(0.0001)	(0.0001)	(0.0001)	(0.0003)	(0.0004)
Age	0.0251**	0.0068	0.0342***	0.0894**	-0.0297
	(0.0115)	(0.0124)	(0.0121)	(0.0358)	(0.0437)
Gender	-0.0549	0.0183	-0.1015	0.0903	0.0750
	(0.0681)	(0.0687)	(0.0735)	(0.1835)	(0.2137)
Ethnicity	-0.0493	-0.0437	-0.1818	0.4524***	-0.1921
	(0.0891)	(0.1631)	(0.1345)	(0.1596)	(0.4256)
Household size	-0.1121***	-0.0993***	-0.1240***	-0.2588***	-0.1222***
	(0.0133)	(0.0155)	(0.0145)	(0.0422)	(0.0427)
Education dummy	0.0838	0.0625	-0.0125	-0.0304	0.2150
	(0.0745)	(0.0836)	(0.0744)	(0.1732)	(0.2061)
Constant	6.9819***	6.7319***	5.2627***	2.9014***	4.6242***
	(0.3396)	(0.4047)	(0.3491)	(0.9927)	(1.2357)
Observations	4597	4597	4597	3174	4127
Household FE	Yes	Yes	Yes	Yes	Yes
Wave FE					

Note: "Rain month" is measured by the number of times monthly rainfall is three standard deviation away from the mean.
Standard errors (in parentheses) are clustered at the village level. The symbols ***, ** and * denote p<0.01, p<0.05 and p<0.10, respectively.

Table A.2.11: Impact of weather shocks on rural household per capita income (ln) - Rainfall deviation

	Total income	Crop income	Crop revenue	Livestock	Hunting	Off Farm	Remittance
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Rainfall deviation	-0.1757	-0.4237*	-0.0320	-0.0102	-1.2857***	-0.5880**	-0.3559
	(0.1691)	(0.2417)	(0.2102)	(0.2719)	(0.4102)	(0.2506)	(0.3757)
Crop area	0.1023***	0.1241**	0.3140***	-0.0378	-0.0524	0.0141	0.1651
	(0.0313)	(0.0491)	(0.0444)	(0.0686)	(0.0568)	(0.0533)	(0.1175)
Age × Age	-0.0010***	-0.0007***	-0.0009***	-0.0003	0.0002	-0.0019***	-0.0006
	(0.0002)	(0.0003)	(0.0002)	(0.0005)	(0.0004)	(0.0005)	(0.0007)
Age	0.1068***	0.0705**	0.0879***	0.0195	-0.0395	0.1875***	0.0315
	(0.0272)	(0.0294)	(0.0261)	(0.0495)	(0.0490)	(0.0592)	(0.0702)
Gender	-0.4116***	-0.1565	-0.1738	-0.4893*	-0.4406	-0.0582	-0.0476
	(0.1373)	(0.2004)	(0.1896)	(0.2598)	(0.2689)	(0.2826)	(0.3419)
Ethnicity	-0.1749	-0.1368	0.2717	0.0275	-0.7102	0.2354	0.8973
	(0.2832)	(0.3660)	(0.4748)	(0.6230)	(0.6156)	(0.4057)	(1.4703)
Household size	-0.0991***	-0.1301***	-0.1460***	-0.1304***	-0.0586	-0.0530	-0.0767
	(0.0292)	(0.0477)	(0.0317)	(0.0492)	(0.0465)	(0.0635)	(0.0739)
Education dummy	0.0021	0.2201	-0.0714	0.0573	0.1268	0.0446	0.5862
	(0.1242)	(0.1867)	(0.1837)	(0.2509)	(0.5199)	(0.2064)	(0.4033)
Constant	4.8717***	4.5582***	4.2469***	5.4102***	5.7646***	0.0067	3.9873*
	(0.8481)	(0.9579)	(0.8281)	(1.4400)	(1.6381)	(1.7791)	(2.3085)
Observations	4594	4045	4605	3371	2517	2655	2200
Household FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wave FE							

Note: "Rainfall deviation" is measured by the natural log of the year rainfall minus the natural log of mean annual rainfall in a given village. Standard errors (in parentheses) are clustered at the village level. The symbols ***, ** and * denote p<0.01, p<0.05 and p<0.10, respectively.

Table A.2.12: Impact of weather shocks on rural household per capita consumption (ln) - Rainfall deviation

	Total	Food	NonFood	Education	Health
	(1)	(2)	(3)	(4)	(5)
Rainfall deviation	-0.2182***	-0.2443***	-0.1364	-0.6026***	0.3485
	(0.0713)	(0.0873)	(0.0882)	(0.1867)	(0.2576)
Crop area	0.0262*	0.0155	-0.0051	0.0693**	0.0100
	(0.0143)	(0.0152)	(0.0166)	(0.0316)	(0.0493)
Age × Age	-0.0001	0.0000	-0.0003**	-0.0009**	0.0003
	(0.0001)	(0.0001)	(0.0001)	(0.0003)	(0.0004)
Age	0.0166	-0.0019	0.0295**	0.0795**	-0.0382
	(0.0114)	(0.0124)	(0.0118)	(0.0356)	(0.0436)
Gender	-0.0623	0.0130	-0.1039	0.0996	0.0321
	(0.0682)	(0.0685)	(0.0729)	(0.1848)	(0.2176)
Ethnicity	-0.0728	-0.0659	-0.1934	0.4424***	-0.2516
	(0.0808)	(0.1577)	(0.1312)	(0.1592)	(0.4410)
Household size	-0.1144***	-0.1018***	-0.1254***	-0.2626***	-0.1216***
	(0.0133)	(0.0152)	(0.0144)	(0.0415)	(0.0443)
Education dummy	0.0811	0.0620	-0.0123	-0.0322	0.1800
	(0.0752)	(0.0832)	(0.0738)	(0.1717)	(0.2108)
Constant	7.1755***	6.9376***	5.3753***	3.1844***	4.6619***
	(0.3360)	(0.4013)	(0.3419)	(0.9893)	(1.2387)
Observations	4597	4597	4597	3174	4127
Household FE	Yes	Yes	Yes	Yes	Yes
Wave FE					

Note: "Rainfall deviation" is measured by the natural log of the year rainfall minus the natural log of mean annual rainfall in a given village.
Standard errors (in parentheses) are clustered at the village level. The symbols ***, ** and * denote p<0.01, p<0.05 and p<0.10, respectively.

Appendix B

Appendix for Chapter 3

B.1 General Health and Well-being Questionnaires

PART A: GENERAL HEALTH AND WELL-BEING
(SF-36 Health Survey)

This first set of questions seeks your views about your health, how you feel and how well you are able to do your usual activities.

Please take the time to read and answer each question carefully by crossing the box corresponding to your response. If you are unsure about how to answer a question, please give the best answer you can.

A1 In general, would you say your health is:

*(Cross ☒ **ONE** box)*

☐ Excellent 1	☐ Very good 2	☐ Good 3	☐ Fair 4	☐ Poor 5	tgh1

A2 <u>Compared to one year ago</u>, how would you rate your health in general <u>now</u>?

*(Cross ☒ **ONE** box)*

tgh2

- ☐ 1 Much better now than a year ago
- ☐ 2 Somewhat better now than a year ago
- ☐ 3 About the same as one year ago
- ☐ 4 Somewhat worse now than one year ago
- ☐ 5 Much worse now than one year ago

A3 The following questions are about activities you might do during a typical day. Does <u>your health now limit you</u> in these activities? If so, how much?

*(Cross ☒ **ONE** box on **EACH** line)*

	ACTIVITIES	*Yes, limited a lot*	*Yes, limited a little*	*No, not limited at all*	
a	<u>Vigorous activities</u>, such as running, lifting heavy objects, participating in strenuous sports	☐ 1	☐ 2	☐ 3	tgh3a
b	<u>Moderate activities</u>, such as moving a table, pushing a vacuum cleaner, bowling or playing golf	☐ 1	☐ 2	☐ 3	tgh3b
c	Lifting or carrying groceries	☐ 1	☐ 2	☐ 3	tgh3c
d	Climbing <u>several</u> flights of stairs	☐ 1	☐ 2	☐ 3	tgh3d
e	Climbing <u>one</u> flight of stairs	☐ 1	☐ 2	☐ 3	tgh3e
f	Bending, kneeling, or stooping	☐ 1	☐ 2	☐ 3	tgh3f
g	Walking <u>more than one kilometre</u>	☐ 1	☐ 2	☐ 3	tgh3g
h	Walking <u>half a kilometre</u>	☐ 1	☐ 2	☐ 3	tgh3h
i	Walking <u>100 metres</u>	☐ 1	☐ 2	☐ 3	tgh3i
j	Bathing or dressing yourself	☐ 1	☐ 2	☐ 3	tgh3j

A4 During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of your physical health?

(Cross ☒ ONE box on EACH line)

		YES	NO	
a	Cut down the amount of time you spent on work or other activities	☐	☐	tgh4a
b	Accomplished less than you would like	☐	☐	tgh4b
c	Were limited in the kind of work or other activities	☐	☐	tgh4c
d	Had difficulty performing the work or other activities (for example, it took extra effort)	☐	☐	tgh4d

A5 During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of any emotional problems (such as feeling depressed or anxious)?

(Cross ☒ ONE box on EACH line)

		YES	NO	
a	Cut down the amount of time you spent on work or other activities	☐	☐	tgh5a
b	Accomplished less than you would like	☐	☐	tgh5b
c	Didn't do work or other activities as carefully as usual	☐	☐	tgh5c

A6 During the past 4 weeks, to what extent has your physical health or emotional problems interfered with your normal social activities with family, friends, neighbours, or groups?

(Cross ☒ ONE box)

☐ 1 Not at all	☐ 2 Slightly	☐ 3 Moderately	☐ 4 Quite a bit	☐ 5 Extremely	tgh6

A7 How much bodily pain have you had during the past 4 weeks?

(Cross ☒ ONE box)

☐ 1 No bodily pain	☐ 2 Very mild	☐ 3 Mild	☐ 4 Moderate	☐ 5 Severe	☐ 6 Very severe	tgh7

A8 During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and housework)?

(Cross ☒ ONE box)

☐ 1 Not at all	☐ 2 Slightly	☐ 3 Moderately	☐ 4 Quite a bit	☐ 5 Extremely	tgh8

A9 **These questions are about how you feel and how things have been with you during the past 4 weeks. For each question, please give the one answer that comes closest to the way you have been feeling.**

How much of the time during the past 4 weeks:

(Cross ☒ ONE box on EACH line)

		All of the time	Most of the time	A good bit of the time	Some of the time	A little of the time	None of the time	
a	Did you feel full of life?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	tgh9a
b	Have you been a nervous person?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	tgh9b
c	Have you felt so down in the dumps that nothing could cheer you up?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	tgh9c
d	Have you felt calm and peaceful?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	tgh9d
e	Did you have a lot of energy?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	tgh9e
f	Have you felt down?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	tgh9f
g	Did you feel worn out?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	tgh9g
h	Have you been a happy person?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	tgh9h
i	Did you feel tired?	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	tgh9i

A10 **During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting friends, relatives, etc.)?**

(Cross ☒ ONE box)

tgh10

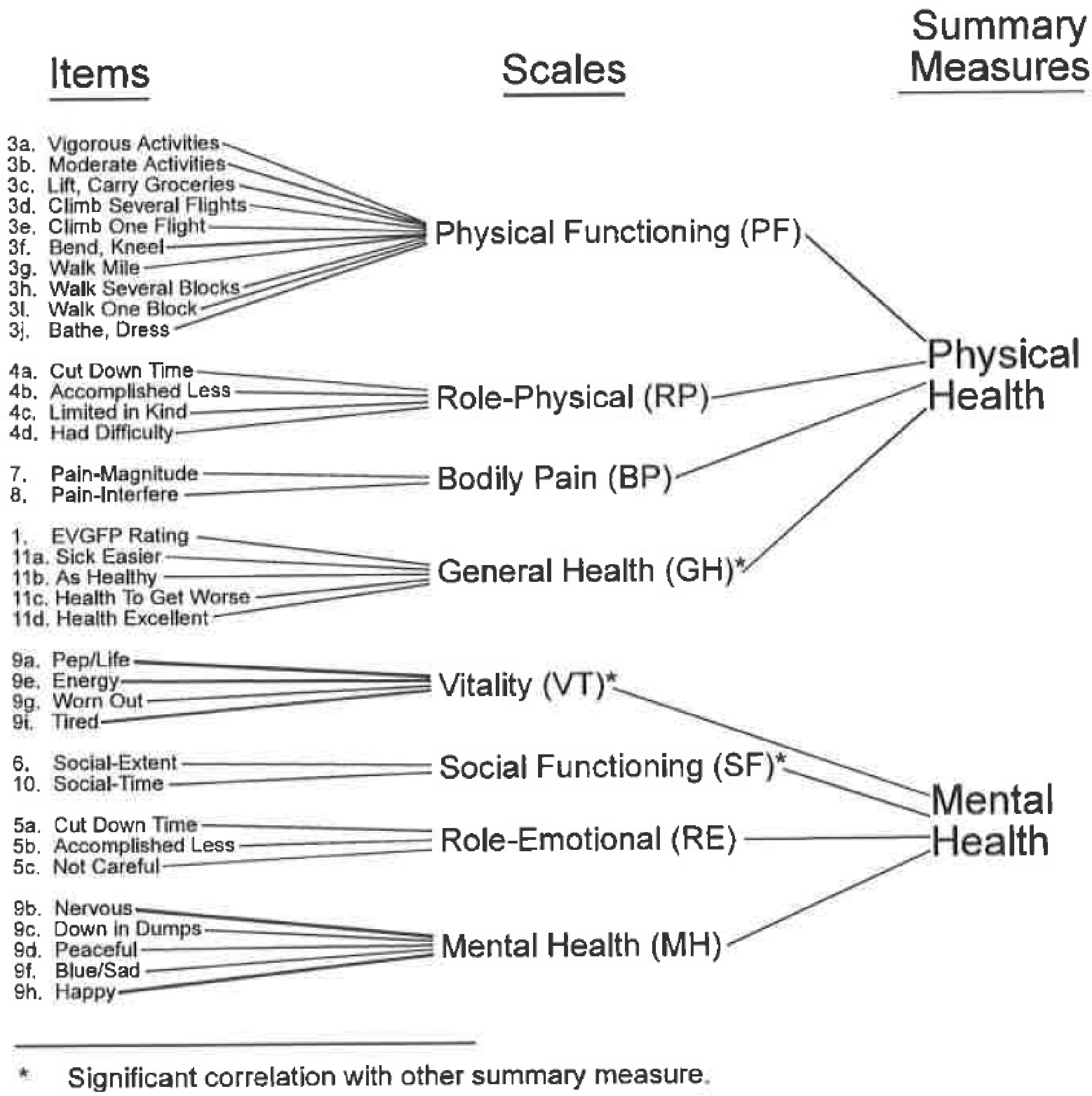
- ☐ 1 All of the time
- ☐ 2 Most of the time
- ☐ 3 Some of the time
- ☐ 4 A little of the time
- ☐ 5 None of the time

A11 **How TRUE or FALSE is each of the following statements for you?**

(Cross ☒ ONE box on EACH line)

		Definitely True	Mostly True	Don't know	Mostly False	Definitely False	
a	I seem to get sick a little easier than other people	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	tgh11a
b	I am as healthy as anybody I know	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	tgh11b
c	I expect my health to get worse	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	tgh11c
d	My health is excellent	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	tgh11d

B.2 SF-36 Measurement Model



Source: Ware et al. (1994)

B.3 Distributional Statistics

To study social welfare, inequality, poverty or other measures that describe the distributional characteristics of an outcome of interest, we need access to one of the following types of data. First, it is common to have access to the full set of values corresponding to each observation in the population or sample. If we have a finite sample of size n, then the full set of values is $Y = [y_1, \cdots, y_n]$ where each y_i is the outcome of interest of individual ith. In our context, y_i is a mental health measure (i.e. SF-36 mental health scores).

If we do not have the information about the mental health level of each individual in the sample, we can still derive any distributional statistic of mental health if we know the relative position of all individuals compared to the rest of the population or how frequent it is to observe any given level of mental health. This means that we either know the cumulative distribution function ($F_Y()$) or the probability density function ($f_Y()$) of mental health. The vector of information can be written as a set of ordered pairs, $F_Y = [(y, F_Y(y))|y \in \mathbb{R}]$ or $f_Y = [(y, f_Y(y))|y \in \mathbb{R}]$. Note that the range of y might be restricted depending on how it is constructed.

Let $v(.)$ be a functional that uses the information contained in Y, F_Y, or f_Y to estimate a distributional statistic of Y. This functional can be used to estimate statistics relevant to policy analysis such as inequality indices. To measure the impact a change in the distribution of mental health will have on the inequality indices, we can compare the observed distribution F_Y with the ex-post distribution G_Y. The change in the distributional statistic generated by a change in distribution from $F_Y \to G_Y$ is

$$\Delta v = v(G_Y) - v(F_Y). \tag{B.1}$$

B.4 Influence Function

The influence function of a statistics tells us how an individual observation influences that statistics. The influence functions can be used to measure the robustness of functionals to data outliers (Hampel, 1974).

The magnitude of this change depends on the magnitude of the change from $F_Y() \to G_Y()$. For example, changes to the distribution caused by one additional person will be larger for a small population compared to adding this person to a large population. Therefore, we might want to standardize the change in the statistic Δv with respect to some measure that quantifies the change of the distribution

$$\Delta^s v = \frac{\Delta v}{\Delta(G_Y - F_Y)} = \frac{v(G_Y) - v(F_Y)}{\Delta(G_Y - F_Y)} \tag{B.2}$$

B.4.1 Properties of Influence Functions

1. $\int \text{IF}\left(y; v\left(F_Y\right)\right) dF_Y = 0$

2. $v\left(F_y\right) \sim N\left(v\left(F_Y\right), \frac{\sigma^2_{\text{IF}}}{N}\right)$

3. $\sigma^2_{\text{IF}} = \int \text{IF}\left(y; v\left(F_Y\right)\right)^2 dF_Y$

B.4.2 Calculating influence functions

The easiest way to calculate influence functions is via directional derivatives. Consider families of distributions indexed by a parameter $\theta \in \mathbb{R} : P(X;\theta)$. Then we get that

$$\phi(P(.;\theta)) : \mathbb{R} \to \mathbb{R}$$

is a function we can easily differentiate. Next, do some algebra to get the resulting expression into the form

$$\frac{\partial \phi}{\partial \theta} = \int \mathrm{IF}(X) \frac{\partial}{\partial \theta} dP(X;\theta)$$

Finally, normalize, by adding a constant, so that IF has a mean zero.

Example:

$$\phi(P) = \mathrm{Var}(X) = \int X^2 dP - \left(\int X dP \right)^2$$

Thus

$$\begin{aligned} \frac{\partial \phi}{\partial \theta} &= \int X^2 \frac{\partial}{\partial \theta} dP - 2 \left(\int X dP \right) \int X \frac{\partial}{\partial \theta} dP \\ &= \int \left(X^2 - 2E[X] \cdot X \right) \frac{\partial}{\partial \theta} dP \end{aligned}$$

Normalizing, as required, to $E[\mathrm{IF}] = 0$, we get

$$\mathrm{IF}(X) = X^2 - 2E[X] \cdot X - \mathrm{Var}(X) + E[X]^2$$

B.4.3 Influence Function of Quantiles

If F is a distribution function, then the quantile function corresponding to F is given by:

$$q(\tau) = \inf\{x : F(x) \geq \tau\} \text{ where } \tau \in (0,1) \tag{B.3}$$

We do not define $q(\tau) = F^{-1}(\tau)$ because F might be discontinuous or flat. Quantile function is non-decreasing, left-continuous and has limits to the right.

Assume that Y with distribution F_Y is observed in the presence of covariates X with distribution F_X and support $\mathcal{X} \in \mathbb{R}^k$. We have a joint distribution $F_{Y,X} : \mathbb{R} \times \mathcal{X} \to [0,1]$. We would like to learn the effect of a small change in the location of the F_X on the τth quantile of F_Y. By definition, we have

$$F_Y(y) = \int F_{Y|X}(y|X = x) \cdot dF_X(x) \tag{B.4}$$

Let the shifted distribution of X be G_X and assume that the conditional distribution $F_{Y|X}$ does not change with the small change in the distribution of X from F_X to G_X, then the corresponding distribution of Y can be obtained by replacing F_X by G_X. Thus

$$G_Y^*(y) \equiv \int F_{Y|X}(y|X = x) \cdot dG_X(x) \tag{B.5}$$

For the τth quantile, the influence function is

$$\mathrm{IF}(Y, q_\tau, F_Y) = \frac{\tau - \mathbb{1}\{Y \le q_\tau\}}{f_Y(q_{\tau)}} \tag{B.6}$$

The recentered influence function (RIF) is then

$$\mathrm{RIF}(Y, q_\tau, F_Y) = q_\tau + \frac{\tau - \mathbb{1}\{Y \le q_\tau\}}{f_Y(q_\tau)} \tag{B.7}$$

B.5 RIF Regression

$$\mathrm{RIF}\left(y_i; v\left(F_Y\right)\right) = X_i{}'\beta + \varepsilon_i, E\left(\varepsilon_i\right) = 0 \tag{B.8}$$

RIF-OLS[1] differs from the standard OLS model in the sense that it uses the estimated $\mathrm{RIF}(y_i; v(F_Y))$ for each observation y_i in the data as the dependent variable and regresses it against all the variables of interest.

The interpretation is also slightly different. In the standard OLS, the typical interpretation of the coefficients is that a one-unit increase in X will increase y in β units on average, everything else held constant. In the RIF-OLS model, to obtain the unconditional partial effect on the statistic v, we will need to obtain unconditional expectations on both sides of equation (B.8):

$$v\left(F_Y\right) = E\left(\mathrm{RIF}\left(y_i; v\left(F_Y\right)\right)\right) = E\left(X_i'\beta\right) + E\left(\varepsilon_i\right) = \bar{X}'\beta \tag{B.9}$$

Then the unconditional partial effect is calculated as

$$\beta = \frac{\partial v\left(F_Y\right)}{\partial \bar{X}_k} \tag{B.10}$$

As a result, the interpretation of the unconditional partial effect is that if the distribution of x_k changes such that the unconditional average increases by one unit $\left(\Delta\bar{X}_k = 1\right)$, then the expected change in the distributional statistic v is equal to β. This interpretation also implies that if we use linear regression, we can only capture one aspect of change in the distribution of X which is the unconditional mean. We could, however, include higher-order polynomials and interactions to capture nonlinear relationships across explanatory variables. One example is to specify

$$\mathrm{RIF}\left(y_i; v\left(F_Y\right)\right) = \beta_0 + \beta_1 X_i + \beta_2\left(X_i - \bar{X}\right)^2 + \varepsilon_i$$

Then,

$$v\left(F_Y\right) = \beta_0 + \beta_1\bar{X} + \beta_2 E\left(\left(X_i - \bar{X}\right)^2\right) = \beta_0 + \beta_1\bar{X} + \beta_2\,\mathrm{Var}(X).$$

Hence, we can capture the changes in both the mean and the variance of the unconditional distribution of X (Rios Avila, 2019).

[1] Here OLS stands for Ordinary Least Squares.

B.6 RIF Decomposition

Let assume a joint distribution function $f_{Y,X,T}(y_i, x_i, t_i)$. In our context, Y is mental health scores, X are the explanatory variables (i.e. demographic, economic and life event variables), and T represents two groups of men and women. By definition, we can write

$$f^k_{Y,X}(y,x) = f^k_{Y|X}(Y \mid X) f^k_X(X)$$
$$F^k_Y(y) = \int F^k_{Y|X}(Y \mid X) dF^k_X(X)$$

where $k \in [0,1]$ indicates that the density is conditional on $T = k$. To calculate the gap in a distributional statistic v between two groups, we can use the cumulative conditional distribution of Y:

$$\Delta v = v_1 - v_0 = v\left(F^1_Y\right) - v\left(F^0_Y\right) = v\left(\int F^1_{Y|X}(Y \mid X) dF^1_X(X)\right) - v\left(\int F^0_{Y|X}(Y \mid X) dF^0_X(X)\right) \tag{B.11}$$

Equation (B.11) shows that differences in the statistics of interest Δv can arise either because of differences in the distribution of the explanatory variables X or because of differences in the relationships between the dependent variable Y and the explanatory variables X.

B.7 Full RIF Regression Results for Quantiles

Table B.7.1: Full RIF Regression Results for Men

	Q10	Q25	Q50	Q75	Q90
	(1)	(2)	(3)	(4)	(5)
Marital status (ref: Single)					
Marital Status: Legally married or de facto	3.8047***	4.9766***	2.8572***	1.5560***	0.9726***
	(5.48)	(7.29)	(7.57)	(5.46)	(4.02)
Marital Status: Divorced/separated/widowed	-1.4989	-1.1898	-0.1629	-0.2391	-0.3037
	(-1.35)	(-1.08)	(-0.27)	(-0.50)	(-0.73)
Household size	0.5505***	0.5949***	0.2391***	0.2279***	0.1679***
	(3.82)	(4.00)	(2.61)	(3.10)	(2.69)
Age	-0.7049***	-0.8411***	-0.3565***	-0.1880***	-0.1621***
	(-9.20)	(-11.26)	(-8.50)	(-5.53)	(-5.39)
Age square	0.0091***	0.0103***	0.0050***	0.0033***	0.0028***
	(11.24)	(13.24)	(11.52)	(9.34)	(8.56)
Living outside a major city	2.0662***	1.4209**	0.9259**	0.7297**	0.7969***
	(3.42)	(2.15)	(2.32)	(2.16)	(2.66)
Young family	-1.0371**	-1.4837***	-0.6037*	-0.5062*	-0.3002
	(-2.01)	(-2.72)	(-1.81)	(-1.94)	(-1.35)
Foreign born	-0.1559	-0.9857*	-1.0345***	-0.5595*	-0.2137
	(-0.30)	(-1.73)	(-2.92)	(-1.89)	(-0.82)
Log of after-tax equivalised income	2.5925***	3.4689***	2.4390***	1.5285***	0.9330***
	(8.27)	(10.80)	(12.53)	(9.13)	(6.05)
Labour force status (ref: Not in the labour force)					
Labour force status: Employed	9.5647***	8.7401***	3.5972***	1.8173***	1.1235***
	(13.94)	(14.51)	(11.05)	(6.83)	(4.66)
Labour force status: Unemployed	0.6572	-0.9490	-0.0283	-0.0060	-0.1953
	(0.63)	(-1.08)	(-0.07)	(-0.02)	(-0.74)

Education (ref: Year 12 or below)					
Education: Certificates and diploma	-0.0185	0.8665	0.6038*	0.2260	0.1781
	(-0.03)	(1.55)	(1.82)	(0.83)	(0.73)
Education: Undergraduate	1.0112*	2.2119***	0.1304	-0.5425	-0.7760***
	(1.67)	(3.36)	(0.31)	(-1.56)	(-2.60)
Education: Postgraduate	0.9662	2.7461***	0.6920	-0.3455	-0.7807*
	(1.11)	(2.85)	(1.07)	(-0.64)	(-1.68)
Death of spouse or child	-8.6758***	-8.9625***	-4.9791***	-2.8283***	-2.0984***
	(-3.80)	(-5.01)	(-5.99)	(-4.33)	(-3.74)
Death of close relative or family member	-1.6002***	-1.7416***	-0.6258***	-0.4953***	-0.3902**
	(-3.89)	(-4.50)	(-2.90)	(-2.89)	(-2.54)
Serious injury or illness to family member	-3.2562***	-3.7891***	-2.3313***	-1.7381***	-1.5022***
	(-7.40)	(-9.13)	(-10.03)	(-9.34)	(-9.42)
Natural disaster damaged/destroyed home	-3.7757***	-4.3044***	-1.8353***	-1.3340***	-0.8970**
	(-3.26)	(-4.21)	(-3.34)	(-3.01)	(-2.29)
Fired or made redundant	-3.6356***	-3.3375***	-1.9318***	-1.5969***	-1.2075***
	(-4.72)	(-4.92)	(-5.29)	(-5.81)	(-5.18)
Victim of physical violence	-17.4942***	-12.3104***	-5.5307***	-2.8199***	-1.6649***
	(-10.38)	(-9.95)	(-10.00)	(-7.03)	(-5.01)
Close family member detained in jail	-2.6129*	-3.7574***	-1.8020***	-0.8495*	-0.6998*
	(-1.87)	(-3.00)	(-2.78)	(-1.72)	(-1.68)
Detained in jail	-16.0369***	-10.6467***	-3.1745***	-0.8848	-0.6286
	(-5.15)	(-4.62)	(-3.27)	(-1.21)	(-1.04)
Residential area index of socioeconomic advantage (ref: 1st quintile)					
2nd quintile	1.4266**	2.1901***	1.0504***	0.5361*	0.3798
	(2.13)	(3.17)	(2.66)	(1.69)	(1.35)
3rd quintile	2.2640***	3.5959***	1.4417***	0.5505*	0.1826
	(3.52)	(5.28)	(3.59)	(1.69)	(0.64)
4th quintile	2.8677***	4.3004***	1.5757***	0.7256**	0.5677*

	(4.41)	(6.32)	(3.86)	(2.16)	(1.91)
5th quintile	3.4326***	4.8127***	2.0495***	0.5273	0.1349
	(5.21)	(6.90)	(4.88)	(1.54)	(0.45)
Constant	25.2133***	31.4837***	54.1135***	71.2021***	84.0031***
	(7.34)	(8.84)	(25.47)	(39.22)	(50.42)
Observations	82975	82975	82975	82975	82975
R^2	0.0447	0.0557	0.0467	0.0383	0.0270
WE	Yes	Yes	Yes	Yes	Yes

Q10, Q25, Q50, Q75, and Q90 are the 10th, 25th, 50th, 75th and 90th quantiles. Clustered robust standard errors at the individual level. The symbols ***, **, and * denote that p<0.01, p<0.05, and p<0.10, respectively.

Table B.7.2: Full RIF Regression Results for Women

	Q10	Q25	Q50	Q75	Q90
	(1)	(2)	(3)	(4)	(5)
Marital status (ref: Single)					
Marital Status: Legally married or de facto	4.9545***	4.2649***	2.2664***	0.9987***	-0.1077
	(6.66)	(7.33)	(5.54)	(3.30)	(-0.40)
Marital Status: Divorced/separated/widowed	-0.3745	-0.6511	-0.7690	-0.8817**	-1.1286***
	(-0.37)	(-0.80)	(-1.33)	(-2.00)	(-2.71)
Household size	0.2737	0.0296	-0.0286	-0.0209	-0.0421
	(1.64)	(0.21)	(-0.27)	(-0.27)	(-0.61)
Age	-0.4264***	-0.3678***	-0.0930**	0.0609*	0.0581*
	(-5.80)	(-6.20)	(-2.14)	(1.83)	(1.82)
Age square	0.0073***	0.0063***	0.0028***	0.0011***	0.0009***
	(9.71)	(10.35)	(6.37)	(3.20)	(2.59)
Living outside a major city	1.5188**	1.0120*	1.0443**	0.7613**	0.8469**
	(2.20)	(1.75)	(2.29)	(2.15)	(2.48)
Young family	3.8730***	3.2544***	1.9309***	0.8988***	0.3934*

	(6.69)	(6.91)	(5.50)	(3.37)	(1.75)
Foreign born	0.6676	-0.8427*	-0.9383**	-0.6334**	-0.4255
	(1.23)	(-1.72)	(-2.43)	(-2.10)	(-1.54)
Log of after-tax equivalised income	1.9719***	2.8623***	2.5637***	1.6515***	1.0638***
	(5.98)	(10.22)	(11.93)	(9.65)	(6.49)
Labour force status (ref: Not in the labour force)					
Labour force status: Employed	8.6687***	6.7399***	3.0664***	1.3093***	0.2621
	(14.76)	(14.93)	(9.59)	(5.32)	(1.12)
Labour force status: Unemployed	-1.4798	-2.2026***	-1.2901***	-0.6314*	-0.4006
	(-1.29)	(-2.73)	(-2.61)	(-1.81)	(-1.39)
Education (ref: Year 12 or below)					
Education: Certificates and diploma	1.2768**	1.3045***	1.0079***	0.6144**	0.2557
	(2.23)	(2.66)	(2.73)	(2.13)	(0.93)
Education: Undergraduate	2.4511***	3.1718***	1.5967***	0.2224	-0.6528**
	(4.20)	(6.25)	(3.88)	(0.69)	(-2.21)
Education: Postgraduate	2.8352***	2.4625***	0.5771	-0.8050	-1.5130***
	(3.28)	(3.13)	(0.83)	(-1.48)	(-3.20)
Death of spouse or child	-10.2787***	-10.2712***	-5.7620***	-3.8009***	-3.0206***
	(-5.46)	(-7.94)	(-7.76)	(-7.38)	(-6.69)
Death of close relative or family member	-2.5401***	-2.2527***	-1.4745***	-1.0125***	-0.7926***
	(-5.71)	(-6.71)	(-6.39)	(-5.92)	(-5.05)
Serious injury or illness to family member	-4.5806***	-4.0107***	-2.8514***	-2.2031***	-1.8147***
	(-10.14)	(-11.72)	(-12.08)	(-12.75)	(-11.60)
Natural disaster damaged/destroyed home	-4.2828***	-3.5856***	-2.2729***	-1.9417***	-0.9131**
	(-3.33)	(-3.68)	(-3.66)	(-4.23)	(-2.16)
Fired or made redundant	-7.9633***	-5.9125***	-4.4727***	-2.8502***	-1.5674***
	(-6.59)	(-7.23)	(-9.17)	(-8.65)	(-5.72)
Victim of physical violence	-25.8405***	-16.8173***	-7.8593***	-4.0886***	-1.8104***
	(-13.56)	(-14.71)	(-13.18)	(-10.67)	(-5.65)

Close family member detained in jail	-4.6872***	-3.5495***	-2.4150***	-1.2543**	-0.6944
	(-3.13)	(-3.40)	(-3.49)	(-2.46)	(-1.61)
Detained in jail	-3.5027	-2.5117	-3.2023*	-1.1419	0.3133
	(-0.57)	(-0.70)	(-1.96)	(-1.08)	(0.31)
Residential area index of socioeconomic advantage (ref: 1st quintile)					
2nd quintile	2.1634***	2.5254***	2.1729***	1.1692***	0.4301
	(2.97)	(4.23)	(5.04)	(3.54)	(1.37)
3rd quintile	3.0514***	3.6412***	2.5524***	1.3348***	0.4614
	(4.29)	(6.19)	(5.76)	(3.94)	(1.45)
4th quintile	3.0194***	3.5146***	2.5589***	1.2331***	0.3324
	(4.24)	(5.91)	(5.72)	(3.58)	(1.03)
5th quintile	4.7782***	5.0572***	3.7012***	1.6157***	0.6462*
	(6.98)	(8.49)	(8.01)	(4.47)	(1.91)
Constant	20.9046***	27.3819***	45.1824***	64.3291***	78.8634***
	(5.76)	(8.89)	(19.25)	(34.70)	(44.06)
Observations	94337	94337	94337	94337	94337
R^2	0.0495	0.0617	0.0520	0.0453	0.0304
Year FE	Yes	Yes	Yes	Yes	Yes

Q10, Q25, Q50, Q75, and Q90 are the 10th, 25th, 50th, 75th and 90th quantiles. Clustered robust standard errors at the individual level. The symbols ***, **, and * denote that p<0.01, p<0.05, and p<0.10, respectively.

B.8 RIF Regression Results for Other Inequality Measures

Table B.8.1: RIF Regression - Other Inequality Measures

	IQR(90 10)	IQRATIO(90 10)	IQSR(10 90)	GINI x 100	Variance
	(1)	(2)	(3)	(4)	(5)
Marital status (ref: Single)					

Marital Status: Legally married or de facto	-4.0176***	-0.1401***	-0.4306***	-0.0205***	-67.6891***
	(-12.48)	(-13.38)	(-14.32)	(-19.34)	(-15.56)
Marital Status: Divorced/separated/widowed	-0.0210	0.0086	-0.0221	0.0002	-4.5283
	(-0.05)	(0.60)	(-0.52)	(0.12)	(-0.75)
Household size	-0.2704***	-0.0103***	-0.0325***	-0.0014***	-4.5921***
	(-3.53)	(-4.14)	(-4.62)	(-5.42)	(-4.47)
Age	0.5357***	0.0185***	0.0586***	0.0026***	9.5571***
	(16.75)	(17.87)	(19.36)	(24.35)	(21.82)
Age square	-0.0067***	-0.0002***	-0.0008***	-0.0000***	-0.1180***
	(-20.47)	(-23.55)	(-24.87)	(-31.57)	(-26.06)
Living outside a major city	-1.0664***	-0.0478***	-0.1435***	-0.0056***	-16.7178***
	(-3.76)	(-5.27)	(-5.81)	(-6.12)	(-4.58)
Young family	-1.7171***	-0.0577***	-0.1719***	-0.0070***	-26.5080***
	(-5.70)	(-5.94)	(-6.56)	(-7.15)	(-6.87)
Foreign born	-0.6650***	-0.0171**	-0.0476**	0.0014*	-4.5153
	(-2.96)	(-2.39)	(-2.47)	(1.88)	(-1.57)
Log of after-tax equivalised income	-1.4418***	-0.0629***	-0.1852***	-0.0116***	-27.7587***
	(-8.17)	(-11.10)	(-11.85)	(-20.04)	(-12.10)
Labour force status (ref: Not in the labour force)					
Labour force status: Employed	-8.5981***	-0.2936***	-0.8921***	-0.0376***	-138.5882***
	(-30.17)	(-31.74)	(-31.89)	(-39.29)	(-34.46)
Labour force status: Unemployed	0.0636	0.0085	-0.1390**	0.0006	-15.4110*
	(0.09)	(0.38)	(-2.15)	(0.29)	(-1.69)
Education (ref: Year 12 or below)					
Education: Certificates and diploma	0.0079	-0.0029	-0.0125	-0.0021***	-2.6380
	(0.03)	(-0.39)	(-0.60)	(-2.76)	(-0.86)
Education: Undergraduate	-2.0995***	-0.0583***	-0.1359***	-0.0104***	-34.5714***
	(-8.20)	(-7.13)	(-6.18)	(-12.55)	(-10.54)
Education: Postgraduate	-2.2565***	-0.0572***	-0.0943***	-0.0090***	-30.4294***

	(-5.73)	(-4.58)	(-2.89)	(-7.04)	(-6.14)
Death of spouse or child	7.5568***	0.2887***	0.9074***	0.0427***	137.4624***
	(5.32)	(6.17)	(5.96)	(8.92)	(6.48)
Death of close relative or family member	1.7698***	0.0675***	0.1816***	0.0086***	26.5748***
	(5.69)	(6.70)	(6.28)	(8.43)	(6.35)
Serious injury or illness to family member	2.7667***	0.1166***	0.3194***	0.0150***	41.2445***
	(9.45)	(12.27)	(11.76)	(15.60)	(10.52)
Natural disaster damaged/destroyed home	3.0436***	0.1146***	0.3606***	0.0172***	55.3957***
	(3.42)	(3.95)	(4.13)	(5.80)	(4.39)
Fired or made redundant	4.4993***	0.1703***	0.5033***	0.0216***	69.3216***
	(6.94)	(8.01)	(7.91)	(10.02)	(7.65)
Victim of physical violence	21.1554***	0.7246***	2.3987***	0.0898***	354.0504***
	(17.01)	(17.60)	(16.07)	(20.35)	(17.08)
Close family member detained in jail	2.8812***	0.1049***	0.2308**	0.0145***	40.8986***
	(3.11)	(3.45)	(2.53)	(4.79)	(3.15)
Detained in jail	10.5287***	0.3563***	1.1434***	0.0483***	192.1825***
	(3.62)	(3.69)	(3.24)	(4.71)	(3.88)
Residential area index of socioeconomic advantage (ref: 1st quintile)					
2nd quintile	-1.4552***	-0.0545***	-0.1299***	-0.0092***	-24.3872***
	(-4.50)	(-5.21)	(-4.39)	(-8.78)	(-5.67)
3rd quintile	-2.3135***	-0.0816***	-0.1905***	-0.0138***	-37.9789***
	(-7.28)	(-7.95)	(-6.60)	(-13.33)	(-9.02)
4th quintile	-2.5046***	-0.0902***	-0.2053***	-0.0147***	-40.7123***
	(-7.92)	(-8.85)	(-7.20)	(-14.27)	(-9.76)
5th quintile	-3.8469***	-0.1331***	-0.3166***	-0.0204***	-59.1390***
	(-12.18)	(-13.11)	(-11.37)	(-19.87)	(-14.42)
Is female	0.9807***	0.0506***	0.1534***	0.0075***	16.0861***
	(5.21)	(8.38)	(9.33)	(12.30)	(6.63)
Constant	59.2811***	2.5752***	4.5915***	0.2584***	592.2831***

	(30.64)	(41.37)	(26.78)	(40.79)	(23.53)
Observations	177312	177312	177312	177312	177312
Year FE	Yes	Yes	Yes	Yes	Yes

The symbols ***, **, and * denote that p<0.01, p<0.05, and p<0.10, respectively.

B.9 RIF Decomposition Results

Table B.9.1: RIF Decomposition - Quantiles

	Q10	Q25	Q50	Q75	Q90
	(1)	(2)	(3)	(4)	(5)
Overall					
Women	48.8132***	63.4395***	78.0359***	87.7448***	93.9473***
	(382.91)	(628.01)	(1086.01)	(1606.03)	(1832.31)
Counterfactual	50.6065***	65.1678***	79.8064***	89.0903***	94.6564***
	(362.48)	(583.89)	(1017.05)	(1625.62)	(1904.67)
Men	52.6727***	66.8981***	80.7696***	89.5043***	94.8237***
	(455.62)	(601.34)	(1242.80)	(1679.30)	(1988.90)
Total difference	-3.8595***	-3.4586***	-2.7337***	-1.7595***	-0.8764***
	(-22.43)	(-23.02)	(-28.22)	(-23.05)	(-12.52)
Total composition effect	-2.0663***	-1.7302***	-0.9632***	-0.4140***	-0.1673**
	(-11.40)	(-10.98)	(-9.45)	(-5.42)	(-2.43)
Total mental health structure	-1.7933***	-1.7284***	-1.7705***	-1.3455***	-0.7091***
	(-9.49)	(-11.48)	(-16.64)	(-17.39)	(-9.93)
Total composition effect	-2.0663***	-1.7302***	-0.9632***	-0.4140***	-0.1673**
	(-11.40)	(-10.98)	(-9.45)	(-5.42)	(-2.43)

Pure composition effect	-1.5179***	-1.6097***	-0.8142***	-0.4793***	-0.3491***
	(-22.91)	(-24.54)	(-21.82)	(-16.09)	(-13.54)
Specification error	-0.5484***	-0.1205	-0.1490	0.0653	0.1818**
	(-2.95)	(-0.74)	(-1.43)	(0.82)	(2.53)
Pure composition effect					
Marital status (ref: Single)					
Marital Status: Legally married or de facto	-0.2517***	-0.3294***	-0.1904***	-0.1054***	-0.0666***
	(-9.71)	(-12.65)	(-12.51)	(-8.89)	(-6.41)
Marital Status: Divorced/separated/widowed	-0.1567***	-0.1405***	-0.0291	-0.0248	-0.0284
	(-3.46)	(-3.24)	(-1.15)	(-1.19)	(-1.51)
Household size	-0.0535***	-0.0591***	-0.0237***	-0.0209***	-0.0153***
	(-5.85)	(-6.51)	(-4.82)	(-5.11)	(-4.28)
Age	-0.0275	-0.0329	-0.0139	-0.0073	-0.0062
	(-0.41)	(-0.41)	(-0.41)	(-0.41)	(-0.41)
Age square	0.0662	0.0755	0.0364	0.0240	0.0197
	(0.79)	(0.79)	(0.79)	(0.79)	(0.79)
Living outside a major city	-0.0061**	-0.0027	-0.0025**	-0.0026**	-0.0031**
	(-2.15)	(-1.56)	(-1.97)	(-2.11)	(-2.24)
Young family	0.0056**	0.0080**	0.0032**	0.0027**	0.0016*
	(2.11)	(2.50)	(2.13)	(2.14)	(1.69)
Foreign born	0.0014	0.0095***	0.0101***	0.0055***	0.0021*
	(0.50)	(2.88)	(3.89)	(3.20)	(1.68)
Log of after-tax equivalised income	-0.1440***	-0.1933***	-0.1275***	-0.0754***	-0.0452***
	(-10.58)	(-12.28)	(-12.77)	(-11.24)	(-8.95)
Labour force status (ref: Not in the labour force)					

Labour force status: Employed	-0.8584***	-0.7924***	-0.3254***	-0.1638***	-0.1020***
	(-22.75)	(-22.23)	(-17.18)	(-11.28)	(-7.99)
Labour force status: Unemployed	-0.0048	0.0091*	0.0007	0.0002	0.0017
	(-0.91)	(1.76)	(0.25)	(0.10)	(0.79)
Education (ref: Year 12 or below)					
Education: Certificates and diploma	-0.0141	-0.1066***	-0.0668***	-0.0249**	-0.0189*
	(-0.53)	(-4.17)	(-4.45)	(-2.02)	(-1.71)
Education: Undergraduate	0.0659***	0.1262***	0.0179**	-0.0206***	-0.0329***
	(4.31)	(8.16)	(2.11)	(-2.94)	(-5.13)
Education: Postgraduate	-0.0044*	-0.0098**	-0.0028**	0.0008	0.0022**
	(-1.89)	(-2.31)	(-2.00)	(1.03)	(2.01)
Death of spouse or child	-0.0322***	-0.0334***	-0.0184***	-0.0104***	-0.0077***
	(-4.89)	(-5.15)	(-4.96)	(-3.78)	(-3.22)
Death of close relative or family member	-0.0238***	-0.0261***	-0.0095***	-0.0073***	-0.0057**
	(-4.07)	(-4.52)	(-3.04)	(-2.86)	(-2.53)
Serious injury or illness to family member	-0.1268***	-0.1471***	-0.0908***	-0.0676***	-0.0587***
	(-8.66)	(-10.14)	(-10.56)	(-9.75)	(-9.46)
Natural disaster damaged/destroyed home	0.0026	0.0030	0.0013	0.0009	0.0006
	(1.13)	(1.15)	(1.12)	(1.10)	(1.06)
Fired or made redundant	0.0567***	0.0520***	0.0301***	0.0248***	0.0188***
	(5.94)	(5.71)	(5.64)	(5.65)	(4.85)
Victim of physical violence	-0.0443***	-0.0313***	-0.0140***	-0.0072***	-0.0043***
	(-4.17)	(-4.08)	(-3.94)	(-3.52)	(-2.95)
Close family member detained in jail	-0.0180***	-0.0258***	-0.0123***	-0.0056*	-0.0045*
	(-2.75)	(-3.94)	(-3.28)	(-1.90)	(-1.71)

Detained in jail	0.0440***	0.0298***	0.0090***	0.0026	0.0018
	(6.87)	(5.48)	(3.16)	(1.14)	(0.88)
Specification error					
Marital status (ref: Single)					
Marital Status: Legally married or de facto	-0.2171	-0.5411*	0.2012	-0.0533	-0.0536
	(-0.62)	(-1.80)	(1.03)	(-0.36)	(-0.40)
Marital Status: Divorced/separated/widowed	-0.2623**	-0.1683	-0.1462**	-0.0024	-0.0006
	(-1.97)	(-1.45)	(-1.96)	(-0.04)	(-0.01)
Household size	0.2164	0.1028	0.3125	0.0008	-0.0050
	(0.54)	(0.30)	(1.39)	(0.00)	(-0.03)
Age	-19.0154***	-10.8522***	-13.5542***	-2.9808***	-1.6745
	(-7.12)	(-4.71)	(-9.06)	(-2.64)	(-1.63)
Age square	12.1832***	6.6705***	8.0126***	1.9625***	1.1949**
	(8.42)	(5.34)	(9.88)	(3.20)	(2.15)
Living outside a major city	0.0174	-0.0440	-0.0382	-0.0434	-0.0420*
	(0.26)	(-0.77)	(-1.03)	(-1.55)	(-1.66)
Young family	0.0247	-0.0105	-0.0420	-0.0013	-0.0066
	(0.31)	(-0.15)	(-0.95)	(-0.04)	(-0.22)
Foreign born	0.0145	-0.0184	-0.0500	-0.0276	-0.0238
	(0.16)	(-0.24)	(-0.98)	(-0.72)	(-0.68)
Log of after-tax equivalised income	1.4496	-0.6081	7.5337***	0.7608	0.6326
	(0.43)	(-0.21)	(3.97)	(0.53)	(0.49)
Labour force status (ref: Not in the labour force)					
Labour force status: Employed	2.0245***	0.4952*	0.9823***	0.0151	0.0054
	(6.68)	(1.88)	(5.78)	(0.12)	(0.05)

Labour force status: Unemployed	0.1034***	0.0276	-0.0047	0.0018	-0.0017
	(2.90)	(0.90)	(-0.24)	(0.12)	(-0.13)
Education (ref: Year 12 or below)					
Education: Certificates and diploma	-0.0859	0.0936	0.0833	-0.0096	-0.0323
	(-0.70)	(0.89)	(1.22)	(-0.19)	(-0.69)
Education: Undergraduate	-0.0234	0.1830*	0.2413***	0.0788	0.0191
	(-0.20)	(1.81)	(3.69)	(1.59)	(0.42)
Education: Postgraduate	0.0257	0.0549	0.0761***	0.0135	0.0109
	(0.57)	(1.41)	(3.01)	(0.71)	(0.63)
Death of spouse or child	-0.0227	-0.0013	-0.0220**	-0.0033	-0.0037
	(-1.14)	(-0.07)	(-1.97)	(-0.39)	(-0.47)
Death of close relative or family member	-0.1163*	-0.0412	-0.0651	-0.0027	0.0027
	(-1.65)	(-0.67)	(-1.64)	(-0.09)	(0.10)
Serious injury or illness to family member	-0.1648*	-0.0448	-0.0768	-0.0093	-0.0103
	(-1.95)	(-0.61)	(-1.63)	(-0.26)	(-0.32)
Natural disaster damaged/destroyed home	-0.0106	-0.0099	-0.0130	-0.0000	0.0008
	(-0.50)	(-0.54)	(-1.09)	(-0.01)	(0.09)
Fired or made redundant	-0.0049	0.0055	-0.0070	0.0003	0.0004
	(-0.18)	(0.24)	(-0.46)	(0.03)	(0.04)
Victim of physical violence	-0.0443*	-0.0168	-0.0370***	-0.0037	-0.0029
	(-1.89)	(-0.82)	(-2.81)	(-0.38)	(-0.32)
Close family member detained in jail	0.0230	-0.0073	-0.0158	0.0004	-0.0003
	(0.83)	(-0.30)	(-1.02)	(0.04)	(-0.03)
Detained in jail	0.0055	0.0013	-0.0024	-0.0006	-0.0009
	(0.96)	(0.27)	(-0.75)	(-0.25)	(-0.44)

Constant	4.0584	4.8679	-3.3438*	0.3948	0.1005
	(1.15)	(1.59)	(-1.69)	(0.26)	(0.07)
Total mental health structure					
Total	-1.7933***	-1.7284***	-1.7705***	-1.3455***	-0.7091***
	(-9.49)	(-11.48)	(-16.64)	(-17.39)	(-9.93)
Reweighting error	0.0808*	0.0661	0.0625**	0.0626***	0.0485***
	(1.86)	(1.64)	(2.37)	(4.04)	(4.08)
Pure mental health structure	-1.8741***	-1.7944***	-1.8330***	-1.4081***	-0.7576***
	(-10.12)	(-12.26)	(-17.64)	(-18.54)	(-10.74)
Pure mental health structure					
Marital status (ref: Single)					
Marital Status: Legally married or de facto	0.8940**	0.0967	-0.6059***	-0.3108**	-0.6284***
	(2.40)	(0.33)	(-2.89)	(-2.03)	(-4.43)
Marital Status: Divorced/separated/widowed	0.4387***	0.2640***	0.0232	-0.1177**	-0.1411***
	(3.41)	(2.60)	(0.32)	(-2.23)	(-2.88)
Household size	-1.0072**	-1.7623***	-1.0532***	-0.6908***	-0.5892***
	(-2.35)	(-5.21)	(-4.38)	(-3.92)	(-3.60)
Age	32.3323***	33.3024***	25.7933***	14.3318***	11.7610***
	(11.80)	(15.39)	(16.79)	(12.77)	(11.29)
Age square	-16.9748***	-16.8683***	-13.3709***	-7.3806***	-5.7753***
	(-11.41)	(-14.35)	(-16.02)	(-12.11)	(-10.21)
Living outside a major city	-0.0971	0.0042	0.0317	0.0336	0.0427
	(-1.40)	(0.08)	(0.81)	(1.18)	(1.61)
Young family	0.6916***	0.7089***	0.4144***	0.2067***	0.1110***
	(7.64)	(9.90)	(8.16)	(5.58)	(3.23)
Foreign born	0.1522	0.0455	0.0702	0.0118	-0.0196

	(1.57)	(0.59)	(1.29)	(0.30)	(-0.53)
Log of after-tax equivalised income	-6.7149*	-5.9795**	-4.4111**	1.7258	1.2332
	(-1.89)	(-2.13)	(-2.21)	(1.18)	(0.91)
Labour force status (ref: Not in the labour force)					
Labour force status: Employed	-2.5672***	-1.7320***	-1.2698***	-0.3027**	-0.5254***
	(-8.74)	(-7.47)	(-7.71)	(-2.52)	(-4.70)
Labour force status: Unemployed	-0.1753***	-0.0677**	-0.0382*	-0.0231	-0.0052
	(-4.65)	(-2.28)	(-1.81)	(-1.50)	(-0.36)
Education (ref: Year 12 or below)					
Education: Certificates and diploma	0.4132***	-0.0172	0.0195	0.1141**	0.0510
	(3.22)	(-0.17)	(0.27)	(2.17)	(1.04)
Education: Undergraduate	0.3473***	-0.0105	0.1086	0.1140**	0.0194
	(2.94)	(-0.11)	(1.64)	(2.35)	(0.43)
Education: Postgraduate	0.0728	-0.0795**	-0.0754***	-0.0317*	-0.0448**
	(1.56)	(-2.16)	(-2.88)	(-1.66)	(-2.52)
Death of spouse or child	0.0076	-0.0107	0.0143	-0.0065	-0.0053
	(0.41)	(-0.72)	(1.36)	(-0.84)	(-0.74)
Death of close relative or family member	-0.0106	-0.0300	-0.0503	-0.0667**	-0.0544**
	(-0.15)	(-0.53)	(-1.25)	(-2.27)	(-1.99)
Serious injury or illness to family member	-0.0519	0.0074	-0.0084	-0.0667*	-0.0394
	(-0.62)	(0.11)	(-0.18)	(-1.94)	(-1.24)
Natural disaster damaged/destroyed home	0.0027	0.0192	0.0061	-0.0086	-0.0008
	(0.12)	(1.10)	(0.49)	(-0.95)	(-0.10)
Fired or made redundant	-0.1044***	-0.0696***	-0.0572***	-0.0319**	-0.0093
	(-3.45)	(-2.91)	(-3.37)	(-2.57)	(-0.80)

Victim of physical violence	-0.0856***	-0.0549***	-0.0021	-0.0168*	0.0004
	(-3.70)	(-3.02)	(-0.16)	(-1.77)	(0.05)
Close family member detained in jail	-0.0638**	0.0105	0.0014	-0.0096	0.0000
	(-2.46)	(0.52)	(0.10)	(-0.90)	(0.00)
Detained in jail	0.0119*	0.0102*	0.0023	0.0002	0.0022
	(1.70)	(1.83)	(0.59)	(0.05)	(0.84)
Constant	-9.1498**	-8.9848***	-6.4466***	-7.8521***	-5.6075***
	(-2.47)	(-3.07)	(-3.10)	(-5.15)	(-3.96)
Reweighting error					
Marital status (ref: Single)					
Marital Status: Legally married or de facto	0.0645***	0.0761***	0.0605***	0.0282***	0.0172***
	(5.47)	(6.57)	(6.83)	(5.78)	(4.52)
Marital Status: Divorced/separated/widowed	0.0115*	0.0090*	0.0041*	0.0010	0.0011
	(1.82)	(1.82)	(1.71)	(1.11)	(1.25)
Household size	0.0301***	0.0311***	0.0168***	0.0104***	0.0075***
	(4.41)	(4.97)	(4.40)	(4.11)	(3.54)
Age	-0.4676***	-0.4528***	-0.2724***	-0.1051***	-0.0822***
	(-4.45)	(-4.47)	(-4.46)	(-4.31)	(-4.24)
Age square	0.5598***	0.5216***	0.3292***	0.1631***	0.1272***
	(4.40)	(4.41)	(4.40)	(4.36)	(4.33)
Living outside a major city	0.0013	0.0002	0.0003	0.0002	0.0003
	(0.51)	(0.44)	(0.48)	(0.49)	(0.50)
Young family	-0.0136*	-0.0246***	-0.0141***	-0.0079***	-0.0055**
	(-1.93)	(-4.04)	(-3.41)	(-2.78)	(-2.15)
Foreign born	-0.0003	-0.0049**	-0.0059**	-0.0032**	-0.0015*
	(-0.20)	(-2.05)	(-2.27)	(-2.18)	(-1.78)

Log of after-tax equivalised income	-0.0125	-0.0159	-0.0134	-0.0066	-0.0040
	(-1.30)	(-1.31)	(-1.31)	(-1.31)	(-1.30)
Labour force status (ref: Not in the labour force)					
Labour force status: Employed	-0.0998***	-0.0747***	-0.0405***	-0.0143***	-0.0089***
	(-3.23)	(-3.23)	(-3.22)	(-3.14)	(-3.04)
Labour force status: Unemployed	-0.0011	0.0001	0.0001	-0.0000	0.0001
	(-0.37)	(0.29)	(0.30)	(-0.07)	(0.34)
Education (ref: Year 12 or below)					
Education: Certificates and diploma	0.0004	-0.0037	-0.0025	-0.0006	-0.0002
	(0.44)	(-1.16)	(-1.16)	(-0.96)	(-0.56)
Education: Undergraduate	0.0042	0.0109	0.0043	-0.0004	-0.0020
	(1.36)	(1.45)	(1.43)	(-0.72)	(-1.40)
Education: Postgraduate	-0.0022	-0.0049	-0.0026	0.0000	0.0006
	(-0.96)	(-1.00)	(-0.99)	(0.11)	(0.93)
Death of spouse or child	-0.0005	-0.0004	-0.0003	-0.0001	-0.0001
	(-0.10)	(-0.10)	(-0.10)	(-0.10)	(-0.10)
Death of close relative or family member	0.0026	0.0022	0.0012	0.0005	0.0004
	(0.65)	(0.65)	(0.64)	(0.64)	(0.63)
Serious injury or illness to family member	0.0018	0.0017	0.0012	0.0007	0.0006
	(0.23)	(0.23)	(0.23)	(0.23)	(0.23)
Natural disaster damaged/destroyed home	0.0008	0.0009	0.0005	0.0002	0.0002
	(0.32)	(0.32)	(0.32)	(0.32)	(0.32)
Fired or made redundant	0.0024	0.0020	0.0014	0.0010	0.0008
	(0.81)	(0.81)	(0.81)	(0.81)	(0.81)
Victim of physical violence	0.0110	0.0072	0.0043	0.0017	0.0010

	(0.91)	(0.91)	(0.91)	(0.90)	(0.90)
Close family member detained in jail	0.0008	0.0022	0.0013	0.0004	0.0004
	(0.68)	(0.74)	(0.74)	(0.71)	(0.70)
Detained in jail	-0.0001	-0.0001	-0.0000	-0.0000	-0.0000
	(-0.05)	(-0.05)	(-0.05)	(-0.05)	(-0.05)
Year FE	Yes	Yes	Yes	Yes	Yes

Notes: Q10, Q25, Q50, Q75, and Q90 are the 10th, 25th, 50th, 75th and 90th quantiles. The symbols ***, **, and * denote that p<0.01, p<0.05, and p<0.10, respectively.

Table B.9.2 shows that overall, all models suggest that mental health inequality is higher among women compared with men. Inequality among women is between 7% to 11.7% higher than for men. The smallest gaps are measured for the interquartile ratio, while the largest is observed for the variance of mental health. Most of the negative life events such as the death of a spouse or child, death of a close relative or family member, serious injury or illness to a family member, a victim of physical violence and close family member detained in jail increase the mental health inequality. Meanwhile, experiencing a natural disaster, being fired or made redundant, or being detained in jail themselves appears to reduce the mental health inequality.

Table B.9.2: RIF Decomposition - Other Inequality Measures

	IQR(90 10)	IQRATIO(90 10)	IQSR(10 90)	GINI x 100	Variance
	(1)	(2)	(3)	(4)	(5)
Overall					
group_1	45.1341***	1.9219***	2.7522***	0.1367***	324.0602***
	(344.55)	(424.53)	(202.75)	(313.02)	(189.71)
group_2	42.1317***	1.7948***	2.4687***	0.1239***	290.2070***

	(358.19)	(528.88)	(213.62)	(286.65)	(168.22)
difference	3.0023***	0.1271***	0.2836***	0.0128***	33.8532***
	(17.05)	(22.47)	(15.91)	(20.84)	(13.94)
explained	1.1044***	0.0368***	0.1357***	0.0062***	20.2308***
	(17.04)	(19.29)	(20.52)	(23.86)	(20.38)
unexplained	1.8979***	0.0903***	0.1478***	0.0066***	13.6224***
	(10.36)	(15.54)	(8.04)	(10.42)	(5.40)
explained					
Marital status (ref: Single)					
Marital Status: Legally married or de facto	0.1367***	0.0047***	0.0200***	0.0010***	3.0892***
	(7.18)	(8.39)	(10.08)	(12.33)	(10.34)
Marital Status: Divorced/separated/widowed	0.0985**	0.0033**	0.0003	0.0002	0.0950
	(2.21)	(2.55)	(0.06)	(1.36)	(0.15)
Household size	0.0170***	0.0006***	0.0023***	0.0001***	0.3264***
	(3.46)	(4.00)	(4.21)	(4.63)	(4.14)
Age	0.2411***	0.0081***	0.0356***	0.0016***	5.5145***
	(4.67)	(4.74)	(4.83)	(4.86)	(4.83)
Age square	-0.2895***	-0.0103***	-0.0451***	-0.0020***	-6.6232***
	(-4.98)	(-5.07)	(-5.13)	(-5.16)	(-5.13)
Living outside a major city	0.0040*	0.0002*	0.0006*	0.0000*	0.0718*
	(1.76)	(1.86)	(1.90)	(1.90)	(1.84)
Young family	0.0072*	0.0003**	0.0006*	0.0000***	0.1070*
	(1.81)	(2.17)	(1.67)	(2.72)	(1.85)
Foreign born	0.0003	-0.0000	-0.0000	-0.0000*	-0.0120
	(0.20)	(-0.14)	(-0.19)	(-1.86)	(-0.54)
Log of after-tax equivalised income	0.0852***	0.0032***	0.0117***	0.0007***	1.7201***
	(6.96)	(8.59)	(9.14)	(11.87)	(9.02)
Labour force status (ref: Not in the labour force)					
Labour force status: Employed	0.8043***	0.0254***	0.1033***	0.0042***	15.5889***

	(20.86)	(22.34)	(25.27)	(26.74)	(25.46)
Labour force status: Unemployed	0.0074	0.0002	0.0027***	0.0001***	0.3868***
	(1.31)	(1.22)	(4.42)	(2.63)	(4.24)
Education (ref: Year 12 or below)					
Education: Certificates and diploma	-0.0193	-0.0003	-0.0041	0.0001	-0.4266
	(-0.69)	(-0.42)	(-1.51)	(0.69)	(-1.05)
Education: Undergraduate	-0.0821***	-0.0020***	-0.0028*	-0.0003***	-1.0179***
	(-4.84)	(-4.05)	(-1.72)	(-5.05)	(-4.14)
Education: Postgraduate	0.0066**	0.0002**	0.0001	0.0000**	0.0741**
	(2.34)	(2.09)	(0.58)	(2.57)	(1.99)
Death of spouse or child	0.0242***	0.0008***	0.0025***	0.0001***	0.3665***
	(3.92)	(4.47)	(4.12)	(5.34)	(4.05)
Death of close relative or family member	0.0161***	0.0006***	0.0022***	0.0001***	0.3166***
	(3.04)	(3.52)	(4.02)	(4.48)	(3.91)
Serious injury or illness to family member	0.0673***	0.0028***	0.0119***	0.0005***	1.4755***
	(4.80)	(6.75)	(8.27)	(9.55)	(7.04)
Natural disaster damaged/destroyed home	-0.0025	-0.0001	-0.0003	-0.0000	-0.0529
	(-1.37)	(-1.41)	(-1.44)	(-1.48)	(-1.45)
Fired or made redundant	-0.0391***	-0.0014***	-0.0069***	-0.0003***	-0.9528***
	(-3.99)	(-4.99)	(-6.84)	(-7.26)	(-6.42)
Victim of physical violence	0.0313***	0.0010***	0.0039***	0.0001***	0.5724***
	(3.44)	(3.45)	(3.47)	(3.47)	(3.47)
Close family member detained in jail	0.0111*	0.0004**	0.0003	0.0001**	0.0777
	(1.83)	(2.20)	(0.44)	(2.27)	(0.89)
Detained in jail	-0.0412***	-0.0012***	-0.0047***	-0.0002***	-0.7610***
	(-6.55)	(-6.76)	(-7.18)	(-7.26)	(-7.49)
Residential area index of socioeconomic advantage (ref: 1st quintile)					
2nd quintile	-0.0025	-0.0001	-0.0002	-0.0000	-0.0365
	(-1.15)	(-1.19)	(-1.09)	(-1.23)	(-1.15)

3rd quintile	-0.0100**	-0.0003**	-0.0008**	-0.0001**	-0.1620**
	(-2.31)	(-2.34)	(-2.22)	(-2.45)	(-2.35)
4th quintile	0.0031	0.0001	0.0003	0.0000	0.0508
	(0.69)	(0.69)	(0.69)	(0.69)	(0.69)
5th quintile	0.0248***	0.0008***	0.0020***	0.0001***	0.3801***
	(3.52)	(3.55)	(3.38)	(3.70)	(3.55)
unexplained					
Marital status (ref: Single)					
Marital Status: Legally married or de facto	-1.3740***	-0.0503***	-0.0899***	-0.0012	-7.9694*
	(-4.05)	(-4.63)	(-2.63)	(-1.05)	(-1.71)
Marital Status: Divorced/separated/widowed	-0.3286**	-0.0078*	-0.0004	-0.0005	-1.4666
	(-2.57)	(-1.93)	(-0.03)	(-1.07)	(-0.83)
Household size	0.1906	0.0089	0.0558	0.0040***	9.8998*
	(0.47)	(0.69)	(1.38)	(2.90)	(1.80)
Age	-2.6536	-0.0945	-1.1770***	-0.0683***	-190.3757***
	(-1.04)	(-1.16)	(-4.59)	(-7.79)	(-5.43)
Age square	-0.1390	-0.0457	0.4127***	0.0302***	89.5276***
	(-0.10)	(-1.03)	(2.95)	(6.31)	(4.68)
Living outside a major city	0.0722	0.0011	0.0054	0.0003	1.2655
	(1.10)	(0.53)	(0.81)	(1.29)	(1.41)
Young family	-0.6190***	-0.0231***	-0.0706***	-0.0030***	-9.5498***
	(-7.37)	(-8.61)	(-8.36)	(-10.34)	(-8.29)
Foreign born	-0.2118**	-0.0067**	-0.0227**	-0.0006*	-2.7350**
	(-2.34)	(-2.30)	(-2.50)	(-1.91)	(-2.21)
Log of after-tax equivalised income	7.9738**	0.1088	0.5205	0.0281**	133.2373***
	(2.28)	(0.97)	(1.48)	(2.35)	(2.79)
Labour force status (ref: Not in the labour force)					
Labour force status: Employed	0.0204	-0.0230**	0.0569**	0.0046***	19.1516***
	(0.07)	(-2.52)	(1.97)	(4.65)	(4.81)

Labour force status: Unemployed	0.0672*	0.0024**	0.0117***	0.0004***	1.7270***
	(1.94)	(2.14)	(3.36)	(3.40)	(3.65)
Education (ref: Year 12 or below)					
Education: Certificates and diploma	-0.3299***	-0.0121***	-0.0474***	-0.0013***	-5.4749***
	(-2.79)	(-3.19)	(-3.99)	(-3.27)	(-3.40)
Education: Undergraduate	-0.2983**	-0.0128***	-0.0452***	-0.0017***	-5.3249***
	(-2.56)	(-3.46)	(-3.87)	(-4.17)	(-3.32)
Education: Postgraduate	-0.1356***	-0.0045***	-0.0117***	-0.0003*	-1.5955***
	(-3.09)	(-3.20)	(-2.66)	(-1.96)	(-2.66)
Death of spouse or child	0.0067	0.0009	0.0053***	0.0001	0.4966*
	(0.34)	(1.44)	(2.69)	(1.30)	(1.82)
Death of close relative or family member	0.0681	0.0045**	0.0081	0.0003	0.3928
	(0.99)	(2.05)	(1.17)	(1.31)	(0.42)
Serious injury or illness to family member	0.1656**	0.0097***	0.0152*	0.0004	0.5264
	(2.01)	(3.71)	(1.84)	(1.40)	(0.46)
Natural disaster damaged/destroyed home	0.0069	0.0006	0.0006	-0.0000	-0.1294
	(0.34)	(0.84)	(0.27)	(-0.33)	(-0.46)
Fired or made redundant	0.1016***	0.0043***	0.0081***	0.0003***	0.6493*
	(3.88)	(5.04)	(3.05)	(3.22)	(1.84)
Victim of physical violence	0.1241***	0.0060***	0.0182***	0.0004***	1.5432***
	(5.44)	(8.19)	(7.87)	(5.61)	(4.94)
Close family member detained in jail	0.0396	0.0017**	0.0066**	0.0002*	0.7127*
	(1.48)	(2.00)	(2.47)	(1.72)	(1.92)
Detained in jail	-0.0160***	-0.0005**	-0.0019***	-0.0001***	-0.3249***
	(-2.81)	(-2.46)	(-3.24)	(-3.63)	(-4.12)
Residential area index of socioeconomic advantage (ref: 1st quintile)					
2nd quintile	-0.1357	-0.0066*	-0.0278**	-0.0009**	-3.0608**
	(-1.22)	(-1.86)	(-2.49)	(-2.42)	(-2.01)
3rd quintile	-0.1015	-0.0073**	-0.0191*	-0.0006	-1.2463

	(-0.90)	(-2.01)	(-1.68)	(-1.44)	(-0.80)
4th quintile	-0.0801	-0.0054	-0.0170	-0.0002	-0.9047
	(-0.68)	(-1.42)	(-1.43)	(-0.58)	(-0.56)
5th quintile	-0.1714	-0.0122***	-0.0374***	-0.0010**	-2.8182*
	(-1.41)	(-3.14)	(-3.06)	(-2.30)	(-1.69)
Constant	-0.8310	0.2173*	0.5606	0.0157	-7.0951
	(-0.23)	(1.89)	(1.56)	(1.28)	(-0.14)
Year FE	Yes	Yes	Yes	Yes	Yes

Appendix C

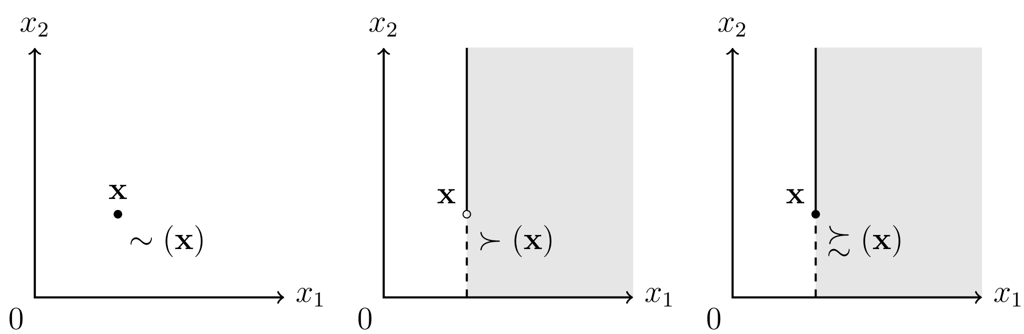
Appendix for Chapter 4

C.1 Lexicographic Preferences

Lexicographic preferences describe comparative preferences where an individual prefers any amount of one good to any amount of another good. For example, take an arbitrary point or bundle $\mathbf{x} \in \mathbb{R}^2_+$. A consumer has **lexicographic** preferences over $\mathbb{R}^2_+$ if the relation $\succsim$ satisfies $\mathbf{y} \succsim \mathbf{x}$ whenever $y_1 > x_1$, or $y_1 = x_1$ and $y_2 \geq x_2$. The indifference set, strict upper- and weak upper-contour sets of a typical bundle in $\mathbb{R}^2_+$ are illustrated in Figure C.1.1.

Lexicographic preference is a total order (complete, transitive and antisymmetric) but it is not continuous. Lexicographic preference is a classic example in microeconomics to show that not all preference orderings can be represented by utility functions because it is not continuous (the strict upper contour set is not open or the weak upper contour set is not closed).

Figure C.1.1: The "indifference", "preferred to", and "at least as good as" sets (from left to right)



C.2 Stochastic Dominance

The stochastic dominance concepts allow one to determine the preference of an expected utility maximizer when he faces some lotteries, with minimal knowledge of his utility function. The *first-order stochastic dominance* says that a decision maker prefers a lottery F which first-order stochastically dominates G whenever he has a weakly increasing utility function u. The *second-order stochastic dominance* says that a decision maker prefers a lottery F which second-order stochastically dominates G as long as he is risk averse and u is weakly increasing.

C.2.1 First Order Stochastic Dominance

Theorem 1 (Equivalance of first-order stochastic dominance definitions)**.** *The following are equivalent.*

1. *For every weakly increasing utility function u, i.e., $\int u(x)\,dF \geq \int u(x)\,dG$.*
2. *$F(x) \leq G(x)$ for all x.*

C.2.2 Second Order Stochastic Dominance

Theorem 2 (Equivalance of second-order stochastic dominance definitions)**.** *Assume that $\int xdF = \int xdG$. The following are equivalent.*

1. *$\int u(x)\,dF(x) \geq \int u(x)\,dG(x)$ for every weakly increasing concave utility function u.*
2. *G is a mean-preserving spread of F i.e $y = x + \epsilon$ for some $x \sim F, y \sim G$ and ϵ such that $E(\epsilon|x) = 0$ for all x.*
3. *For every $t \geq 0, \int_a^t G(x)\,dx \geq \int_a^t F(x)\,dx$.*

C.3 Comparing the Lexicographic Serial Rule and the Hylland-Zeckhauser Procedure

In the model of Hylland and Zeckhauser (1979) (henceforth HZ), the agents have vNM utilities over the objects, and they trade probabilities of receiving the objects on a market. This is not an ordinal concept, so it is in one sense not comparable to random priority, the serial rule, and the lex rule, which depend only on ordinal preferences. However, we have motivated the lex rule in terms of particular cardinal preferences, so it makes sense to compare the lex rule with the equilibria of the HZ model for these utilities.

Although the HZ model can be defined somewhat more generally, we restrict attention to the simplest case, which is that $|I| = |A|$, so that each agent receives exactly one object and no one is unassigned. Let

$$\Delta = \{\, p \in \Re^A_+ : \sum_a p_a = 1 \,\}$$

be the set of probability measures on A, and let $\omega = \sum_a \frac{1}{|A|}\delta_a$ be its barycenter.[1] Let $\nabla = (1, \dots, 1) \in \Re^A$. An *allocation* is a matrix $P = (p_{ia})_{i \in I, a \in A}$ whose rows p_i are elements of Δ such that $\sum_i p_i = \nabla$. A *price vector* can be any $z \in \Re^A$. The (common) *budget set* for such a z is

$$B(z) = \{\, p \in \Delta : \langle p, z \rangle \leq \langle \omega, z \rangle \,\},$$

so each agent's initial endowment is the probability measure ω that assigns equal probability to all objects. A *utility* is an element of $\Re^A$, and *utility profile* is a n-tuple $u = (u_i)_{i \in I}$ of utilities. For $i \in I$ let agent i's demand correspondence be given by

$$D_i(z) = \mathrm{argmax}_{p_i \in B(z)} \langle p_i, u_i \rangle.$$

An *HZ equilibrium* for u is a price-allocation pair (z, P) such that $p_i \in D_i(p)$ for all i.

[1]Here $\delta_a \in \Delta$ is the Dirac measure that assigns all probability to a.