Geography, reference groups, and the determinants of life satisfaction

by

Christopher Paul Barrington-Leigh

S.B. Physics, Massachusetts Institute of Technology, 1995
Ph.D. Applied Physics, Stanford University, 2001

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

The Faculty of Graduate Studies

(Economics)

The University of British Columbia

(Vancouver)

January 2009

Abstract

This dissertation combines three contributions to the literature on the determinants of well-being and the social nature of preferences. Departures from self-centred, consumption-oriented decision making are increasingly common in economic theory and are empirically well motivated by a wide range of behavioural data from experiments, surveys, and econometric inference. The first two contributions are focused on the idea that reference levels set by others' consumption may figure prominently in both experienced well-being and in decision making. In the first paper, the well-being question is addressed empirically through the use of self-reported life satisfaction and high-resolution census and survey data in Canada. Strong income externalities are found at multiple spatial scales after controlling for various confounding factors. The second paper explores the general equilibrium consequences of a utility function having an explicit comparison with neighbours' consumption. The question is investigated in a model in which decision makers knowingly choose their neighbours — and hence their consumption reference level — as well as their own consumption expenditure, thereby helping to set the reference level for nearby others. For both discrete and continuous distributions of types in an economy with a heterogeneous population undergoing such endogenous formation of consumption reference groups, there exist general equilibria in which differentiation of neighbourhoods occurs endogenously. The novel welfare implications of growth in such economies are described. The final paper addresses econometric reservations about the use of subjective reports as dependent variables. The date and location of survey interviews are combined with weather and climate records to construct the random component of weather conditions experienced by respondents on the day of their interview. Standard inferences about the determinants of life satisfaction remain robust after taking into account this significant source of affective bias.

Contents

Abstract . . . ii

Contents . . . iii

List of tables . . . vii

List of figures . . . ix

Epistemological preface . . . xi

Acknowledgements . . . xiv

Dedication . . . xvi

Co-authorship statement . . . xvii

1 Introduction . . . 1
- 1.1 Happiness in economics . . . 1
- 1.2 Veblen preferences . . . 2
- 1.3 Contributions . . . 3
- Bibliography for Chapter 1 . . . 5

2 The geographic scale of urban Veblen effects . . . 1
- 2.1 Introduction . . . 1
- 2.2 Data and method . . . 5
- 2.3 Results and interpretation . . . 9
  - 2.3.1 Classical regression . . . 9
  - 2.3.2 Veblen effects . . . 12
  - 2.3.3 Exposure response . . . 12
  - 2.3.4 Price levels . . . 15
  - 2.3.5 Wealth and income . . . 17
  - 2.3.6 Life in the big city . . . 17
  - 2.3.7 Status and signalling . . . 20

2.3.8 Symmetry of income effects 23
2.3.9 Geo-demographic reference groups 26
2.3.10 Further robustness checks 28
2.3.11 Absolute and relative benefits of health 28
2.4 Discussion 28
2.5 Conclusion 33
Bibliography for Chapter 2 35

3 A model of neighbourhoods and endogenous reference group choice 39
3.1 Introduction 39
3.2 Discrete types and unpriced land 41
3.2.1 Summary 45
3.3 A Continuum of types and a market for land 49
3.3.1 Agents' problem 50
3.3.2 Firms' problem 51
3.3.3 Definition of equilibrium 52
3.3.4 Land markets are required for separating equilibria 53
3.3.5 Some general properties of equilibrium with a land market 54
3.3.6 Log-exp-log utility with equitable ownership 54
3.3.7 General equilibrium averages 57
3.3.8 Concavity 59
3.3.9 Existence 60
3.3.10 Welfare analysis of interior equilibria 60
3.3.11 Empirical interpretation 64
3.3.12 Log-exp-log utility with absentee landlords 64
3.3.13 Pooling equilibria 64
3.3.14 Planner's problem 65
3.4 Numerical analysis 66
3.5 Conclusion 69
Bibliography for Chapter 3 71

4 Weather and life satisfaction 72
4.1 Introduction 72
4.1.1 Reliability: does SWL vary too much? 73
4.1.2 Meaningfulness: does SWL not vary enough? 74
4.1.3 Stock markets and behaviour 75
4.1.4 Sunlight and depression 76
4.1.5 Climate, geography, and well-being 76
4.2 Data and Method 77
4.2.1 Assignment of weather stations 78

4.3 Evidence and discussion 80
4.3.1 Weather and well-being 80
4.3.2 Weather and other determinants of well-being 86
4.3.3 Climate and well-being 90
4.3.4 Cyclic temporal effects 92
4.4 Conclusions 94
Bibliography for Chapter 4 98

5 Conclusions and further work 101
Bibliography for Chapter 5 105

A Appendix to Chapter 2 106
A.1 Supplementary tables for urban geography of life satisfaction 106
A.2 Survey descriptions, consistency, and summary statistics 123
A.2.1 Survey descriptions 123
A.2.2 Consistency of place-based characteristics 124
A.2.2.1 Trust in neighbours 124
A.2.2.2 Life satisfaction 124
A.2.2.3 Other variables 125
A.2.3 Variation across geographical regions 125
A.2.4 Life satisfaction rankings based on ESC2 alone 126
Bibliography for Appendix to Chapter 2 150

B Appendix to Chapter 3 151
B.1 Endogenous reference groups are not club goods 151
B.2 Neighbourhood segregation 152
B.2.1 Exogenous segregation and Veblen consumption 152
B.2.2 Endogenous segregation without neighbourhood benefits 153
B.2.3 Neighbourhood benefits 154
B.3 Functional forms for Veblen preferences 155
B.4 Nonexistence of separating equilibrium for discrete types model 156
B.4.1 Direct neighbourhood benefits 156
B.4.2 “Log-log-log” preferences with two types 158
B.4.3 Mixed strategies 160
B.4.4 Neighbourhood benefits compared with other neighbourhoods 160
B.4.5 “Log-log-exp” preferences with two types 162
B.5 Proofs 162
B.6 Construction of equilibrium 169
Bibliography for Appendix to Chapter 3 171

C Appendix to Chapter 4 172
C.1 Detailed Tables 172

List of tables

2.1 A classical regression . . . 10
2.2 Baseline estimates of relative income effects . . . 13
2.3 Summary of Veblen coefficients for various subgroups . . . 14
2.4 Effect of CMA price correction . . . 16
2.5 Alternate measures of wealth and income . . . 18
2.6 Own and neighbours' dwelling sizes . . . 19
2.7 Income effects and age . . . 21
2.8 Summary coefficients with urban life controls . . . 22
2.9 Income effects, sex, and marriage . . . 24
2.10 Symmetry in comparison effect . . . 25
2.11 Demographic / geographic subpopulations as reference groups . . . 29
2.12 Robustness checks . . . 30
2.13 Spillover effects of others' health . . . 31

4.1 Weather and SWL, without geographic controls . . . 81
4.2 Weather and SWL, allowing for monthly fixed effects . . . 83
4.3 Weather and SWL, allowing for local fixed effects . . . 84
4.4 Weather and SWL, controlling for local climate . . . 85
4.5 Weather and other covariates of SWL . . . 87
4.6 Weather and a compressed measure of SWL . . . 89
4.7 Comparison with naïve models . . . 91
4.8 Climate and satisfaction with life . . . 93
4.9 Days of the week and SWL . . . 95
4.10 Calendar months and SWL . . . 96

A.1 Detailed regressions for main estimates and subpopulations . . . 107
A.2 Detailed regressions with alternate wealth measures . . . 112
A.3 Detailed regressions with measures of dwelling size . . . 113
A.4 Detailed regressions with demographic comparison groups . . . 114
A.5 Detailed regressions for income effects, sex, and marriage . . . 117
A.6 Detailed regressions for spillover effects of others' health . . . 122
A.7 Summary of survey variables. . . . 127

A.7 Summary of survey variables. 128
A.7 Summary of survey variables. 129
A.7 Summary of survey variables. 130

C.1 Weather effects on well-being, trust, and income: details 173
C.2 Climate and satisfaction with life 177

List of figures

2.1 Life satisfaction and mean income (k$/yr) averaged by CMA 6
2.2 Life satisfaction and self-reported trust in neighbours, averaged by CMA 6
2.3 Veblen coefficients as a function of income 27

3.1 Contingent existence of separating equilibrium 44
3.2 Additional cases of equilibrium under "log-exp-log" preferences 46
3.3 Further cases of equilibrium under "log-exp-log" preferences 47
3.4 Separating equilibrium parameter relationships 48
3.5 An equilibrium with monotonically increasing price amongst occupied neighbourhoods 66
3.6 An equilibrium for which $r > \bar{\bar{h}}$ 67
3.7 An equilibrium with non-monotonic price 68
3.8 An equilibrium in which all households prefer the separating equilibrium to the pooling one 68
3.9 Total marginal change to welfare in an economy subject to uniform growth . . . 69

4.1 Comparison of the "nearest" and "clustered" algorithms for assigning weather stations to respondents 79

A.1 Histograms of reported *life satisfaction* in several Canada-wide surveys 131
A.2 Correlation between surveys of mean trust in neighbours by province 132
A.3 Correlation between surveys of mean trust in neighbours by CMA 133
A.4 Correlation between surveys of mean trust in neighbours by CSD 133
A.5 Comparison of provincial mean life satisfaction from different surveys 134
A.6 Comparison of CMA mean life satisfaction from different surveys 134
A.7 Comparison of CSD mean life satisfaction from different surveys 135
A.8 Comparison of provincial mean *importance of religion* from different surveys . 135
A.9 Comparison of CMA mean *importance of religion* from different surveys . . . 136
A.10 Comparison of CSD mean *importance of religion* from different surveys 136
A.11 Comparison of provincial mean trust in colleagues from different surveys . . . 137
A.12 Comparison of CMA mean trust in colleagues from different surveys 137
A.13 Comparison of CSD mean trust in colleagues from different surveys 138

A.14 Comparison of provincial mean trust in family from different surveys 138
A.15 Comparison of CMA mean trust in family from different surveys 139
A.16 Comparison of CSD mean trust in family from different surveys 140
A.17 Comparison of provincial mean subjective health from different surveys 140
A.18 Comparison of CMA mean subjective health from different surveys 141
A.19 Comparison of CSD mean subjective health from different surveys 141
A.20 Life satisfaction means by province 142
A.21 Life satisfaction means by province, corrected for survey averages 143
A.22 Life satisfaction means by CMA 144
A.23 Life satisfaction means by CMA, corrected for survey averages 145
A.24 Trust in neighbours by province 146
A.25 Trust in neighbours by CMA 147
A.26 Life satisfaction from ESC2 by province 148
A.27 Life satisfaction from ESC2 by CMA 149

B.1 Non-existence of separating equilibrium 161
B.2 Non-existence of separating equilibrium 163

Epistemological preface

The two most fundamental ingredients in the study of economics must be the answers to the questions: "How do people make decisions?" and "What is good for people?" In recent decades, the majority of economists have generally not had their ears and minds adequately open to the best available answers coming from other disciplines.

When I came to economics, it seemed clear to me that while utility theory is a useful model at the margin, to think of individual self-centred (consumption) utility functions as more stable (exogenous) than a myriad of other varying conditions is absurd, particularly in the context of market preferences. In the sciences we always talk about equilibrium in the context of a particular time scale, but what are the appropriate temporal scales in which consumption preferences are static when we are subjected to changing influences of information and norms?

This fundamental difference between my intuition and that of my peers' and mentors' contributed to a frustrating time in my early studies. While economics championed its success at fine-tuning, it seemingly had failed to describe or value the large directions in which my society had evolved.

Now, the fields of behavioural economics and subjective well-being have opened up some of the economist's methods to the ugly complexities of the real world. While I came with lofty ambitions to address particular grand theoretical, or rhetorical, questions about macroeconomies, I have not been able to pass quickly over my two fundamental questions with both an intact conscience and a mature theoretical toolbox in hand with which to approach the bigger issues. I thus have come to focus on these questions directly in my study of economics.

I start from the point of view that understanding how we work is an empirical question. Evolutionary principles can tell us little, *a priori*, about the details of human choice behaviour, and observed behaviour (revealed preference) in a given society can tell us only so much about what does actually result in a positive life experience.

I believe it is of critical importance to expose our assumptions about social aims and to assess them scientifically wherever possible. As I complete this dissertation, the rise of concern about global climate change has finally come to dominate many domestic and international policy debates. Although only one piece of a large picture of ecological externalities gone awry, it has highlighted the deficiency of any framework which treats individualised incentives as a good approximation to the bigger picture of benefits and costs. It is possible that another form of externality, in addition to the evident ecological ones, is equally powerful in shaping our fortunes and equally threatening to market-first thinking. These externalities are the social ones, such as the effect of one's own behaviour on others' preferences for consumption, and such as the positive benefits which occur when individuals reach out to each other and thereby build community.

If we get through this climate crisis having only changed our greenhouse gas emissions, without more explicitly realising and rationalising our social objectives and without correcting other similarly overwhelming social and ecological externalities, we will have missed the whole point.

Consider the following. We humans are social beings. We come into the world as the result of others' actions. We survive here in dependence on others. Whether we like it or not, there is hardly a moment of our lives when we do not benefit from others' activities. For this reason it is hardly surprising that most of our happiness arises in the context of our relationships with others.

— HH Dalai Lama

Our desires and pleasures spring from society; we measure them, therefore, by society and by the objects which serve for their satisfaction. Because they are of a social nature, they are of a relative nature. ... A house may be large or small; as long as the surrounding houses are equally small it satisfies all social demands for a dwelling. But let a palace arise beside the little house, and it shrinks from a little house to a hut ... the occupant of the relatively small house will feel more and more uncomfortable, dissatisfied and cramped within its four walls.

— Karl Marx[1]

...men do not desire to be rich, but richer than other men.

— John Stuart Mill[2]

Unless you've measured it, you don't know what you're talking about.

— Lord Kelvin

[1] *Marx and Engels* [1848, p. 163], cited in *Kingdon and Knight* [2007]
[2] Cited by *Graham and Pettinato* [2002]

Acknowledgements

I have received help and support from so many.

I still owe thanks to my earlier academic mentors for believing in me (see my acknowledgements from last time) and to Umran Inan in particular for helping to set the example that grad school can be a fun and rewarding time. Thanks also to Umran, Steve Cummer, David Smith, and others who respected, supported, and tried to understand my transition and best of all made it smooth through continued collaborations and conversations in physics.

For insightful teaching during my generally conventional economics coursework in the first two years of the PhD program, I would like to thank Thomas Lemieux and expecially Francisco Gonzalez. In a course on Conflict, Francisco inspired me to think that some mathematical details and rhetorical applications of theoretical stories were still interesting and also inspired and encouraged me to write more openly about my beliefs and interpretations of the impact of faulty ideas in economics.

“Wait for me! I want to come too!” Ben, Chris, and (not visible) Ken and Kelly, pulling the laggard anchorman over the wall. (2006)

I am especially and deeply grateful to fellow students Ken Jackson and Chris Bidner, with whom I met, nearly weekly at times, to discuss our work and provide peer supervision as our ideas progressed from first year explorations to completed job talks. Others in my class were also especially supportive of me in my first year of trying to become an economist.

For useful encouragement, criticism, and other support, I am grateful to many other faculty in the department, especially Patrick Francois, Thomas Lemieux, Nicole Fortin, Mukesh Eswaran, and David Green. Patrick kindly supervised my work in year three, constructively offering both encouragement and firm advice, and he stood on my exam committee for my department and final defences. Thomas fearlessly became my committee chair and both he and Nicole Fortin gave valuable feedback on the empirical papers. The department’s faculty are characterised by open minds and open doors and I received from the department compassion and support during the crisis surrounding the loss of my brother.

In 2006 I finally met John Helliwell, the unrelenting skeptic and inspiring optimist who,

though vastly experienced as a prominent economist, seemed to support nearly every part of my unconventional views and interests. He has exemplified the interdisciplinary breadth of perspective and the civic engagement that makes a renaissance person out of an academic economist, and I have benefited tremendously from his support and nurturing in many ways. John, thank you for transforming my relationship with economics.

I have been fortunate to have Haifang Huang and Anthony Harris as colleagues in the Helliwell group. Lee Grenon and the staff of Statistics Canada's Research Data Centre at UBC have my gratitude for years of support, efficiency, and flexibility which made possible much of the work in this dissertation. I am also especially grateful to Shelley Phipps of Dalhousie University who served as my external examiner and provided many useful comments on the work.

Hadi Dowlatabadi has been both an inspiration as an academic with reach outside the ivory tower and has also been a friend and supporter. Hadi also helped to make my graduate student experience much more comfortable by providing me desk space amongst friends in the Aquatic Ecosystems Research Lab as well as by sharing discussions, advice, and always a smile and interesting story or idea (or chocolate) as he passed by.

To the founders and residents of Green College — thank you. My time there was professionally, academically, and socially enriching beyond expectation, and I maintain that the College is an academic gem of global significance.

To Louise, for so much clear thought, wise advice, and loving encouragement during the last year and more: thank you for this and everything else.

To my dear sister Rosalind, who has been my closest friend during these years and at times a crucial support: you have my admiration always.

My brother Robert inspired me with his intellect but more so with his centredness and exquisite gentleness and humour. Robert took his life in unexplained circumstances during my work on this degree. I will continue to hold my memory of him as an aspiration for how to treat my neighbours. Robert, though 12 years my junior, helped me with my PhD coursework at UBC and would have been a great support to me in this work.

To my parents Iris and John, who have been unwavering in their support for my path in life, despite its twists and turns, and who endowed me with a passion to live ethically and independently, I am endlessly grateful. And to the rest of my nearby family — Stephen, Clara, Thomas, and John: I am lucky to have you.

CHRISTOPHER P. BARRINGTON-LEIGH
Vancouver, B.C.
13 January 2008

This work was generously supported by external funding from the Social Sciences and Humanities Research Council of Canada, by Statistics Canada, and by the Canadian Institute for Advanced Research.

For Iris: this one — and I promise it's the last — is for you,
who set me on this hidden path in 2000,

And in memory of Robert, who had so much left to say.

Co-authorship statement

Chapter 2 is a manuscript co-authored with John F. Helliwell. C. P. Barrington-Leigh is the primary author in all regards. The *identification and design of the research program* for this paper were carried out jointly. Background *research*, the *data analysis*, and the *preparation of the manuscript* were performed by C. P. Barrington-Leigh, with comments on revisions provided by John F. Helliwell.

Chapter 1

Introduction

Homo sapiens is an extraordinarily social animal and economics has done much to contribute to the understanding of how its individual actions can add up to interesting group behaviour. In treating neat and tractable problems, however, economics may leave itself open to straying from the most important questions it ostensibly addresses. Humans like to seek hedonistic comforts and consumption, but they also like to help others, to be socially engaged, and to effect change in their lives and communities. Seeking non-market intangibles such as competence, autonomy, and relatedness [*Ryan and Deci*, 2000] are no less fundamental nor important parts of our human nature. In the last four decades empirical evidence has become increasingly available to address scientifically the question of what aspects of a life actually promote the outcome of subjective well-being or, loosely, our "happiness".

1.1 Happiness in economics

Should "happiness" be an objective in economics? It is not a natural one. In one sense, the social goal of humanity has in the past been genetic and cultural adaptation of the species to the given, possibly changing, environment. Individually, humans are assumed to have generally pursued the promulgation of their own genes.

What should be a measure of success for modern humanity? This is a hard question, but for economics to answer it without asking is worse than transparently failing to justify the answer. There is no naturally preferred direction for genetic evolution; for instance, bigger brains have often been a maladaptation due to their large energetic requirements. As a means of organising our perceptions and motivations, we are endowed by evolution with various internal measures of how things are going — for instance, (instantaneous) pleasure and (reflective) happiness or satisfaction.

Dystopian literature [*e.g., Huxley, 1932]* has pointed out the perils of taking one evolved, psycho-physiological motivating device, such as pleasure, as a primary objective. Nevertheless, the pursuit of reflective happiness is more easily defended as an important *a priori* goal for a civilisation than are many others, including consumption volume, production output, or sizes of choice sets which are ubiquitously used as proxies for economic well-being.

Numerous other personal performance measures (health, longevity, caloric intake, literacy, mobility, and so on) are available and relatively accessible in objective terms. A fundamental premise of the subjective well-being approach is that the relative importance of these metrics as social outcomes cannot be divined from economic or evolutionary theory. Ultimately, the question "What makes a good life?" cannot be answered with an arbitrary weighting of seemingly important factors. More significantly, it also cannot be answered satisfactorily by appealing to revealed preferences, especially for non-market goods for which choices and marginal tradeoffs are less easily observed and for which the variation in salient conditions may be small. Indeed, countless examples of dysfunctional or inconsistent individual behaviour are now available in economics as well as psychology and account for many societal constraints and supports such as the prohibition of illegal drugs and the inducement to save for one's old age.

To assess the quality of life experience one must then necessarily appeal to a subjective response. By asking people in some way to report the quality of their experience, we recognise that finding the determinants of well-being given our current evolutionary makeup and civilised situation is an empirical question.

Whether or not one is compelled by lofty policy implications of happiness research or by elevating subjective well-being to some kind of a "prime directive," understanding how circumstances and choices affect people's own life satisfaction, and that of others in their society, bears enough on the fundamentals of economics and policy to attract widespread attention.

Subjective measures of overall life satisfaction have been a growing component of psychological literature since the late 1960's [*Cantril*, 1965] and come in the form of single-question measures [*Diener*, 1984] and multi-question indices [*Diener et al.*, 1985]. For a recent assessment of the rôle of life satisfaction measures in economics, see *Dolan et al.* [2007].

Self-reported life satisfaction is a global assessment of life quality according to the respondent's own criteria and standards. It is not framed by any objective conditions of health, wealth, comfort, or any other of an interviewer's possible preconceptions of what might be an important factor. It refers to all aspects of life deemed relevant to the respondent and provides no information directly about an assessment of more specific domains of life, nor about which aspects the respondent considers important.

A large literature has dealt with the influence of personality and culture on self-reported life satisfaction [*Diener et al.*, 2003], and a pursuit of prominent contemporary interest is to evaluate the social and economic determinants of differences in life satisfaction across and within countries. The empirical papers in this dissertation contribute to this effort by focusing on the variation within one country, Canada. Chapter 4 provides some more background on the psychological literature on happiness.

1.2 Veblen preferences

One important feature of international well-being comparisons is the observation, currently in some dispute [*Stevenson and Wolfers*, 2008a], that economic growth gives rise to very limited benefits in terms of subjective well-being [*Easterlin*, 1974]. An early explanation of this fact was that the personal standards, or reference levels, for material well-being are not fixed in absolute terms but rather adapt to the norms perceived by the respondent.

A focus of two of the three manuscripts comprising this dissertation is the role of such consumption norms in setting individuals' preferences over consumption. That markets might be driven in part by this kind of feedback has always been acknowledged in economics [*Marx and Engels*, 1848; *Veblen*, 1899; *Pigou*, 1920; *Duesenberry*, 1949; *Galbraith*, 1958; *Duncan*, 1975], and utility functions which refer to others' consumption levels have come loosely to be called Veblen preferences. In fact, the *Veblen effect* originally referred to utility of consumption increasing with the price of a good [*Leibenstein*, 1950]. *Conspicuous consumption* is closely related but the nature of the conspicuousness is more general, while a *snob good* is one for which the utility is decreasing in the number of other people consuming like goods. I use *Veblen* as an adjective in a sense which conflates these meanings. Rather, it refers to general contemporary relativity in the well-being benefits of or preferences over consumption. In synonymous but possibly more conventional microeconomic terms, by *consumption externality* I refer to the effect of others' consumption when it acts directly in one's utility function. In terms both of behaviour (Chapter 3) and well-being (Chapter 2), I consider the possibility that our tendency to emulate others' average behaviour sets internal standards which affect the evaluation of outcomes.

1.3 Contributions

Below are outlined the main contributions of this dissertation, in the form of two empirical and one theoretical paper, to the literature on consumption externalities and the measurement and determinants of life satisfaction.

Chapter 2 uses a geographical multi-level regression approach to investigate the relationship between individual life satisfaction and the income of others. This empirical work assesses the degree to which the fulfillment we derive from consumption is lessened by the consumption of others. It also investigates the geographic scale characterising the population of others who comprise the implicit reference group in this comparison. Two main contributions are made.

- The work features multi-scale measurement of income externalities in Canadian urban regions, resolving different net effects of positive and negative influences of nearby households' mean income on individual life satisfaction.

- It makes innovative use of variables available from Canadian surveys and the census to explore variation in relative income effects between subpopulations and to test alternative explanations for these effects.

Chapter 3 tackles the theoretical problem of finding economic general equilibrium solutions with agents who have explicitly interdependent preferences. There are two main generalisations made over past modeling efforts of Veblen effects:

- Equilibria with heterogenous agents and heterogenous outcomes are admitted.
- Income reference levels are endogenous: rather than choose an arbitrary reference level, agents choose neighbours whose consumption levels are also features of equilibrium.

The resulting contributions are also twofold:

- I find a functional form for utility which admits an analytic solution of a general equilibrium market with endogenous reference group formation.
- I analyse welfare in resulting equilibria and find that eliminating pure Veblen goods is no longer necessarily Pareto preferred in a heterogeneous Veblen good economy. A political economy implication is that the wealthy may support production of and public investment in status goods while the poor prefer to eliminate them.

Chapter 4 returns to the empirical questions regarding life satisfaction but takes a step back to assess and address some of the difficulties of working with a subjective measure in economic analysis. Local weather conditions experienced by survey respondents on the day of their interview are used to assess the size of any bias resulting from transient affective influences on subjective response data and to test the validity of statistical inference about the determinants of subjective well-being.

- This represents the first time individual social survey data have been aligned with detailed weather records from the day and location of the interview to assess the importance of subjective response bias.
- The findings are supportive of standard empirical inference made about the determinants of life satisfaction and cast doubt on the general basis for concerns often held by economists regarding the use of any subjective measure as a dependent variable.

Chapter 5 concludes and mentions some avenues for extensions of the work presented here.

Bibliography for Chapter 1

Cantril, H., *The Pattern of Human Concerns*, Rutgers University Press, New Brunswick, N.J., 1965.

Diener, E., Subjective well-being, *Psychological Bulletin*, *95*, 542–575, 1984.

Diener, E., R. Emmons, R. Larsen, and S. Griffin, The Satisfaction With Life Scale, *Journal of Personality Assessment*, *49*, 71–75, 1985.

Diener, E., S. Oishi, R. Lucas, et al., Personality, culture, and subjective well-being: Emotional and Cognitive Evaluations of Life, *Annual Review of Psychology*, *54*, 403–425, 2003.

Dolan, P., T. Peasgood, and M. White, Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being, *Journal of Economic Psychology*, 2007.

Duesenberry, J., *Income, Saving, and the Theory of Consumer Behavior*, Harvard University Press, 1949.

Duncan, O., Does money buy satisfaction?, *Social Indicators Research*, *2*, 267–274, 1975.

Easterlin, R., Does Economic Growth Improve the Human Lot? Some Empirical Evidence, *Nations and Households in Economic Growth: Essays in Honour of Moses Abramovitz*, pp. 98–125, 1974.

Galbraith, J., *The Affluent Society*, Houghton Mifflin Books, 1958.

Graham, C., and S. Pettinato, Frustrated Achievers: Winners, Losers, and Public Perceptions in the New Global Economy, *Journal of Development Studies*, 2002.

Huxley, A., *Brave New World*, 1932.

Kingdon, G., and J. Knight, Community, Comparisons and Subjective Well-being in a Divided Society, *Journal of Economic Behaviour and Organization*, 2007.

Leibenstein, H., Bandwagon, Snob, and Veblen Effects in the Theory of Consumers' Demand, *The Quarterly Journal of Economics*, *64*, 183–207, 1950.

Marx, K., and F. Engels, Marx Engels selected works, Progress Publishers (1986), 1848.

Pigou, A., The Economic of Welfare, *London: Macmillam*, 1920.

Ryan, R., and E. Deci, Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being, *American Psychologist*, *55*, 68–78, 2000.

Stevenson, B., and J. Wolfers, Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox, 2008a.

Veblen, T., *The theory of the leisure class*, A.M. Kelley, 1899.

Chapter 2

Empathy and emulation: life satisfaction and the urban geography of comparison groups

2.1 Introduction

In many contexts of predictive analysis and policy framing, economists assume without evidence that desirable benefits accrue to humans based primarily on their absolute levels of consumption.[1] More broadly, it is conventional to focus without empirical justification on models in which (1) individual returns to behaviour greatly outweigh externalities and in which (2) changes in any reference levels intrinsic to utility vary less quickly than other factors relevant to behaviour. The first assumption may be counter-factual to the external marginal effect of, for instance, the intrinsic and shared pleasure of human social interaction, or of the reference-framing comparisons which can motivate consumption and determine satisfaction.[2] In the modern tra-

[1] A version of this chapter will be submitted for publication as Barrington-Leigh, C.P. and Helliwell, J.F., 'Empathy and emulation: life satisfaction and the urban geography of comparison groups.' The electronic version of this document offers fully hyperlinked cross-references throughout. We are grateful for support from SSHRC, the Canadian Institute for Advanced Research (CIFAR) and from Statistics Canada through UBC's Interuniversity Research Data Centre. This research forms part of the CIFAR Program on Social Interactions, Identity and Well-Being.

[2] There is often confusion over the claims and implications of happiness research, resulting mainly from a common confusion in teaching and practice of economics at the earliest stage. To make this explicit, it is necessary to remember that economists mean two entirely separate things by *utility*. One use of utility is in describing behaviour. This is a "positive" undertaking; that is, it is characterised by a falsifiable proposition. The proposition is that on average, at least in somewhat static situations, human behaviour is characterised by the optimisation of some well-behaved and stable function, the *decision-making* utility. A second meaning of utility is the original sense of the word, by which Jeremy Bentham meant *well-being*. There is no falsifiable proposition associated with the normative choice of using revealed preferences as an objective of policy. This is simply a value choice. One does not need to believe that humans maximise some preference function in order to hold as a value that policy ought to maximise people's choice sets. One does not need to hold the maximisation of economic choice as a core value in order to believe that behaviour is roughly characterised by rational (decision-making) utility maximisation. The two claims are orthogonal. One is subject to scientific testing and one is not. The two are, however, commonly confounded by the use of the word utility to imply both welfare and revealed preferences.

Advocates of taking self-assessed life satisfaction as a powerful measure of well-being need not question both of these claims. Their position is a normative one, a simple value judgment that the well-being we care about may be, or ultimately can only be, assessed by those experiencing it. An advocate of life satisfaction as an important objective dition, the focus of psychological and anthropological research differs starkly with that of economics in that regarding both motivation and well-being, economists focus on fixed preferences over absolute consumption while others view social comparisons and behaviour emulation as central phenomena in human societies.

Economists tend to be sympathetic to concerns about these missing aspects of human nature but often counter that "allowing" a broad range of influences on utility in models undermines the ability of economic arguments to explain anything non-tautologously. In fact, discussion of interdependent and non-constant preferences — in the context of status-seeking, habituation, conspicuous consumption and affluence, and relative versus absolute poverty — has steadily pervaded the literature on consumption behaviour and the labour-leisure choice since the early modern economists [*Marx and Engels*, 1848; *Veblen*, 1899; *Pigou*, 1920; *Duesenberry*, 1949; *Galbraith*, 1958; *Duncan*, 1975]. Modern evolutionary economic arguments [*Rayo and Becker*, 2004; *Eaton and Eswaran*, 2003] and corroborating neurological measurements [*Tobler et al.*, 2005; *Fliessbach et al.*, 2007], psychological studies, and economic inference provide overwhelming support for the claim that relative assessments figure prominently in our utility over consumption,[3] yet the detailed nature of these comparisons remains hard to measure and hard to incorporate into theory.

We pursue instead a more empirical approach. In recent decades the measurement of self-reported satisfaction with life (SWL) has increasingly been espoused as a new tool to assess the form of the utility function in a direct and quantitative way. The steadfast exclusive reliance on observed behaviour to reveal (or to compare) marginal utilities is giving way among economists to an increased interest in and acceptance of SWL as a window into well-being. This may allow the assumptions mentioned above to be assessed head on and invites the possibility of disentangling questions about behaviour from those about normative goals.

---

in policy may believe one way or another about behaviour being well explained by an optimisation process. Another proposition entirely separate from the value judgment would be that satisfaction with life is also a proxy for *decision-making utility* — that is, that people act to maximise their happiness. Just like the simpler neoclassical assumption of the existence of a utility function which rationalises behaviour, this is a falsifiable claim, unlike the value judgment that life satisfaction is an important policy objective.

The present work does not address the issue of rational decision making or behavioural maximisation of happiness. It *does* take for granted the untestable value statement that life satisfaction is a proxy for well-being. This serves as a motivation for the endeavour to determine empirically the influences on that well-being.

[3]*Rayo and Becker* [2004] argue, using a principal-agent framework, that our internal reward circuitry has finite bounds and therefore must have evolved with features that engineers would call automatic gain control and a (temporal) high-pass filter. That is, the offset and the scale for processing a consumption level into a psychological reward adapt to make best use of the available range of the reward experience. *Tobler et al.* [2005] mention a similar argument in explaining their observed neuronal activity. Dopamine neurons respond (*i.e.,* reward their host) in relation to the difference between the received versus anticipated payoffs rather than to absolute levels. In a controlled experiment using functional MRI to measure brain activity response to relative rewards, *Fliessbach et al.* [2007] find that midbrain regions known to be influenced both by primary rewards like food delivery and by more abstract incentives responded according to relative payment rewards, independently of the absolute level of payment. *Eaton and Eswaran* [2003] suggest a specific sense in which innate preferences should evolve to be jealous of one's competitors.

Nevertheless, while SWL scores can in statistical applications generally be treated as a cardinal measure of well-being[4] [*Frey and Stutzer*, 2002; *Ferrer-i Carbonell and Frijters*, 2004; *Krueger and Schkade*, 2008], the task of unravelling SWL's individual and average determinants remains a complex one. Indications point towards profound ramifications for policy, and some may already be obvious but the details lie ahead.

To date, a number of panel studies, particularly using European data, have addressed the question of relativities in life satisfaction due to income or consumption. Several of these studies show complete adaptation of reference levels for income over only a few years [*van de Stadt et al.*, 1985; *Clark*, 1999]. More generally, in a review of the literature, *Clark et al.* [2008] conclude that, due to the combined effects of comparison with contemporaries and adaptation over time, only about 13% of the short term marginal benefit of individual income changes would accrue after several years if the changes applied to everyone.

Such studies which resolve individual-level changes in fortune support predictions made earlier in explaining a lack of improvement of nationally averaged life satisfaction in nations experiencing rapidly increasing affluence [*Easterlin*, 1974].

In this paper, we address the question, "to whom do people compare their fortunes?" We focus on geographic aspects of consumption and income reference levels and on the counteracting social benefits of having prospering neighbours. Only a few studies have included geographically localised reference groups in the context of competitive consumption effects on SWL.[5] Using geography for delineating reference groups is partly a matter of convenience or, rather, a crude approximation to more probable and specifically matched comparison groups based on social distance. Nevertheless, the evidence corroborates the suspicion that individuals often exhibit implicit comparisons to geographically localised averages in determining their overall satisfaction. Our work is closely related to that of *Kingdon and Knight* [2007], who analyse both positive and negative externalities of average incomes on household satisfaction in South Africa. They use averages at two scales — village clusters and broader districts — and conclude from amongst several possible explanations that their findings are evidence of intrinsic empathy for those nearby and comparison with those slightly further away.

Helliwell and Putnam [2007] innovate by using geographic groups defined by census regions to assess the relativities and additivities in social capital due to education. They conclude that, at least for explaining a variety of measures of social engagement, such spatially defined reference groups are more appropriate than those constructed on the basis of similarity in personal characteristics without regard for geographic proximity. Their analysis does not relate as directly to the subjective evaluation of overall well-being, such as we pursue here, yet inasmuch as people compare themselves with those they know or see, one may expect a similarly

[4] Or what economists call utility in the original sense of Jeremy Bentham. Where utility implies instead a value whose maximisation motivates behaviour, the question is, as already mentioned, distinct and partly still open.

[5] We mention here only studies in which reference groups are more localised than an entire country. *Ferrer-i Carbonell* [2005a] also separates reference groups according to East and West Germany even after unification. See *Clark et al.* [2008] for a review of the SWL effects of income more generally.

important influence of neighbours in our study.

In further confirmation of the importance of proximity, *Knight and Song* [2006] report preliminary results from a survey in rural China in which respondents were asked explicitly about their comparison groups. The vast majority reported that their main comparison group consisted of either neighbours or fellow villagers rather than kin or people in the township or from broader geographic regions.

In Canada, a first look at our current question using Canadian surveys was conducted by *Helliwell and Huang* [2005], who included average incomes at the level of the Census Tract in a regression for SWL. They found that the externalities of reference levels at this scale mostly or entirely negated the individual benefits to marginal variation in income.

With considerably more detailed analysis on this question, *Luttmer* [2005] uses individual data from the U.S. National Survey of Families and Households to estimate the SWL effect of local average incomes on individuals. He also finds no net social benefit to increasing incomes using reference localities consisting of about 150,000 inhabitants. In contrast we will at our finest scale make use of regions in Canada with a median of 530 inhabitants.

In a rare natural experiment over neighbourhood selection, *Oreopoulos* [2002] found no neighbourhood effects on labour market outcomes in a small sample of households randomly assigned to housing projects in different Canadian locations.

The principle of invidious income seeking has also been used in revealed preference models. In a working paper, *Vigdor* [2006] uses the shape of local income distributions across the United States to explain voters' differing tendencies to support redistributive policy. When local geographic reference groups appear in preferences over relative income, the seemingly counterintuitive support for regressive taxation by the poor can be explained as a rational response intended to optimise local relative position.

Given the pervasiveness and remarkable magnitude of the interdependence of welfare functions on geographic neighbours, understanding the scale and nature of local reference groups and mutually beneficial social groups is a desirable goal with possibly important implications for urban planning and all levels of fiscal and even trade policy. Our objective in this paper is to look for geographically localised influences on SWL at a variety of spatial scales in order to determine which are most important in a developed country like Canada. Popular accounts of "keeping up with the Joneses" next door suggest that at least in some neighbourhoods, emulation of conspicuous consumption by others is made at a very local scale. On the other hand, some research suggests that even national status is relevant, in a kind of competitive economic nationalism. Our contribution is distinguished from others by its focus on multiscale geography, its emphasis on urban inhabitants, and its use of Canadian data. Although we are able to resolve income gradients on the scale of $\sim$100 m, our main finding is that in Canada income comparisons exist and significantly dominate any counteracting effects primarily at the scale of census tracts and metropolitan regions, the latter being typically several tens of kilometers in scale.

Below we discuss the data and approach (Section 2.2), present the results of reduced form linear regressions in light of possible confounding effects and competing interpretations (Section 2.3) and discuss the implications of our findings (Section 2.4).

2.2 Data and method

We use life satisfaction reports, among other variables, from three surveys conducted across Canada: the second wave of the Equality, Security, and Community survey (ESC2, 2002-2003), the Ethnic Diversity Survey (EDS, 2002), and the General Social Survey Cycle 17 (GSS, 2003). See Appendix A.2 for more detail on, and differences between, these surveys.

The surveys comprise a total of $\sim$70,000 individuals and they have some key questions in common. Most importantly, respondents were asked to rate their overall life satisfaction on a 5 or 10 point scale. Numerous other questions relevant to social interactions and socioeconomic and cultural backgrounds were posed, and some of these variables will be employed below. A significant feature of the surveys which facilitates our geographic analysis is that all three provide six-digit postal codes of respondents' residences at the time of the survey. In dense urban regions these correspond to a resolution of about one street block, or $\sim$200 m. In this work we include only urban respondents of age 15 years or older and in all estimates we make use of household weights provided for EDS and GSS.

Complementing these individual samples is the public use version of the 2001 Canadian Census, which is released with many variables available at the *Dissemination Area* (DA) level. In cities, DAs are composed of one or more neighbouring blocks, with a population of 400 to 700 people, and they cover all of Canada. Recall that in *Luttmer* [2005] the resolution available is $\sim$150,000 inhabitants, and in the popular German panel, GSOEP, individual locations are poorly resolved. The availability of both survey and census information with extremely fine resolution makes the Canadian data attractive for our purpose, even though the surveys are cross-sectional and preclude modelling with individual fixed effects. Figure 2.1 and Figure 2.2 demonstrate some superficial relationships between geographically averaged survey and census data, and foreshadow certain results to come. In particular, when comparing urban regions within Canada, mean life satisfaction and mean income show an inverse relationship. By contrast, there is a strong positive relationship between life satisfaction and a measure of social capital, the mean reported trust in neighbours.

We also make use of the 2001 Census data at the individual level, but only to aggregate census-tract income means according to certain population subgroups detailed later in Section 2.3.9.

An equation representative of the majority of estimates to follow is the "ordered logit" equation for the log odds of individual i reporting a value $j+1$ or higher:

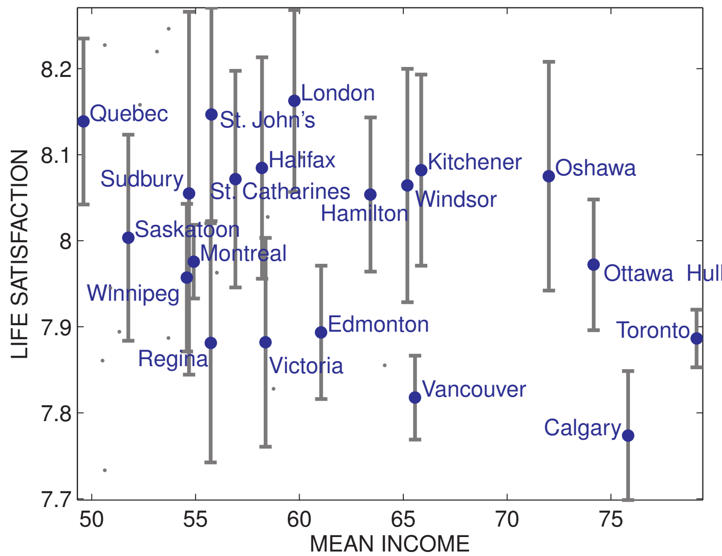


Figure 2.1: Life satisfaction and mean income (k$/yr) averaged by CMA.

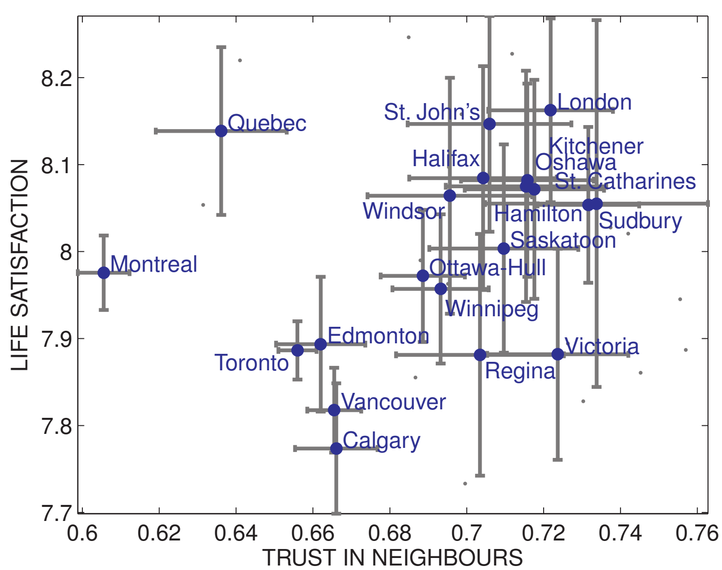


Figure 2.2: Life satisfaction and self-reported trust in neighbours, averaged by CMA.

$$\log\left(\frac{\text{Prob}\left(\text{SWL}_i > j\right)}{\text{Prob}\left(\text{SWL}_i \leq j\right)}\right) = c_j + \alpha \cdot \mathbf{X}_i + \tilde{\beta} \cdot \mathbf{Y}_i + \varepsilon_i + \sum_r \left(\Delta_r \cdot \left[\mathbf{Y}_i - \bar{\mathbf{Y}}_{R_{ir}}\right] + \gamma_r \cdot \mathbf{Z}_{R_{ir}} + \nu_{R_{ir}}\right) \tag{2.1}$$

Here the $\mathbf{X}_i$ are personal characteristics affecting individual i's well-being such as employment, marital status, health, and personality. In the empirical analysis to follow, a distinction will be made between relatively objective characteristics and those that rely strongly on a subjective self-assessment. $\mathbf{Y}_i$ are variables such as income which may influence SWL both absolutely and relatively. The region R_{ir} is the census region of scale r around individual i. Coefficients on relative levels $\mathbf{Y}_i - \bar{\mathbf{Y}}_{R_{ir}}$ are allowed to vary independently for each comparison region scale r. $\mathbf{Z}_{R_{ir}}$ represents other variables describing the geographic scale r around individual i which either do not have individual counterparts or are not thought to enter the utility function in a relative way.

We use an ordered logit[6] model(2.1) in order to estimate the underlying, or experienced, well-being through the discrete measure available from surveys. By using this formulation we need not rely on a cardinal interpretation of SWL.

Equation (2.1) is equivalent to a slightly more convenient form,

$$\log\left(\frac{\text{Prob}\left(\text{SWL}_i > j\right)}{\text{Prob}\left(\text{SWL}_i \leq j\right)}\right) = c_j + \alpha \cdot \mathbf{X}_i + \beta \cdot \mathbf{Y}_i + \varepsilon_i + \sum_r \left[\delta_r \cdot \bar{\mathbf{Y}}_{R_{ir}} + \gamma_r \cdot \mathbf{Z}_{R_{ir}} + \nu_{R_{ir}}\right] \tag{2.2}$$

when $\tilde{\beta} \equiv \beta - \sum_r \Delta_r = \beta + \sum_r \delta_r$. For instance, consider the case when Y_i represents own income and $\bar{\mathbf{Y}}_{R_{ir}}$ average local incomes. Then β represents the marginal effect on the log odds of an increase to own income, while the marginal effect of a simultaneous, uniform increase to everyone's income is $\tilde{\beta}$, the sum of all the coefficients on incomes in equation (2.1). It may be noted that since we use logarithms[7] of dollar income values, the functional form of equation (2.1) constrains the comparisons to enter the well-being function in the form of ratios.

For the estimates in this paper, the geographic reference areas are simply the fixed regions defined by the census.[8] These are each one of: 49,000 Dissemination Areas (DAs) with median

[6]This is also known as a "proportional (log) odds model" for obvious reasons. An alternative ordinal model, the ordered probit, is often used in the subjective well-being literature. However, similar results are typically found from OLS and either ordinal method [*Ferrer-i Carbonell and Frijters*, 2004]. Ordered logit has the advantage of a simple interpretation of coefficients, since the marginal effect of a covariate on the log odds is constant, regardless of the values of other covariates. We test other methods below.

[7]*Helliwell and Huang* [2005] find that linear income can be excluded when both linear and logarithm forms are included in a regression of life satisfaction.

[8] We used two techniques to provide contextual variable averages around each individual for a subset of census

population ∼540; 4800 Census Tracts[9] (CTs) with median population ∼4300; 5600 Census Subdivisions (CSDs) which in urban areas usually correspond to city boundaries; 27 Census Metropolitan Areas (CMAs); and 10 Provinces (PRs) containing at least one CMA.

The use of an invariant set of census regions makes possible another tool for isolating the contextual effects under study. This is to include geographic fixed effects at a given level of geography in order to identify spatial relationships at the next smaller scale. The appropriate modification to equation equation (2.2) is then:

$$\log\left(\frac{\mathrm{Prob}\,(\mathrm{SWL}_i > j)}{\mathrm{Prob}\,(\mathrm{SWL}_i \leq j)}\right) = c_j + \alpha \cdot \mathbf{X}_i + \beta \cdot \mathbf{Y}_i + \varepsilon_i + \sum_{r<R}\left[\delta_r \cdot \bar{\mathbf{Y}}_{\mathrm{CR}_{ir}} + \gamma_r \cdot \mathbf{Z}_{\mathrm{CR}_{ir}} + \nu_{\mathrm{CR}_{ir}}\right] + \phi_{\mathrm{CR}_{iR}} \qquad (2.3)$$

where CR_{ir} is the census region of scale r which contains respondent i's residence, r now indexes the census region scale in order of increasing size (DA, CT, CSD, CMA, and PR), $\phi_{\mathrm{CR}_{iR}}$ is a geographic fixed effect for some scale R, and where the equation only resolves local relative income effects at spatial scales r smaller than R. A source of endogeneity of particular interest in this study arises when unmeasured and geographically autocorrelated factors are related to both income and life satisfaction. In equation equation (2.3), the coefficient δ_{R-1} on the contextual effects of the largest resolved scale is unbiased by any unmeasured geographic variation present at the scale of R. For instance, consider the unmeasured influence of regional price levels, differences in government quality, cultural factors affecting community strength or lifestyle choices, and variation in climatic or other geographic amenities. Each of these possible missing variables represents a source of endogeneity because geographic variation captured only in the error term $\nu_{\mathrm{CR}_{ir}}$ may be causally correlated with a component of SWL captured only in ε_i. As a result, all coefficients on smaller-scale contextual effects would be biased. If these unmeasured influences exist, for instance, at the CMA level, then including dummy variables $\{\phi\}$ for each CMA will eliminate bias on the remaining coefficients. By separately running a series of estimates using

and survey variables and for a range of spatial scales. In one computationally intensive method, circles are drawn around each respondent's location at radii of 100 m, 800 m, 2 km, 4 km, 10 km, 20 km, and 100 km. Survey variable aggregates are formed by averaging over respondents lying in the inner circle or in one of the annuli defined between successive circles. The respondent is excluded from the inner circle. Census variable aggregates are formed by overlaying the circles on a map of polygons defining one size of census region (for instance, the DAs). For the inner circle and for each annulus, a weight is assigned to each census polygon according to its fraction lying within the aggregation region, and these weights are multiplied by population counts in each census region to generate appropriately weighted means of the desired variables. We do not find a qualitative difference in results between this method and the simpler one with fixed regions and thus prefer the simpler one. In order to eliminate spurious correlation of error terms, each reference region calculated for an individual in this simpler method also excludes the next smallest census region containing the individual.

[9] Census Tracts and Metropolitan Areas are special in that they exist only in urban regions and that some variables, such as those to do with the detailed distribution of income, are only offered by Statistics Canada for CTs. For urban regions we are able to aggregate these variables up from the CTs to the larger regions.

fixed effects at different values of geographic scale R, the set of coefficients $\{\delta_{R-1}\}$ for R corresponding to CT, CSD, and CMA[10] can be extracted and interpreted as the local Veblen effect at each scale.

We make use of a number of objective and some subjective controls in $\mathbf{X}$ and $\mathbf{Z}$. See *Helliwell* [2003] for a study of similar individual variables and national measures of social capital which prove to be significantly correlated with SWL in 46 countries. Our controls also include a measure of psychological coping resources from a series of questions in the GSS. As discussed by *Helliwell and Huang* [2005], this measure of "mastery" is likely to over-correct for personality since it is likely correlated with outcomes (in particular, incomes) but it is useful in the absence of panel data and individual fixed effects.

2.3 Results and interpretation

In this section we present our main findings and test them against several reasonable "classical" explanations for the observed correlations between own and others' income. We find evidence of a strong relative income effect at certain geographic scales. This effect appears to be stronger for those who are likely to know their region better, which is consistent with an explanation based on contemporary reference setting. We further show that not all determinants of well-being contribute in a predominantly relative way.

2.3.1 Classical regression

Table 2.1 on page 10 presents results from a fairly conventional series of regressions for life satisfaction among urban survey respondents. Each non-shaded column reports coefficients and standard errors for one regression using data from the survey indicated in the row labeled "survey". In all cases shown, coefficients are from an ordered logit model and are displayed in raw, unexponentiated form.[11] For example, column (3) in Table 2.1 indicates that a factor 10 increase in household income, holding other variables constant, is associated with a 34% increase (since $e^{0.29} \approx 1.34$) in the predicted odds of being at least one step higher on the standard ten-point SWL scale.

The first three columns of Table 2.1 record estimates of similar models carried out separately on data from urban respondents in each of three surveys. Missing coefficients reflect the lack

[10]Limits of the sample size and available computing power both made the use of fixed effects at the DA level impractical. Because many provinces are dominated by one or a few CMAs, province-level fixed effects were generally not used either.

[11]This provides easy identification of positive and negative effects based on the sign of the coefficient. In accordance with equation (2.1), an exponentiated coefficient represents the modeled change, for a one unit increase in the covariate, of the ratio of probabilities of reported SWL being in a higher category to that of it being in any lower one. In the ordered logit model, this marginal influence on probability is the same for any values of the other covariates — and thus at any level of life satisfaction.

	(1)	(2)	(3)	⟨1-3⟩	(4)	(5)	(6)	⟨4-6⟩	(7)	(8)	(9)	⟨7-9⟩	(11)	(12)	⟨11-12⟩
log(HH inc)	**.48***	**.21***	**.29***	**.26***	**.52***	**.23***	**.31***	**.29***	**.57***	**.25***	**.32***	**.29***	**.30***	**.71***	**.37***
	(.12)	(.046)	(.055)	(.034)	(.12)	(.046)	(.044)	(.031)	(.15)	(.049)	(.053)	(.035)	(.097)	(.21)	(.088)
health	**1.64***		**2.73***	**2.55***	**1.61***		**2.74***	**2.40***	**1.56***		**2.77***	**2.56***		**3.25***	**3.25***
	(.21)		(.093)	(.085)	(.17)		(.11)	(.094)	(.25)		(.12)	(.10)		(.39)	(.39)
trust-N	**.50***	**1.73***	**1.03***	**1.13***	**.51***	**1.80***	**1.01***	**1.01***	**.57***	**1.84***	**1.06***	**1.14***	**2.25***	**.99***	**1.77***
	(.11)	(.095)	(.083)	(.054)	(.076)	(.097)	(.12)	(.053)	(.092)	(.092)	(.070)	(.048)	(.18)	(.23)	(.14)
married	**.55***	**.44***	**.41***	**.44***	**.58***	**.47***	**.41***	**.45***	**.57***	**.46***	**.40***	**.44***	**.74***	.29	**.67***
	(.11)	(.060)	(.065)	(.041)	(.12)	(.056)	(.049)	(.035)	(.12)	(.057)	(.059)	(.039)	(.12)	(.27)	(.11)
asmarried	***.34***	**.51***	**.39***	**.42***	***.26***	**.42***	**.32***	**.35***	.19	**.45***	**.29***	**.34***	***.46***	***.56***	**.50***
	(.14)	(.090)	(.073)	(.053)	(.13)	(.066)	(.055)	(.040)	(.16)	(.088)	(.078)	(.055)	(.20)	(.26)	(.16)
separated			**−.44***	**−.44***			**−.46***	**−.46***			**−.46***	**−.46***		**−1.13***	**−1.13***
			(.10)	(.10)			(.078)	(.078)			(.092)	(.092)		(.35)	(.35)
divorced	−.24		−.077	−.11	−.16		−.087	*−.092**	−.26		−.096	*−.11**		.085	.085
	(.16)		(.086)	(.076)	(.21)		(.057)	(.055)	(.21)		(.072)	(.068)		(.36)	(.36)
widowed	.25		.001	.062	.32		−.020	.051	.30		−.050	.004		.14	.14
	(.22)		(.13)	(.11)	(.24)		(.12)	(.11)	(.33)		(.14)	(.13)		(.48)	(.48)
male	*−.12**	*−.068**	**−.16***	**−.12***	**−.12***	*−.073**	**−.17***	**−.13***	−.091	*−.066**	**−.17***	**−.12***	***−.18***	−.18	***−.18****
	(.071)	(.040)	(.039)	(.026)	(.046)	(.038)	(.032)	(.022)	(.078)	(.039)	(.037)	(.025)	(.077)	(.18)	(.071)
noReligion		−.011	**−.19***	**−.13***		.093	**−.14***	**−.10***		.12	**−.12***	−.058	.20	−.22	.060
		(.063)	(.046)	(.037)		(.079)	(.037)	(.034)		(.077)	(.045)	(.039)	(.13)	(.18)	(.10)
godImportance	**.47***	**.54***	**.35***	**.44***	**.57***	**.60***	**.40***	**.51***	**.59***	**.62***	**.41***	**.52***	**.82***	**.61***	**.76***
	(.12)	(.067)	(.059)	(.041)	(.075)	(.070)	(.061)	(.039)	(.088)	(.068)	(.057)	(.039)	(.14)	(.20)	(.11)
student		**1.26***	**.67***	**1.02***		**1.26***	**.66***	**.99***		**1.25***	**.54***	**.93***	**1.65***	.57	**1.48***
		(.14)	(.17)	(.11)		(.15)	(.16)	(.11)		(.16)	(.18)	(.12)	(.29)	(.69)	(.27)
employed		**1.19***	**.59***	**.95***		**1.19***	**.59***	**.93***		**1.19***	**.48***	**.84***	**1.45***	.36	**1.27***
		(.13)	(.16)	(.099)		(.14)	(.15)	(.10)		(.16)	(.16)	(.11)	(.26)	(.60)	(.24)
domestic		**1.07***	**.71***	**.92***		**1.07***	**.71***	**.89***		**1.10***	**.58***	**.82***	**1.26***	.19	**1.08***
		(.15)	(.17)	(.11)		(.15)	(.14)	(.10)		(.17)	(.16)	(.12)	(.29)	(.63)	(.26)
unemployed	**−.85***		−.13	**−.51***	**−.82***		−.14	*−.37**	**−.78***		−.24	**−.45***		−.32	−.32
	(.20)		(.20)	(.14)	(.29)		(.21)	(.17)	(.26)		(.21)	(.17)		(.69)	(.69)
retired		**1.42***	**.85***	**1.17***		**1.40***	**.85***	**1.12***		**1.42***	**.72***	**1.03***	**1.75***	.26	**1.38***
		(.15)	(.17)	(.11)		(.13)	(.13)	(.090)		(.19)	(.17)	(.12)	(.31)	(.53)	(.27)
age	**−.065***	**−.055***	**−.086***	**−.071***	**−.064***	**−.059***	**−.088***	**−.075***	**−.063***	**−.063***	**−.091***	**−.077***	**−.072***	***−.086***	**−.075***
	(.014)	(.009)	(.008)	(.005)	(.014)	(.009)	(.007)	(.005)	(.019)	(.009)	(.008)	(.006)	(.016)	(.034)	(.014)
$(\mathrm{age}/100)^2$	**8.14***	**5.58***	**8.47***	**7.29***	**7.82***	**5.94***	**8.66***	**7.61***	**7.83***	**6.38***	**8.99***	**7.79***	**7.50***	***7.59***	**7.52***
	(1.41)	(.97)	(.92)	(.60)	(1.26)	(.94)	(.77)	(.54)	(1.82)	(.98)	(.91)	(.63)	(1.83)	(3.78)	(1.65)
CMA f.e.					✓	✓	✓	✓							
CSD f.e.									✓	✓	✓	✓			
CT f.e.													✓	✓	✓
survey	E2	ED	G17	⟨3⟩	E2	ED	G17	⟨3⟩	E2	ED	G17	⟨3⟩	ED	G17	⟨2⟩
obs.	2633	24113	12970	39716	2535	24113	12970	39618	2013	23468	12197	37678	8454	1397	9851
pseudo-R^2	.037	.053	.062		.039	.058	.064		.044	.069	.069		.167	.100	
N_{clusters}					30	42	46		47	221	192		762	111	

Table 2.1: A "classical" regression for life satisfaction on household income and personal characteristics. Estimated coefficients are shown from ordered logit models of SWL. Standard errors (in parentheses) are calculated using clustering whenever geographic fixed effects (f.e.) are indicated. Surveys are identified with E2 for ESC2, ED for EDS, and G17 for GSS17. Shaded columns indicating by ⟨3⟩ that multiple surveys are included present weighted means of coefficients from estimations carried out separately for each survey. Not shown are a series of controls for household size. Only urban respondents are included. Significance: **1%*** ***5%*** *10%**

of certain questions in some surveys. The fourth, shaded column contains mean coefficients for each covariate, calculated by weighting each individual estimate by the inverse square of its standard error. When a variable is only available from a subset of the surveys, the mean shown reflects the coefficients from available regressions. The geographic fixed effects described in equation equation (2.3) are accounted for by including dummies at one level of census geography, as indicated by the rows CMA f.e. for metropolitan area fixed effects, CT f.e. for census tract fixed effects, and so on. Standard errors are calculated using clustering with the same groups as used for the geographic fixed effects.

Unlike the majority of results to follow, the explanatory variables in Table 2.1 do not include regional averages of income. A standard interpretation of the positive coefficients for household income (in $\log_{10}$ form) found for this specification is that increasing incomes can be expected to benefit average SWL. The results also show that most coefficients, including that on household income, are relatively unaffected by the inclusion of regional fixed effects. Unsurprisingly in light of the existing literature, measures of self-reported health, trust in neighbours, religiosity, involvement in a marriage-like relationship, youth, and old age have positive and significant partial correlations with SWL. Being unemployed and being male are each negative predictors of reported well-being. Included in all regressions but not shown are dummy variables for household size. Categories of 1, 2, 3, 4, 5, and >5 occupants are admitted in order to account for different impacts of household income on survey respondents.

Aside from the self-reported health and trust variables, the set of non-income controls used in this table will frequently be used but not shown explicitly in subsequent estimates. Although religiosity is included among them, we consider these variables to be relatively objective attributes as compared with health and trust. These latter subjective measures may be influenced by the respondent's personality type and current level of affect at the time of the interview.[12] They are nevertheless considered to be important and distinct determinants of SWL and, if anything, can be expected to correct partly for the individual variation in optimism and personality type which might play into SWL responses.

The row labeled "pseudo-R^2" provides a measure of the explained portion of individual variation in the dependent variable. It is generally believed that all but about 10%-20% of SWL variation between adult individuals is due to predetermined individual characteristics [*Diener*, 1984], which gives rise to a low pseudo-R^2 in all our estimates. The table shows that progressive inclusion of fixed effects at the province, metro area, city, and census tract level has the result of increasing the explained portion of individual variation without significantly changing other coefficients.[13] Similar results to these are obtained (but not shown) using an OLS model. In that case, the R^2 varies as high as 0.39 in the case with local fixed effects at the census tract level.

[12]See, however, *Barrington-Leigh* [2008b] for an effort to quantify this influence.

[13]The number of observations diminishes considerably when CT dummies are included in the equation, so the corresponding rise in the explained portion is less remarkable in this case. The set of included respondents is in each case restricted by the exclusion of regions with few samples. For the regressions shown here, the minimum allowed cluster size was 9.

This suggests that, despite the large idiosyncratic variability in reported SWL, localised factors are an important determinant of SWL.

2.3.2 Veblen effects

Table 2.2 shows estimates of the same equations as Table 2.1 but now augmented with reference income levels. The coefficient β on the logarithm of own household income now represents the individual marginal benefit of income when others' incomes are held constant. Rows (1) and (2) indicate that a factor of 10 increase in own household income, holding others' constant, is associated with only a 20% or 30% increase in the probability of being one point higher out of 10 in SWL. This small value is consistent with previous studies. It is also similar to that found in the previous "classical" regression, likely reflecting the fact that respondents predominantly live in large, high-income cities. The row labelled "$\sum\beta_{\text{inc}}$" shows $\tilde{\beta}$, the sum of the various income coefficients (see (2.2)). This is the net social benefit of marginal changes to the household income of oneself and of everyone else in one's own CMA. This value is significantly negative, indicating that, holding other factors constant, respondents in metro areas with higher average income tend to report a significantly lower satisfaction with life. This reduced-form result appears to be stronger than that found in other studies. It does not, however, imply that raising the income level of all metro regions at once would result in decreased well-being, since all national-level effects, including federal public goods funded by income taxes, are held constant in the present analysis. Clearly to encompass all these channels of influence one must appeal to cross-country comparisons.

Reminiscent of the findings of *Kingdon and Knight* [2007] is the positive coefficient generally found on the most local geographic reference group's income along with negative coefficients on the mean income of wider regions.

As described in Section 2.2, these reference levels are based on mean incomes reported in the 2001 census and exclude residents of the next smallest census region containing each respondent. For instance, the CSD average income is calculated for each survey respondent as the mean household income amongst residents who live in the respondent's CSD but not in his or her CT. CSD mean income, which is likely to be related through taxes to the amount of funding in the civic jurisdiction, receives an insignificant coefficient. In general, progressively incorporating fixed effects does not significantly alter the estimated coefficients. This indicates that our measures of mean census income are not proxying for other, unmeasured geographic characteristics, and that collinearity between income at different geographic scales is not driving the results.

2.3.3 Exposure response

If the estimates of negative spillover effects just described are truly a reflection of an adaptable reference level acting in respondents' assessments of their own well-being, then we can expect

	(1)	(2)	(3)	⟨1-3⟩	(4)	(5)	(6)	⟨4-6⟩	(7)	(8)	(9)	⟨7-9⟩	(11)	(12)	⟨11-12⟩
log(HH inc)	.26	**.23***	**.32***	**.27***	.29	**.23***	**.31***	**.26***	***.41***	**.25***	**.30***	**.27***	**.28***	**.70***	**.35***
DA: log(HH inc)	.093	.14	***.41***	**.28***	.24	.15	**.41***	**.28***	.81	.13	***.42***	**.32***	.58	−.18	.27
CT: log(HH inc)	.14	−.33	**−.58***	**−.43***	.031	*−.33**	**−.60***	**−.47***	*−.91**	−.29	**−.61***	**−.51***			
CSD: log(HH inc)	−1.46	.11	−.43	−.24	−.72	.21	*−.36**	−.12							
CMA: log(HH inc)	−.56	**−1.13***	**−1.08***	**−1.08***											
$\sum\beta_{\text{inc}}$	−1.53	***−.98***	**−1.36***	**−1.19***	−.15	.26	−.25	.26	.31	.093	*.12**	***.11****	**.86***	***.52***	**.66***
health	**1.04***		**2.74***	**2.61***	**.87***		**2.75***	**2.59***	*.76**		**2.78***	**2.66***		**3.29***	**3.29***
trust-N	**.49***	**1.73***	**1.05***	**1.25***	**.40***	**1.80***	**1.05***	**1.27***	**.47***	**1.84***	**1.09***	**1.26***	**2.22***	**1.04***	**1.81***
married	**.63***	**.45***	**.44***	**.45***	**.69***	**.47***	**.45***	**.46***	**.77***	**.46***	**.42***	**.46***	**.73***	.30	**.67***
asmarried	*.37**	**.46***	**.35***	**.40***	.27	**.42***	**.34***	**.37***	.47	**.45***	**.31***	**.38***	***.45***	***.57***	**.49***
separated			**−.45***	**−.45***			**−.44***	**−.44***			**−.47***	**−.47***		**−1.07***	**−1.07***
divorced	−.24		−.098	−.11	−.15		−.086	−.087	−.17		−.10	−.11		.048	.048
widowed	.10		−.029	−.009	.084		−.037	−.008	***.49***		−.072	.072		.026	.026
male	−.018	*−.078**	**−.16***	**−.11***	−.11	***−.082***	**−.16***	**−.12***	−.11	*−.076**	**−.17***	**−.12***	***−.17***	−.21	***−.17****
noReligion		.009	**−.21***	**−.13***		.083	**−.18***	**−.13***		.11	**−.15***	*−.074**	*.21**	−.18	.089
godImportance	***.36***	**.56***	**.38***	**.45***	**.44***	**.59***	**.40***	**.50***	***.45***	**.62***	**.41***	**.51***	**.83***	**.62***	**.76***
student		**1.27***	**.65***	**1.03***		**1.26***	**.64***	**.99***		**1.24***	**.53***	**.93***	**1.57***	.43	**1.39***
employed		**1.20***	**.58***	**.98***		**1.20***	**.57***	**.93***		**1.18***	**.47***	**.84***	**1.38***	.23	**1.20***
domestic		**1.07***	**.69***	**.92***		**1.08***	**.69***	**.88***		**1.08***	**.58***	**.82***	**1.20***	.056	**1.00***
unemployed	**−1.17***		−.19	**−.51***	***−.99***		−.21	***−.42****	**−.94***		−.29	***−.50****		−.24	−.24
retired		**1.41***	**.84***	**1.17***		**1.40***	**.84***	**1.13***		**1.40***	**.71***	**1.03***	**1.65***	.089	**1.28***
age	−.030	**−.059***	**−.091***	**−.071***	*−.027**	**−.061***	**−.093***	**−.078***	*−.035**	**−.065***	**−.093***	**−.075***	**−.070***	***−.084***	**−.073***
$(\text{age}/100)^2$	***4.94***	**5.95***	**8.99***	**7.26***	**4.37***	**6.20***	**9.12***	**7.50***	**5.30***	**6.62***	**9.20***	**7.62***	**7.34***	***7.49***	**7.37***
CMA f.e.					✓	✓	✓	✓							
CSD f.e.									✓	✓	✓	✓			
CT f.e.													✓	✓	✓
survey	E2	ED	G17	⟨3⟩	E2	ED	G17	⟨3⟩	E2	ED	G17	⟨3⟩	ED	G17	⟨2⟩
obs.	1141	23589	12201	36931	1035	23589	12201	36825	806	22955	11429	35190	8257	1363	9620
pseudo-R^2	.038	.054	.064		.041	.058	.066		.050	.069	.070		.166	.099	
N_{clusters}					23	42	46		23	220	185		747	109	

Table 2.2: Baseline estimates of relative income effects. See Table 2.1 on page 10 for a description of the format. For compactness, standard errors are not shown. The $\sum\beta_{\text{inc}}$ row shows the sum $\tilde{\beta}$ of estimated coefficients on own and others' income. Significance: **1%*** ***5%*** *10%**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
log(HH inc)	**.27***	**.35***	**.28***	**.28***	**.35***	**.33***	**.28***	**.27***	**.36***	**.38***	**.31***	**.26***	**.24***	*.27**	**.20***	.14	**.37***	**.39***
DA: log(HH inc)	***.28***	.27	.23	.27	***.53***	*.47**	.30	.27	*.58**	*.56**	.22	.22	.23	.068	***.35***	.74	.050	.13
CT: log(HH inc)	−**.43***	−**.51***	−***.68***	−**.73***	−.42	−.39	−**.72***	−**.71***	−.21	−.44	−.19	−.039	−**.55***	−**.59***	−**.63***	−**.60***	−.28	−.34
CSD: log(HH inc)	−.24	−.12	−.31	−.28	−.68	−.67	−.27	−.24	*−1.04**	−.90	−.26	−.024	−.13	−.032	−.26	−.20	.21	.34
CMA: log(HH inc)	−**1.08***	−**1.08***	*−.91**	*−.91**	*−.96**	*−.96**	−***1.03***	−***1.03***	−.64	−.64	−.61	−.61	−**1.07***	−**1.07***	−***.67***	−***.67***	−**1.98***	−**1.98***
$\sum \beta_{\text{inc}}$	−**1.19***		−**1.40***		*−1.17**		−**1.44***		−.95		−.42		−**1.28***		−**1.03***		−**1.68***	
health	**2.61***		**2.81***		**2.67***		**2.83***		**2.53***		**2.72***		**2.77***		**2.81***		**2.62***	
trust-N	**1.25***		**1.15***		**.88***		**1.10***		**.90***		**1.58***		**1.29***		**1.52***		**1.00***	
controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
geo fixed effects		✓		✓		✓		✓		✓		✓		✓		✓		✓
$\tau_{\text{neigh}} \geq 10$yr			✓	✓	×	×												
$\tau_{\text{city}} \geq 10$yr							✓	✓	×	×								
foreign born											✓	✓	×	×				
own house															✓	✓	×	×
survey	⟨3⟩	⟨2⟩	G17	G17	G17	G17	G17	G17	G17	G17	⟨2⟩	⟨0⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨1⟩	⟨2⟩	⟨0⟩
obs.	36931	≥9620	7087	≥0	5114	≥0	9001	≥0	3200	≥0	9316	≥0	26474	≥3361	25867	≥4424	9923	≥0
pseudo-R^2			.067		.061		.067		.059									

Table 2.3: Summary of Veblen coefficient estimates for various subgroups. Weighted mean coefficients and, in parentheses, standard errors for mean coefficients are shown. These means are averaged across estimates carried out separately for each survey. Odd-numbered columns show estimates carried out without geographic fixed effects but restricted to specific subsets of the survey sample, as indicated by the rows containing ✓'s and ×'s. The even-numbered rows display mean coefficients $\{\delta_{R-1}\}$ extracted from the corresponding set of regressions carried out with geographic controls. In these rows, the coefficients on household (HH) income and on DA-level average income are generated (where possible) in regressions which include a complete set of geographic dummies at the CT level. The coefficients for income at the DA, CT, and CSD level are generated in regressions which include a complete set of geographic dummies at the next highest geographic level. The "survey" and "obs." rows for these columns show the minimum number of surveys and smallest total sample size used for any of the coefficients. For instance, "≥ 0" indicates that sample sizes were insufficient to complete any estimates with dummy variables at the smallest level, the CTs. Not shown are coefficients for the set of more objective (but including religiosity) individual controls which are shown in Table 2.1. The complete results underlying these mean coefficients are shown in the Appendix in Table A.1 on page 107. Significance: **1%*** ***5%*** *10%**

the strength of these externalities to depend not just on location but also on the degree to which respondents have been exposed to other people — or information about people — in each geographic region. Accordingly, we next estimate the geographic spillover effects of income for subpopulations which might be expected to have stronger or weaker ties to their home location.

Outside of the Appendix, a simpler form of tabulated estimates is given in most of the tables to follow. Table 2.3 exemplifies this summary format and its first two columns summarise all of Table 2.2. Column (1) contains the mean coefficients for the baseline equation already recorded in column $\langle 1-2 \rangle$ of Table 2.2. Column (2) of Table 2.3 compiles the mean coefficients on $\{\delta_{R-1}\}$ described in Section 2.2 and taken from columns $\langle 4-6 \rangle$, $\langle 7-9 \rangle$, and $\langle 10-12 \rangle$ in Table 2.2. These correspond to the estimated marginal benefit of a region's income when fixed effects at the next highest geographic scale are controlled for.

The remaining odd-numbered columns similarly show coefficients, averaged over surveys, from regressions without fixed effects but carried out over specific subsets of the survey sample, as indicated by the rows containing ✓'s and ×'s. The even-numbered columns display the coefficients $\{\delta_{R-1}\}$ from the corresponding set of regressions carried out with geographic controls. The row labeled "survey" indicates which survey or how many surveys were used. When fewer than three surveys are used it is because not all offer the criterion defining the particular subpopulation. For instance, columns (5) and (6) reflect the fact that only GSS17 includes a question about the length of neighbourhood tenure.

Columns (3) to (10) show that survey respondents who indicate tenure in their neighbourhood or city of at least ten years are more strongly and negatively affected by a higher income in their local region (CT).

Conversely, those who have relocated more recently appear to benefit more from the affluence of their close neighbours at the DA scale. There is also the suggestion that those who are "new" to the city may be less sensitive to CMA mean income than those who are new to their neighbourhood but may not be new to the city. Columns (11)–(14) indicate that the negative effect of nearby others' income on SWL is much stronger amongst native-born Canadians as compared with immigrants.

Homeowners and non-owners, shown in columns (15) to (18), differ in the dependence of their reported SWL on both their own household income and on others'. One may hypothesize that homeowners are likely to have lived in the same neighbourhood for longer, and therefore be more influenced by its norms. On the other hand, non-homeowners are likely to feel less secure and settled in regions with high incomes and house prices. These suppositions find support in the differences between coefficients on CT and CMA incomes for homeowners and renters, but alternative hypotheses will need to be addressed below for a confident interpretation.

2.3.4 Price levels

All income variables presented so far have been measured in nominal terms, uncorrected for price levels. One natural objection to finding a strongly negative coefficient on the metropolitan

	(1)	(2)	(3)	(4)
log(HH inc)	**.27***	**.35***	**.24***	***.27***
DA: log(HH inc)	***.28***	.27	*.28**	.12
CT: log(HH inc)	**−.43***	**−.51***	***−.41***	**−.43***
CSD: log(HH inc)	−.24	−.12	−.16	−.021
CMA: log(HH inc)	**−1.08***	**−1.08***	**−1.87***	**−1.87***
$\sum \beta_{\text{inc}}$	**−1.19***		**−2.00***	
health	**2.61***		**2.58***	
trust-N	**1.25***		**1.24***	
CMA prices			✓	✓
controls	✓	✓	✓	✓
geo fixed effects		✓		✓
survey	⟨3⟩	⟨2⟩	⟨3⟩	⟨2⟩
obs.	36931	≥9620	24094	≥6793

Table 2.4: Effect of CMA price correction. Summary of estimates in the format described on page 14. Estimates in the columns with "CMA prices" are carried out with all income measures corrected for CMA price level. Only CMAs for with available price indices are included. Significance: **1%*** ***5%*** *10%**

area's mean income in nominal terms is that this average is likely to reflect regional price levels. The negative coefficient could therefore reflect individuals' intrinsic assessment of their real income. Because inter-regional price level comparisons are difficult to carry out,[14] we cannot correct all income measures for local buying power. However, geographic fixed effects naturally account for any possible variation in local costs as well as geographic amenities. Assuming that mobility is high enough for CMA-level fixed effects to capture the main differences in the value of nominal incomes, it remains only to test our estimates of CMA-level effects using the available price comparators. Restricting the sample to ten major city regions for which Statistics Canada calculates comparative cost of living data and repeating our baseline estimate, we find the same pattern of coefficents, as shown in Table 2.4.

[14] Statistics Canada remains cautious in accounting for housing cost differences across locations, and therefore provides only very limited consumer price comparisons across Canada [Personal communication, Erwin Diewert]. In general, when geographic location confers amenity values, prices for real estate and even other local commercial goods may incorporate an associated premium. In principle, such premia may reflect physical characteristics of the location or an endogenous social value of exclusivity. See *Barrington-Leigh* [2008a] for a model of endogenous exclusivity in real estate value driven by pure Veblen consumption.

2.3.5 Wealth and income

Ideally, in a neoclassical formulation a better measure of lifetime expected wealth — or indeed of current consumption — would be included to predict SWL. We next test some alternative specifications in order to account for the possibility of mismeasuring an absolute consumptive contribution to SWL. *Luttmer* [2005] addresses the concern that relying on the log of mean household income as a reference value may just provide flexibility to accomodate an alternative, underlying functional form for households' own income. Table 2.5 shows a test against this possibility by incorporating, along with the dummy variables for household size that are always included, the respondent's own income and his or her household's income adjusted in a conventional way for family size.[15] Also in this specification are the respondent's housing payments, estimated house market value, and the nearby average of reported house values from the census. Because a primary form of savings for many households is in house ownership, living in an affluent area may proxy for owning at least part of a relatively expensive house. While a higher house value might imply higher mortgage payments for house owners and therefore less current consumption of other goods, it may also be a less noisy indicator of total wealth and thus future expectations of affluence than is current income.

The table shows in columns (1) to (4) mean coefficents from the available surveys. The final column summarises the geographic reference effect estimates using fixed effects at each level. The coefficients estimated with CT fixed effects suffer from a small sample size in one survey, which accounts for the large coefficient on own income; see Table A.2 on page 112 in the Appendix. As noted by *Helliwell and Huang* [2005], the dominance of household income over personal income, even for wage earners in a multi-person household, is evidence of empathy dominating over any relative income effects within the household unit.

Available in Canadian census data and the EDS survey is a question about the size of one's primary dwelling. One's own house size is a significant candidate for measures of conspicuous affluence, and thus Veblen effects, but a large and comfortable home may also represent a direct channel through which material consumption promotes SWL. In addition, a measure of local house sizes may be a further proxy for respondents' wealth or indebtedness. With these motivations, Table 2.6 reports a specification that includes measures of own and local house size. Other than a possible decrease in the strength of the CMA-level Veblen coefficient, we find no significant changes in income effects and no significant role for own or neighbours' dwelling sizes.

2.3.6 Life in the big city

Canada has a small number of large metropolitan areas, making it a difficult object of study for unpacking different CMA-level influences on SWL. It is possible that mean incomes are

[15] This "household equivalent" income measure is not used throughout most of the analysis because the inclusion of a set of separate controls for household sizes is a less restrictive specification.

	(1)	(2)	(3)	(4)	(5)
log(own inc)	.049	.038	*.11**	***.35***	***.35***
log(HH inc/$\sqrt{hh}$)	**.18***	**.22***	***.15***	.059	.059
DA: log(HH inc)	***.37***	***.36***	*.30**	.089	.089
CT: log(HH inc)	−**.56***	−**.60***	−**.67***		−**.67***
CSD: log(HH inc)	−.13	−.14			−.14
CMA: log(HH inc)	−**.87***				−**.87***
$\sum \beta_{\text{inc}}$	***−1.06***	.21	−.12	***.96***	
mortgagePayment	−.034	−.030	−.024	.050	
log(houseValue)	.13	.15	.095	−.006	
DA: log(houseValue)	−.15	−.10	.14	.11	.11
health	**2.71***	**2.81***	**2.84***	**3.19***	
trust-N	**1.27***	**1.52***	**1.37***	**1.82***	
trust-G	−.0004	.018	.021	−.14	
controls	✓	✓	✓	✓	✓
Geo dummies		CMA	CSD	CT	✓
survey	⟨3⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩
obs.	27634	26901	25486	4142	≥4142

Table 2.5: Summary of alternate measures of wealth and income. The first four columns represent coefficients averaged over surveys, as in the shaded columns of Table 2.2 on page 13. The fifth column shows summary coefficients of the kind described on page 14. Detailed estimates summarised in this table are found in Table A.2 on page 112. Significance: **1%*** ***5%*** *10%**

	(1)	(2)	(3)	(4)	(5)
log(HH inc)	**.20***	**.21***	**.21***	.14	.14
DA: log(HH inc)	.28	.21	.29	.35	.35
CT: log(HH inc)	***−.60***	**−.79***	***−.59***		***−.59***
CSD: log(HH inc)	−.37	−.22			−.22
CMA: log(HH inc)	*−.59**				*−.59**
$\sum \beta_{\text{inc}}$	**−1.05***	*−1.18**	.088	.49	
houseRooms	.002	.004	.006	.003	
DA: houseRooms	.007	.029	.008	.053	
CT: houseRooms	.006	.019	.001		
health	**2.61***	**2.66***	**2.85***		
trust-N	**1.33***	**1.30***	**1.63***	**2.21***	
trust-G	−.004	.018	.017	.009	
controls	✓	✓	✓	✓	✓
Geo dummies		CMA	CSD	CT	✓
survey	⟨3⟩	⟨3⟩	⟨2⟩	⟨1⟩	⟨1⟩
obs.	26990	26884	24486	4424	≥4424

Table 2.6: Own and neighbours' dwelling sizes. The first four columns represent coefficients averaged over surveys, as in the shaded columns of Table 2.2 on page 13. The fifth column shows summary coefficients of the kind described on page 14. Detailed estimates summarised in this table are found in Table A.3 on page 113. Significance: **1%*** ***5%*** *10%**

correlated with (*i.e.,* proxying for) the size of a metropolis and that the coefficient on mean CMA income is reflecting an omitted variable bias due to unmeasured negative qualities of big city life.

There are a number of such factors missing in the baseline equation which one might suppose to be correlated with both mean incomes and life satisfaction. At the risk of over-correcting for these factors, Table 2.8 summarises estimates from a specification incorporating the fraction of immigrants at CT and CSD scales, the population size and density (ρ), and local average trust levels. These variables are motivated by the fact that high density areas tend to hold a more transient population which may affect social capital and, in turn, SWL. Qualitatively, the results with these controls reproduce the patterns found in the baseline case, especially for the CT-level coefficients.

2.3.7 Status and signalling

A classical explanation for conspicuous consumption is that it confers signalling benefits on the consumer and is thus an investment for the future. Under this hypothesis, even if self assessments of consumption are intrinsically independent of others' fortunes, measures of relative affluence should still be correlated with subjective well being because individuals expect to derive benefits from their status-enhancing investments. These benefits could, for example, relate to a better match in the job market or to a better match in social circles. Indeed, there are compelling theoretical accounts of how conspicuous status-seeking can exist in a signalling equilibrium when unobservable abilities contribute to mutual productivity in business interactions between two people [*Rege*, 2008, for an example and references]. In this case Rolex watches and Armani suits may alleviate an imperfect information problem if investing in them and judging others on the same basis tend to improve the chance of making business connections with high-ability people. In this model utility is only derived classically from absolute material consumption and, *Rege* [2008] argues, the availability of such snob goods may be welfare improving overall as the efficiency gain from better matching can outweigh the "rat-race" loss due to wasteful over-consumption.

Our present empirical focus is not on consumption of specific snob goods but on overall consumption and income. In models of concern for relative wealth more generally, *Cole et al.* [1992] and *Corneo and Jeanne* [1999] explain apparent relative income effects as being generated by (mate) matching considerations, and describe how such endogenously generated concern for relative wealth may be "beneficial for economic growth." On this basis we would expect the negative effects of local reference levels to diminish in those who have married, gained a secure career, finished their work career, and so on.

To address the prediction of a diminishing desire for signalling investments over time, Table 2.7 splits respondents up by age. Controls are included and regions other than the DA and CMA are excluded for simplicity. Between the ages of 25 and 64 it may be seen that the negative comparison effect of income at the CMA level is undiminished.

	25-34	35-54	55-64	65+
	(1)	(2)	(3)	(4)
log(HH inc)	**.24***	**.40***	**.29***	***.13***
	(.06)	(.04)	(.07)	(.06)
DA: log(HH inc)	−.006	−.06	−.14	−.19
	(.10)	(.06)	(.11)	(.12)
CMA: log(HH inc)	**−.64***	**−.74***	**−.81***	−.10
	(.20)	(.14)	(.27)	(.30)
trust-N	**.91***	**1.37***	**1.37***	**1.33***
	(.14)	(.10)	(.19)	(.21)
health	**2.09***	**2.05***	**1.74***	**2.08***
	(.17)	(.11)	(.19)	(.20)
Obs.	6002	12844	3435	2860
pseudo-R^2	.02	.03	.04	.05

Significance: **1%*** ***5%*** *10%**

Table 2.7: Income effects and age. Ordered logit coefficients for SWL using pooled surveys. Table headings indicate age ranges in years. Our standard demographic controls are included.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
log(HH inc)	**.27***	**.35***	**.28***	**.27***	**.36***	**.34***	**.28***	**.27***	**.37***	**.40***	**.30***	**.27***	**.25***	*.26**	**.20***	.14	**.38***	**.39***
DA: log(HH inc)	***.27***	.26	.24	.27	***.52***	***.48***	.30	.27	*.58**	*.58**	.21	.21	.24	.060	***.35***	.74	.075	.14
CT: log(HH inc)	−**.45***	−**.51***	−***.74***	−**.85***	−.47	−.44	−**.77***	−**.82***	−.26	−.43	−.14	.059	−**.57***	−**.64***	−**.64***	−**.64***	−.24	−.31
CSD: log(HH inc)	−*.42**	−***.35***	−.63	−.61	−*.85**	−.79	−.60	−***.59***	−*1.15**	−.97	−.37	−***.75***	−.35	−***.38***	−*.48**	−***.53***	.18	.20
CMA: log(HH inc)	−.33	−.33	.19	.19	.25	.25	−.068	−.068	.97	.97	−.17	−.17	−.35	−.35	−.033	−.033	−*1.36**	−*1.36**
$\sum\beta_{\text{inc}}$	−.81		−.67		−.18		−.85		.51		−.13		−*1.12**		−.57		−1.33	
health	**2.61***		**2.80***		**2.70***		**2.83***		**2.57***		**2.71***		**2.78***		**2.82***		**2.65***	
trust-N	**1.24***		**1.10***		**.87***		**1.07***		**.88***		**1.57***		**1.28***		**1.53***		**1.00***	
trust-G	.013		.051		−.014		.041		−.027		.027		.017		.008		.009	
CT: Fraction: immigrants	.11		.038		.42		.16		.34		.20		.036		.086		.24	
CSD: Fraction: immigrants	−**.78***		−*.74**		−.67		−***.74***		−.81		−.48		−**.83***		−**.72***		−***.85***	
CMA: log(pop)	**.099***	**.099***	.079	.079	−.12	−.12	.079	.079	−*.18**	−*.18**	.009	.009	**.13***	**.13***	***.11***	***.11***	.077	.077
CT: $\log(\rho_{\text{pop}})$	−.025	−*.056**	−.052	−*.11**	−.089	−***.13***	−*.071**	−***.13***	−.055	−.063	−.001	.021	−.024	−.036	−.021	−.050	.006	−.063
CSD: trust-N	−**.82***		−.20		−***1.67***		−.36		−***1.83***		−.50		−***.83***		−*.65**		−***1.62***	
controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
geo fixed effects		✓		✓		✓		✓		✓		✓		✓		✓		✓
$\tau_{\text{neigh}} \geq 10$yr			✓	✓	×	×												
$\tau_{\text{city}} \geq 10$yr							✓	✓	×	×								
foreign born											✓	✓	×	×				
own house															✓	✓	×	×
survey	⟨3⟩	⟨2⟩	G17	G17	G17	G17	G17	G17	G17	G17	⟨2⟩	⟨0⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨1⟩	⟨2⟩	⟨0⟩
obs.	36897	≥9620	7086	≥0	5113	≥0	9000	≥0	3199	≥0	9316	≥0	26468	≥3361	25861	≥4424	9923	≥0
pseudo-R^2			.068		.063		.068		.061									

Table 2.8: Summary coefficients with urban life controls. Summary of estimates in the format described on page 14. The standard controls are included but not shown. Significance: **1%*** ***5%*** *10%**

This is remarkable considering that current income might be considered an increasingly poor measure of investments, wealth, or consumption in one's later career. Nevertheless, until the age of retirement, the effects of the regional average income in framing individuals' assessment of their own household income become, if anything, stronger. Not surprisingly, after the age of retirement income loses much of the meaning it holds amongst the working age. Neither household income nor regional incomes maintain any significant partial correlation with SWL.

Table 2.9 separates the sample by sex and by marriage status. Again, there are no significant differences between the groups. These results do not support the signalling interpretation of the estimated pattern of relative income effects.

2.3.8 Symmetry of income effects

Another way to subdivide the sample is in accordance with income itself. Other studies have reached different conclusions on the question of whether the relatively poor or the relatively rich are especially influenced by the comparison of incomes. One might expect the affluent to be more interested in relative status [*Veblen*, 1899]. Conversely, one might expect the below-mean group to be more affected if emulation behaviour is more influenced by upward than downward comparisons, in accordance with the idea of "loss aversion". *Luttmer* [2005] finds no asymmetry in the effect of neighbours' income between those above and below the median income. *Ferrer-i Carbonell* [2005a] reports mixed results, with West Germans showing an asymmetric and upwards comparison effect but East Germans showing symmetric reference behaviour. *McBride* [2001] reports the opposite — a significantly stronger influence of the comparison group, and correspondingly weaker influence of own income, for high-income respondents in the 1994 USA General Social Survey. Similarly, *Kingdon and Knight* [2007] find in South Africa that relative income is more important at higher levels of absolute income.

We look for deviations from our linear specification by modifying equation (2.1) to allow separate coefficients Δ_r^+, Δ_r^- in each region r for those respondents above ($\mathbb{1}_{ir}^+ = 1$; $\mathbb{1}_{ir}^- = 0$) and below ($\mathbb{1}_{ir}^+ = 0$; $\mathbb{1}_{ir}^- = 1$) the reference level:

$$\log\left(\frac{\text{Prob}\,(\text{SWL}_i > j)}{\text{Prob}\,(\text{SWL}_i \leq j)}\right) = c_j + \alpha \cdot \mathbf{X}_i + \tilde{\beta} \cdot \mathbf{Y}_i + \varepsilon_i + \sum_r \left(\left[\mathbb{1}_{ir}^+ \Delta_r^+ + \mathbb{1}_{ir}^- \Delta_r^-\right] \cdot \left[\mathbf{Y}_i - \bar{\mathbf{Y}}_{R_{ir}}\right] + \gamma_r \cdot \mathbf{Z}_{R_{ir}} + \nu_{R_{ir}}\right) \qquad (2.4)$$

The results in Table 2.10 corroborate the findings of *Luttmer* [2005] by showing an absence of any asymmetry in coefficients between those individuals who are above and below the average at each geographical scale. This pattern is revealed for each of the observed values of income, house value, and house size. This symmetry seems somewhat surprisingly close, but considering that explanations are given above for either of the other alternatives, we may say without identifying the psychological channels more explicitly that our observations might

	males	females	single	married
log(own inc)	.05	−***.07***	−.008	−.02
	(.04)	(.03)	(.04)	(.03)
log(HH inc)	**.20***	**.29***	**.19***	**.23***
	(.05)	(.05)	(.05)	(.06)
DA: log(HH inc)	−.07	−.05	−***.17***	−.09
	(.07)	(.07)	(.08)	(.07)
CMA: log(HH inc)	−**.72***	−**.75***	−**.69***	−**.60***
	(.16)	(.16)	(.19)	(.15)
trust-N	**1.32***	**1.12***	**1.13***	**1.38***
	(.12)	(.11)	(.13)	(.11)
health	**2.08***	**1.97***	**2.11***	**1.92***
	(.13)	(.12)	(.14)	(.12)
mastery	**1.93***	**2.23***	**2.56***	**1.73***
	(.18)	(.19)	(.21)	(.19)
godImportance	**.44***	**.51***	**.40***	**.56***
	(.07)	(.08)	(.09)	(.07)
male			−.07	−**.15***
			(.06)	(.05)
age	−**.07***	−**.06***	−**.10***	−**.08***
	(.01)	(.009)	(.01)	(.01)
$(\text{age})^2$	**.0008***	**.0007***	**.001***	**.0009***
	(.0001)	(.0000911)	(.0001)	(.0001)
Obs.	10614	10805	7528	11969
pseudo-R^2	.04	.04	.05	.04

Significance: **1%*** ***5%*** *10%**

Table 2.9: Income effects, sex, and marriage. Ordered logit coefficients for SWL using pooled surveys. Our standard demographic controls are included.

	SWL (1)	SWL (2)	SWL (3)
log(HH inc)	$-\mathbf{.37}^{*}$ (.08)		
DA: Δ^{-}log(HH inc)	$\mathbf{.18}^{*}$ (.06)		
DA: Δ^{+}log(HH inc)	.05 (.06)		
CMA: Δ^{-}log(HH inc)	$-\mathbf{.82}^{*}$ (.11)		
CMA: Δ^{+}log(HH inc)	$-\mathbf{.78}^{*}$ (.10)		
log(houseValue)		$-\mathbf{.14}^{*}$ (.05)	
DA: Δ^{-}log(houseValue)		.14 (.10)	
DA: Δ^{+}log(houseValue)		$-.11$ (.13)	
CMA: Δ^{-}log(houseValue)		$-\mathbf{.30}^{*}$ (.11)	
CMA: Δ^{+}log(houseValue)		$-\mathbf{.40}^{*}$ (.11)	
houseRooms			$-.04$ (.06)
DA: Δ^{-}houseRooms			$\mathbf{.09}^{*}$ (.03)
DA: Δ^{+}houseRooms			$\mathbf{.07}^{*}$ (.02)
CMA: Δ^{-}houseRooms			$-\mathit{.13}^{*}$ (.07)
CMA: Δ^{+}houseRooms			$-\mathit{.11}^{*}$ (.06)
own house		$\surd$	$\surd$
Obs.	30115	22936	23184
pseudo-R^2	.005	.002	.001

Significance: **1%*** ***5%*** *10%**

Table 2.10: Symmetry in comparison effect. Unlike in other tables, the coefficients on region averages here are $\Delta_r^{\pm}$ rather than δ_r; see equation (2.4).

represent some zero-sum combination of asymmetric effects.

A slightly different question is whether the Veblen effect is stronger for individuals with higher or lower *absolute* income levels. In order to treat this question, we conducted a semi-parametric regression in which the box kernel ranged over absolute household income.[16] Figure 2.3 shows the results both for a simplified equation containing household and CT mean incomes along with our standard controls and trust in neighbours, and for a more complete specification containing reference income levels for three geographic scales as well as the same controls. In both cases, the coefficient on absolute income increases with income, suggesting an imperfect specification. For the simpler equation the CT-level Veblen coefficient is approximately constant, while the more complete specification contains the suggestion that the CT-level reference effect also increases with increasing income.

2.3.9 Geo-demographic reference groups

Various mechanisms by which geographic proximity might help to determine reference group formation are plausible. For example, people are likely to interact with their close neighbours and community members in a number of contexts, are likely to work alongside and commute past people who live in the same city, and are likely to have grown up or attended high school in the same metropolitan region. Effective reference levels may be set by emulating one's friends or coworkers, by absorbing some standard from the broader anonymous population, or through some other process of social interaction or information dissemination. By using individual-level data from the 2001 census, we are able to construct some mean incomes for simple, identifiable sub-samples of the population in each census region. Table 2.11 contains a summary of the findings when local members of one's age group or local members of one's visible minority group are used as a reference set. Age categories are 15-19, 20–24, 25–29, 30–39, 40–49, 50–64, and 65+ years, while the "visible minority" designations are those defined by Statistics Canada: Chinese, South Asian, Black, Filipino, Latin American, Southeast Asian, Arab, West Asian, Korean, Japanese, and Other visible minorities. Only respondents who fall into one of the respective categories are included in the "Age Group" and "Visible minority" estimates. We find a significantly reduced CMA income comparison effect when one's own age group is used as a reference, but at least as strong an effect at the CT level. Suggestive of a similar but weaker finding to that of *Kingdon and Knight* [2007] for the "divided society" of South Africa is our finding of an absense of a comparison effect at the CT level combined with a stronger one at the scale of CSDs when visible minority groups are the candidate reference group.[17] However, our sample sizes becomes small when restricted to these categories. Further

[16] An ordered logit model was estimated separately for numerous subsamples, each subsample corresponding to respondents with incomes in a particular range, noted in Figure 2.3 as the kernel width. Using smaller kernel widths resulted in noisier but consistent patterns.

[17] Because "visible minority" status is only available in one survey, a proper comparison of coefficients considers only the EDS results for the "All" and "Visible minority" cases detailed in Table A.4 on page 114.

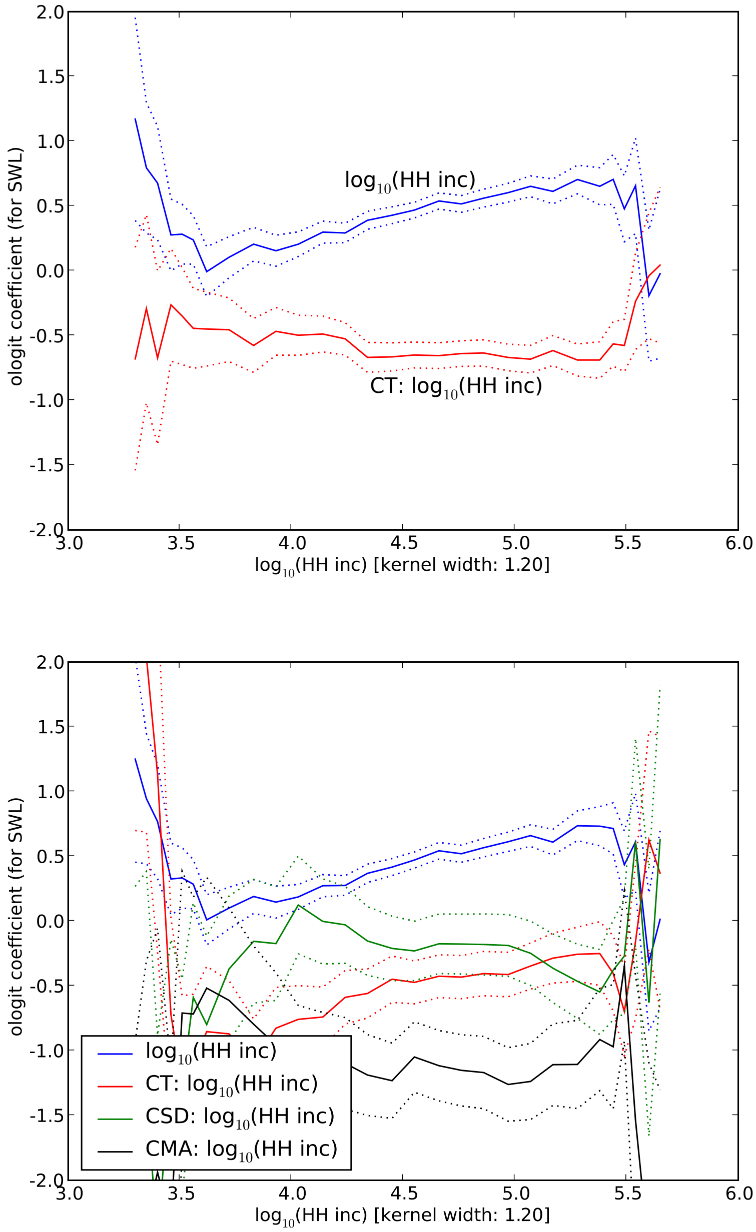


Figure 2.3: Veblen coefficients as a function of income. Dotted lines show the range of one standard error.

investigation along these lines appears to be warranted. For instance, *Charles et al.* [2007] report a stronger conspicuous consumption behaviour for certain visible minority or ethnic groups.

2.3.10 Further robustness checks

Table 2.12 contains a summary of some further checks of the robustness of our estimates. Using OLS or ordered probit in place of ordered logit gives comparable raw coefficients (with the standard factor between probit and logit). Eliminating respondents who reported the highest possible score for SWL does change the picture slightly but leaves unchanged, in particular, the strong negative consumption externality at the CT level.

2.3.11 Absolute and relative benefits of health

For informing policy, empirical well-being research might have little to say if it was found that all determinants of SWL contributed only relatively through context-dependent reference levels. *Alpizar et al.* [2005] posed hypothetical choices to students in order to assess the positional and relative benefits of different kinds of goods. They found that utility from most goods derives from both absolute and relative consumption, although certain goods such as leisure and insurance provide more absolute benefits than housing and income.

We test this proposition for one important determinant of SWL. Table 2.13 shows that, when incomes are controlled for, regional averages of others' health have only positive or insignificant effects on individual SWL. According to this estimate, it may be extrapolated that improving everyone's health at once would make a large positive increase to SWL. Larger scale studies based on cross-country regressions provide further support for the claim that certain non-pecuniary but more objectively measured community attributes, for instance those relating to social capital, are highly valuable social aims from a well-being perspective.

2.4 Discussion

In investigating the effects of geographic income comparison groups, the focus of our analysis has been on *ex post* welfare, as measured by SWL. Before drawing any practical implications from our findings, we point out several complications in interpreting this subject.

Welfarism and relativism

It is common to recognise three ways in which one's own outcomes are put into perspective in the subjective assessment of satisfaction. These correspond to memories of one's past (hence accomodation to *status quo*), comparison to contemporary norms, and reference to aspirations. By focusing on a cross-section in time and by looking at different potential geographic comparison groups, we are probably learning most about the geographic structure of contemporary

	All	All	All	All		Age groups	Age groups	Age groups	Age groups		Vismin groups	Vismin groups	Vismin groups	Vismin groups	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
log(HH inc)	**.28***	**.27***	**.29***	**.37***	**.37***	**.27***	**.29***	**.32***	**.37***	**.37***	.069	.091	.075		.075
	(.036)	(.032)	(.037)	(.087)	(.087)	(.036)	(.034)	(.039)	(.087)	(.087)	(.10)	(.061)	(.10)		(.10)
CT: log(HH inc)	***−.29***	**−.32***	**−.36***		**−.36***	*−.22**	*−.20**	**−.58***		**−.58***	.35	.33	.082		.082
	(.13)	(.097)	(.091)		(.091)	(.12)	(.11)	(.088)		(.088)	(.32)	(.25)	(.29)		(.29)
CSD: log(HH inc)	−.21	−.16			−.16	−.11	−.17			−.17	−.28	***−1.25***			***−1.25***
	(.28)	(.26)			(.26)	(.26)	(.11)			(.11)	(.91)	(.52)			(.52)
CMA: log(HH inc)	**−1.67***				**−1.67***	−.38				−.38	−1.40				−1.40
	(.39)				(.39)	(.26)				(.26)	(1.19)				(1.19)
$\sum \beta_{\text{inc}}$	**−1.88***	.19	−.004	**.37***		***−.97***	.058	.11	**.37***		−1.26	−.83	.16		
	(.40)	(.24)	(.047)	(.087)		(.48)	(.042)	(.069)	(.087)		(.98)	(.79)	(100.0)	(0)	
health	**2.63***	**2.60***	**2.69***	**3.26***		**2.64***	**2.61***	**2.70***	**3.26***						
	(.091)	(.099)	(.11)	(.39)		(.091)	(.10)	(.11)	(.39)						
trust-N	**1.24***	**1.24***	**1.26***	**1.73***		**1.23***	**1.22***	**1.27***	**1.73***		**1.57***	**1.59***	**1.57***		
	(.063)	(.064)	(.058)	(.15)		(.063)	(.064)	(.059)	(.15)		(.22)	(.22)	(.24)		
trust-G	.007	.024	.033	.049		.002	.026	.035	.049		.092	**.10***	*.12**		
	(.029)	(.020)	(.026)	(.071)		(.029)	(.021)	(.026)	(.071)		(.088)	(.036)	(.073)		
controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
geo fixed effects		CMA	CSD	CT	✓		CMA	CSD	CT	✓		CMA	CSD	CT	✓
survey	⟨3⟩	⟨3⟩	⟨3⟩	⟨2⟩	⟨2⟩	⟨3⟩	⟨3⟩	⟨3⟩	⟨2⟩	⟨2⟩	ED	ED	ED	ED	ED
obs.	37701	37601	35937	9851	≥9851	37647	37547	35896	9851	≥9851	4581	4541	4425	0	≥0
pseudo-R^2											.057	.061	.066		
N_{clusters}												18	50		

Table 2.11: Demographic / geographic subpopulations as reference groups. Columns labeled "All" show ordered logit estimates for all respondents using overall means as reference levels. "Age Group" estimates use mean incomes from respondents' own age group as reference levels. "Vismin" estimates include only visible minority respondents and their own-group's mean incomes. Columns marked with a ✓for "geo fixed effects" show summary coefficients of the kind described on page 14. Other columns represent coefficients averaged over surveys, as in the shaded columns of Table 2.2 on page 13. Detailed estimates summarised in this table are found in Table A.4 on page 114. Significance: **1%*** ***5%*** *10%**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(HH inc)	**.27***	**.35***	**.33***	**.36***	**.17***	**.21***	**.52***	**.92***
	(.037)	(.090)	(.035)	(.080)	(.021)	(.048)	(.052)	(.24)
DA: log(HH inc)	***.28***	.27	**.31***	.31	***.17***	.15	**.45***	−.085
	(.13)	(.29)	(.11)	(.26)	(.074)	(.16)	(.16)	(.75)
CT: log(HH inc)	−**.43***	−**.51***	−**.48***	−**.53***	−**.28***	−**.32***	−.32	−**.47***
	(.16)	(.12)	(.14)	(.10)	(.091)	(.070)	(.20)	(.17)
CSD: log(HH inc)	−.24	−.12	−.14	−.050	−.10	−.043	−.33	−***.43***
	(.22)	(.16)	(.18)	(.12)	(.12)	(.095)	(.27)	(.17)
CMA: log(HH inc)	−**1.08***	−**1.08***	−**.91***	−**.91***	−**.64***	−**.64***	−***.72***	−***.72***
	(.26)	(.26)	(.22)	(.22)	(.15)	(.15)	(.33)	(.33)
$\sum \beta_{\text{inc}}$	−**1.19***		−**.92***		−**.69***		***.53***	
	(.28)		(.29)		(.16)		(.27)	
health	**2.61***		**2.03***		**1.44***		**2.32***	
	(.091)		(.074)		(.052)		(.099)	
trust-N	**1.25***		**.99***		**.70***		**1.15***	
	(.061)		(.052)		(.034)		(.071)	
controls	✓	✓	✓	✓	✓	✓	✓	✓
geo fixed effects		✓		✓		✓		✓
survey	⟨3⟩	⟨2⟩	⟨3⟩	⟨2⟩	⟨3⟩	⟨2⟩	⟨3⟩	⟨2⟩
SWL≠10							✓	✓
ologit	✓	✓					✓	✓
OLS			✓	✓				
oprobit					✓	✓		
obs.	36931	≥9620	36931	≥9620	36931	≥9620	24893	≥1969

Table 2.12: Robustness checks for estimates of SWL. Summary of estimates in the format described on page 14. Significance: **1%*** ***5%*** *10%**

	(1)	(2)	(3)	(4)	(5)
log(HH inc)	**.35***	**.34***	**.34***	**.65***	**.65***
	(.057)	(.056)	(.061)	(.20)	(.20)
DA: log(HH inc)	**.49***	**.50***	**.56***	.16	.16
	(.16)	(.15)	(.17)	(.52)	(.52)
CT: log(HH inc)	*−.37**	***−.41***	**−.46***		**−.46***
	(.20)	(.17)	(.16)		(.16)
CSD: log(HH inc)	−.41	−.29			−.29
	(.30)	(.25)			(.25)
CMA: log(HH inc)	**−1.24***				**−1.24***
	(.36)				(.36)
health	**2.72***	**2.71***	**2.78***	**3.39***	**3.39***
	(.089)	(.088)	(.10)	(.35)	(.35)
CT: health	.24	.24	.14		.14
	(.15)	(.15)	(.17)		(.17)
CSD: health	−.16	−.094			−.094
	(.49)	(.51)			(.51)
CMA: health	*1.18**				*1.18**
	(.69)				(.69)
controls	✓	✓	✓	✓	✓
Geo dummies		CMA	CSD	CT	✓
survey	⟨2⟩	⟨2⟩	⟨2⟩	⟨1⟩	⟨1⟩
obs.	13695	13588	12596	1474	≥1474
pseudo-R^2					

Table 2.13: Spillover effects of others' health. The first four columns represent coefficients averaged over surveys, as in the shaded columns of Table 2.2 on page 13. The fifth column shows summary coefficients of the kind described on page 14. Detailed estimates summarised in this table are found in Table A.6 on page 122. Significance: **1%*** ***5%*** *10%**

norms. Aspirations might be thought of as a calculation of what is reasonably attainable; like the others, this is a standard which affects our satisfaction [*Stutzer*, 2004]. Aspirations may be determined in part by the other two influences (comparison with one's past and with one's society's outcomes) but other structural factors such as personal and institutional constraints will also affect how these aspirations are cognitively formed.

It may be noted that all three contextual effects follow from the evolutionary arguments of *Rayo and Becker* [2004] and that all result in mean reference levels rising or falling in tandem with consumption levels over time. Thus any of the three can account for the observation that among many nations, average SWL does not grow with national income.

When considered separately, however, these comparison channels may lead to different policy considerations. For instance, aspirations can be expanded upwards for the majority through a relaxation of class constraints; indeed, the formation of a middle class and an increase in social mobility may be a major driver of economic growth through this channel of aspirations [*Galbraith*, 1979]. Well-being effects of reference group emulation can be minimised by decreasing disparities, and evidence of strong adaptation to income levels over time indicates a SWL value of economic growth and of increasing compensation rates as a function of age, though not necessarily beyond those which reflect increasing productivity due to experience.

Some significant warnings have been articulated which lie in the way of such conclusions, especially as they relate to the measurement and alleviation of poverty. *Galbraith* [1979], in his discussion of the impact of economic aspirations, suggests that people adapt to rates of economic growth just as they do to levels of income, and *Sen* [1983] in his description of the "capabilities approach" warns against absurd prescriptions which may result from an entirely relativist view of welfare. *Sen* [1999] has further warned against the metric of utilities, or "welfarism", because it may lead to the implication that limiting people's knowledge or aspirations is good social policy. Nevertheless, and especially in a relatively open and democratic society, SWL meets Sen's own criterion of measuring people's ability to do and to be what they value. *Kingdon and Knight* [2003] and *Helliwell* [2008] argue that SWL may be an excellent candidate for an encompassing welfare measure even for developing economies.

Endogenous choice of comparison groups and maximisation of SWL

> Hardly any choices are as interactive and interdependent as the choice of whom to associate with, live with, work with, or play with ... [*Schelling*, 2006, p. 43]

If people are sophisticated in their selection of where to live and with whom to socialise, they will take into account any repercussions that set standards for their own future emulation. This remains a difficult complication to the normative assessment of reference level effects, yet it is mitigated by our use of controls, including the "mastery" measure, and in part by our finding, evidenced by the coefficient on CMA incomes, that a dominant comparison group is broadly distributed across metropolitan regions. The latter fact means that household relocations within an urban region are less likely to change Canadians' contemporary reference standards. On the other hand, endogenous choice between different metropolitan areas is poorly accounted for in our work, as is the selection of non-geographic social groups. In addition, while mobility between CMAs is quite limited, selection of one's residential CT is much more common. If this decision is made with the milieu of affluence in mind as an influence on one's own consumption standards, it ought, however, to work against our results, attenuating the negative coefficient on others' CT income. *Knight and Song* [2006] explicitly asked respondents about their income reference group. They find that individuals who are the least content are those with the geographically broadest reference group.[18]

Falk and Knell [2004] analyse competing effects in comparison group selection and the formation of aspirations when there are both relative and absolute returns to well-being. They predict a positive correlation between ability and endogenous standards. On the other hand, there is also strong evidence that people do not fully realise that reference standards will change and therefore that some superficially attractive choices will not end up being beneficial [*Loewenstein et al.*, 2003]. The inclusion in surveys of explicit questions concerning subjective reference groups, such as took place in Wave 3 (2006) of the European Social Survey and in the work of *Knight and Song* [2006], is therefore a valuable innovation.

There is, furthermore, evidence of systematic deviations from optimisation of SWL. The question of what contributes to SWL as a welfare measure (utility in Jeremy Bentham's sense) is quite distinct from the question of whether SWL is a good approximation for utility as in choice theory. *Wilson and Gilbert* [2005] discuss humans' limited ability and systematic inability to forecast their own affect. *Dunn et al.* [2003] address specifically the issue of residence location choice; they use a natural experiment of undergraduate housing assignment as evidence of systematic misprediction of the determinants of one's own SWL.

Thus, the emulation of neighbours or social peers as a behaviour needs to be assessed independently from the reference setting that plays a role in SWL assessment. Nevertheless, our results represent a clearly significant effect of *ex post* neighbours' income.

2.5 Conclusion

We have attempted both to identify the geographic scales which best describe income comparison groups in Canadian cities and, to some degree, to separate income comparison effects from social benefits such as are exhibited by a feeling of trust. Our finding that income comparison, or emulation, effects dominate empathetic ones at levels of metropolitan regions and census tracts is not inconsistent with the findings of *Kingdon and Knight* [2007] for South Africa. They report negative spillovers of income at the district level (with mean populations of 125,000) but positive spillovers within smaller clusters (with mean population 2,900). Our evidence for an empathetic pattern of income spillover effects on the most local scale is weaker than theirs,

[18] Their work is reported as preliminary.

although we find that trust in neighbours has spillover effects on an even smaller scale than *Kingdon and Knight* [2007] can resolve, as well as at larger scales beyond the neighbourhood. We find consistently weaker or nonexistent net effects of others' income at the CSD, or municipal, scale, which is suggestive that tax-funded public goods are an important component of the actual consumption which we would ideally have used in place of our measure of income.

Because of the limited number and variability in CMAs that are intrinsic to Canadian data, our conclusions regarding CMA level effects must remain quite tentative. They nevertheless reflect a strong and important negative association between mean CMA income and mean CMA life satisfaction. It may be that inhabitants of cosmopolitan cities, even in developing nations, form their reference groups in a different manner than do rural dwellers. Our findings do not explain this process and suggest either (1) that comparison groups might consist of more individually specific socially connected networks which tend to be dispersed throughout a broad geographical region or (2) that within metropolitan regions there is high accessibility of information about others' living standards or, at least, wages.

On the other hand income externalities at the census tract level appear to be strong and robust. It may be thought that if urban regions are sites of particularly intense competition over consumption or income status, then past and ongoing urbanisation may have an important effect on production and consumption growth for reasons other than efficiency of production due to agglomeration. However, we do not find evidence of an upward bias to the reference setting. As discussed above, others' results on this question vary. *Eaton and Eswaran* [2006] show that even if reference-setting behaviour is mean-reverting emulation rather than a more one-sided high status seeking, then this aspect of preferences can drive needless and welfare-reducing economic growth.

If the results we find for income relativities should withstand further tests and appear robustly in subsequent surveys, the negative sum of the coefficients on own and comparator incomes suggests the existence of strongly negative consumption externalities. Moreover, these results ignore the negative intergenerational environmental externalities that result from rising global levels of material consumption. Further research is needed to unravel the roles that advertising and other forces play in setting standards for emulation. It has been suggested, for example by *Bertrand et al.* [2006], that the aggregate negative externalities are made larger by a preponderance of advertising and other information flows advocating higher levels of material consumption relative to activities with positive externalities. A better understanding of how norms are established could help to permit individuals to increase their own SWL while not damaging that of their neighbours or of subsequent generations.

Bibliography for Chapter 2

Alpizar, F., F. Carlsson, and O. Johansson-Stenman, How much do we care about absolute versus relative income and consumption?, *Journal of Economic Behavior and Organization*, *56*, 405–21, 2005.

Barrington-Leigh, C. P., Veblen goods and neighbourhoods: endogenising consumption reference groups, 2008a.

Barrington-Leigh, C. P., Weather as a transient influence on survey-reported satisfaction with life, 2008b.

Bertrand, M., S. Mullainathan, and E. Shafir, Behavioral Economics and Marketing in Aid of Decision-Making among the Poor, *Journal of Public Policy & Marketing*, *25*, 8–23, 2006.

Charles, K., E. Hurst, and N. Roussanov, Conspicuous Consumption and Race, *NBER Working Paper*, 2007.

Clark, A., Are wages habit-forming? evidence from micro data, *Journal of Economic Behavior and Organization*, *39*, 179–200, 1999.

Clark, A. E., P. Frijters, and M. A. Shields, Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles, *The Journal of Economic Literature*, *46*, 95–144, 2008, doi:10.1257/jel.46.1.95.

Cole, H., G. Mailath, and A. Postlewaite, Social Norms, Savings Behavior, and Growth, *Journal of Political Economy*, *100*, 1092, 1992.

Corneo, G., and O. Jeanne, Social Organization in an Endogenous Growth Model, *International Economic Review*, *40*, 711–726, 1999.

Diener, E., Subjective well-being, *Psychological Bulletin*, *95*, 542–575, 1984.

Duesenberry, J., *Income, Saving, and the Theory of Consumer Behavior*, Harvard University Press, 1949.

Duncan, O., Does money buy satisfaction?, *Social Indicators Research*, *2*, 267–274, 1975.

Dunn, E., T. Wilson, and D. Gilbert, Location, Location, Location: The Misprediction of Satisfaction in Housing Lotteries, *Personality and Social Psychology Bulletin*, *29*, 1421–1432, 2003.

Easterlin, R., Does Economic Growth Improve the Human Lot? Some Empirical Evidence, *Nations and Households in Economic Growth: Essays in Honour of Moses Abramovitz*, pp. 98–125, 1974.

Eaton, B. C., and M. Eswaran, Well-Being and Affluence in the Presence of a Veblen Good, 2006.

Eaton, C., and M. Eswaran, The evolution of preferences and competition: a rationalization of Veblen's theory of invidious comparisons, *Canadian Journal of Economics*, *36*, 832–859, 2003, available at http://ideas.repec.org/a/cje/issued/v36y2003i4p832-859.html.

Falk, A., and M. Knell, Choosing the Joneses: Endogenous Goals and Reference Standards, *Scandinavian Journal of Economics*, *106*, 417–435, 2004.

Ferrer-i Carbonell, A., Income and well-being: an empirical analysis of the comparison income effect, *Journal of Public Economics*, *89*, 997–1019, 2005a.

Ferrer-i Carbonell, A., and P. Frijters, How Important is Methodology for the estimates of the determinants of Happiness?, *The Economic Journal*, *114*, 641–659, 2004.

Fliessbach, K., B. Weber, P. Trautner, T. Dohmen, U. Sunde, C. E. Elger, and A. Falk, Social Comparison Affects Reward-Related Brain Activity in the Human Ventral Striatum, *Science*, *318*, 1305–1308, 2007.

Frey, B., and A. Stutzer, What Can Economists Learn from Happiness Research?, *Journal of Economic Literature*, *40*, 402–435, 2002.

Galbraith, J., *The Affluent Society*, Houghton Mifflin Books, 1958.

Galbraith, J., *The nature of mass poverty*, Harvard University Press, 1979.

Helliwell, J., How's life? Combining individual and national variables to explain subjective well-being, *Economic Modelling*, *20*, 331–360, 2003.

Helliwell, J., Life Satisfaction and Quality of Development, *NBER Working Paper*, 2008.

Helliwell, J., and H. Huang, How's the Job? Well-Being and Social Capital in the Workplace, *Forthcoming in Industrial and Labor Relations Review*, 2005.

Helliwell, J., and R. Putnam, Education and Social Capital, *Eastern Economic Journal*, *33*, 1–19, 2007.

Kingdon, G., and J. Knight, Well-being poverty versus income poverty and capabilities poverty?, *Center for the Study of African Economies: University of Oxford, manuscript*, 2003.

Kingdon, G., and J. Knight, Community, Comparisons and Subjective Well-being in a Divided Society, *Journal of Economic Behaviour and Organization*, 2007.

Knight, J., and L. Song, Subjective well-being and its determinants in rural China, *University of Nottingham, mimeo*, 2006.

Krueger, A. B., and D. A. Schkade, The reliability of subjective well-being measures, *Journal of Public Economics*, *92*, 1833–1845, 2008.

Loewenstein, G., T. O'Donoghue, and M. Rabin, Projection bias in predicting future utility, *The Quarterly Journal of Economics*, *118*, 1209–1248, 2003.

Luttmer, E. F. P., Neighbors as Negatives: Relative Earnings and Well-Being, *Quarterly Journal of Economics*, *120*, 963–1002, 2005.

Marx, K., and F. Engels, Marx Engels selected works, Progress Publishers (1986), 1848.

McBride, M., Relative-income effects on subjective well-being in the cross-section, *Journal of Economic Behavior and Organization*, *45*, 251–278, 2001.

Oreopoulos, P., Do Neighbourhoods Influence Long-term Labour Market Success? A Comparison of Adults Who Grew up in Different Public Housing Projects, *Analytical Studies Branch Research Paper Series 2002185e*, Statistics Canada, Analytical Studies Branch, 2002.

Pigou, A., The Economic of Welfare, *London: Macmillam*, 1920.

Rayo, L., and G. Becker, Evolutionary Efficiency and Mean Reversion in Happiness, *mimeographed, University of Chicago*, 2004.

Rege, M., Why do people care about social status?, *Journal of Economic Behavior & Organization,*, *66*, 233–242, 2008.

Schelling, T., *Micromotives and macrobehavior*, W. W. Norton and Company, New York, 2006, originally published in 1978.

Sen, A. K., Poor, relatively speaking, *Oxford Economic Papers*, *35*, 153–169, 1983.

Sen, A. K., *Development as Freedom.*, Oxford University Press, 1999.

Stutzer, A., The Role of Income Aspirations in Individual Happiness, *Journal of Economic Behavior and Organization*, *54*, 89–109, 2004.

Tobler, P. N., C. D. Fiorillo, and W. Schultz, Adaptive Coding of Reward Value by Dopamine Neurons, *Science*, *307*, 1642–1645, 2005.

van de Stadt, H., A. Kapteyn, and S. van de Geer, The Relativity of Utility: Evidence from Panel Data, *The Review of Economics and Statistics*, *67*, 179–187, 1985.

Veblen, T., *The theory of the leisure class*, A.M. Kelley, 1899.

Vigdor, J., Fifty Million Voters Can't Be Wrong: Economic Self-Interest and Redistributive Politics, *NBER Working Paper*, 2006.

Wilson, T., and D. Gilbert, Affective Forecasting, *Current Directions in Psychological Science*, *14*, 131, 2005.

Chapter 3

Veblen goods and neighbourhoods: endogenising consumption reference groups

3.1 Introduction

A number of studies have shown large negative externalities in individual subjective well-being due to neighbours' income [*Luttmer*, 2005; *Kingdon and Knight*, 2007; *Barrington-Leigh and Helliwell*, 2007b].[1] These externalities appear to reflect the role of nearby households as reference groups acting in individuals' reference-dependent preferences over income or consumption. At the same time, there are many reasons to expect positive spillovers from having prosperous neighbours. For instance, the quantity of tax-funded public goods and certain forms of social capital spillovers can be expected to be correlated with the incomes of nearby residents and thus to generate an apparent empathy effect. Alternatively, an idea pursued in this work is that neighbours' income may contribute to a local status level enjoyed by the entire neighbourhood, for instance through conspicuous displays of affluence.

An unresolved question is how such opposing positive and negative externalities of others' income relate to each other. It may, for instance, be that one effect is concentrated on a finer geographic scale than the other. In this work, I consider the possibility that individuals are fully aware of the structure of such returns. The motivating questions are then, firstly: when households properly anticipate the importance of reference groups and have some choice over where they live, can the simultaneous choice of whom to associate with and how much to consume lead to self-organisation of heterogeneous individuals into differentiated groups? Secondly, in a world in which such comparison effects are dominant, will a policy maker wish to curtail production of the status good or the freedom to sort? If relativities in preferences are to be acknowledged seriously in economics, general equilibrium outcomes including endogenous sorting must be understood.

With the aim to put the empirical work on geographic consumption reference groups in a

[1] A version of this chapter will be submitted for publication as Barrington-Leigh, C.P., 'Veblen goods and neighbourhoods: endogenising consumption reference groups.' Thanks to Chris Bidner, Peter Burton, Mukesh Eswaran, Patrick Francois, John Helliwell, and Ken Jackson for helpful discussion.

more explicit framework, I develop a basic model of geographic organisation when such "Veblen" preferences are relevant.[2] This represents an extension of previous work in two regards. In comparison to the symmetric Veblen equilibrium of *Eaton and Eswaran* [2006], I treat cases when (1) consumption is not homogeneous across individuals and (2) consumption reference groups are neither fixed nor common to all individuals. Thus, interdependent preferences drive both the segregation of types into dissimilar reference groups and the individual consumption choices given those reference groups. That is, reference groups are endogenised.[3] In the context pursued below, households choosing a home take into account the neighbourhood, judged in part by the look of other nearby houses. Simultaneously, within those neighbourhoods when building or maintaining their houses, yards, and even amenities like cars, consumers are influenced by the decisions of their neighbours and, in particular, tend to emulate local consumption norms.

I will not abstract from details of the functional dependence of utility on consumption of Veblen goods, since in investigating geographic disparity one must depart from the symmetric consumption equilibria which provide elegant solutions in the analysis of *Eaton and Eswaran* [2006]. In addition, I depart from the representative agent formulation and assume exogenous heterogeneity. However, non-symmetric equilibria do not afford easy discussion of efficiency, since Veblen goods by their nature generate real utility benefits for some individuals at the expense of others.

Geographic proximity is only one of several plausible factors in delineating reference groups. Other natural reference groups include nuclear and extended family, work colleages, ethnic groups, and socioeconomic classes. Moreover, experience from one's own past and aspirations based on cognitive reasoning also provide reference levels which frame consumption evaluation. These contextual effects are all consistent with the evolutionary arguments of *Rayo and Becker* [2004][4]. Nevertheless, a focus on the interaction between interdependent preferences and settlement patterns that are spatially sorted according to income or consumption level is particularly important for its relevance to urban planning, real estate markets, and the empirical analysis of

[2]The name is due to *Veblen* [1899] but I use the term *Veblen good* in the sense that *Eaton and Eswaran* [2006] do. A pure Veblen good is one whose contribution to utility comes in a purely relative way, such that a simultaneous increase of its consumption in a homogeneous population does not add to welfare.

[3]The subjective well-being and social psychology literature indicates that there are likely systematic biases (generally in the direction of materialism) in individual choice, such that contemporary individuals are not acting to maximise their happiness [*Dunn et al.*, 2003; *Loewenstein et al.*, 2003]. However, there is no clear indication that people are confused more specifically about the competitive nature of consumption. In this work I do not assume any naiveté on the part of decision makers. The outcomes are driven by the collective action problem inherent in the consumption externality.

[4]They use a principal-agent framework to address the task of evolutionary forces in designing our internal reward circuitry, subject to the constraints that it has finite bounds. They argue that it therefore must have evolved with features that engineers would call automatic gain control and a (temporal) high-pass filter. That is, the comparison level and scale used for translating one's own consumption level into a psychological reward adapt to make best use of the available range of the reward experience.

geographic reference group effects[5]. The most obvious source of endogeneity for any spatial analysis, such as the empirical work motivating this study, is that people are mobile. Therefore, if reference effects are in play, households may have consciously chosen their reference group by moving to it.

The paper treats two general model formulations. Section 3.2 addresses the first, in which there are exactly two neighbourhood locations and two types of household. This simple case foreshadows most of the main results, but suffers from analytic intractability and assumes away the possibility of a (land) market being involved in the allocation of groups, or locations, to households. In Section 3.3 both the neighbourhood characteristics and the household types are continuously distributed and a land market regulates who lives where. Counterintuitively, this framework turns out to be more amenable to closed-form analysis than the discrete case. Section 3.4 provides some simulations of sample equilibria, and Section 3.5 concludes. A number of issues are addressed in more detail in the Appendix, which also contains proofs to propositions in the main text.

3.2 Discrete types and unpriced land

Consider a discrete set of household types, exogenously differentiated by their endowment of labour productivity $w \in [w_L, w_H]$. Each household chooses a consumption level of a pure Veblen good and also chooses which peer group to join. The sole industry may be taken to be the production of the pure Veblen good, housing, and the reference groups may be thought of as non-interacting neighbourhoods characterised by the average value of housing chosen by their residents. After choosing a residential neighbourhood, households compare their consumption of the Veblen good to average consumption in their own neighbourhood.[6] Nevertheless, agents are sophisticated rather than naïve in that prior to choosing a location, they are fully aware that their future consumption benefit will be framed by the neighbourhood that they have chosen. I will henceforth use the housing and neighbourhood context to describe model economies, although the relevance of the scenario extends to other Veblen goods with endogenous reference groups.

To elucidate the possibility of self-forming groups amongst Veblen consumers who make disaggregated decisons about their reference groups, I start by incorporating into the utility function a benefit of living in a wealthy neighbourhood, to act in tandem with the disutility

[5] Several empricial studies have, for reasons of empirical convenience and availability of data, assessed income reference groups on a geographic basis. See *Barrington-Leigh and Helliwell* [2007b] and *Clark et al.* [2008].

[6] The simplifying assumption that neighbourhoods are non-interacting in this interpretation makes the model and those that follow non-spatial, strictly speaking. That is, there is no sense of physical proximity of one neighbourhood to another.

imposed by having a higher consumption reference group.[7] Let preferences be defined[8] over leisure $x \geq 0$, the conspicuous extravagance $h \geq 0$ of one's house, the average value $\bar{h}$ of houses in one's choice of a neighbourhood, and the global average value of houses $\bar{\bar{h}}$. For convenience, utility is additively separable into a leisure term $F(\cdot)$, a Veblen term $H(\cdot)$ comparing own consumption with that of one's chosen peers, and a further Veblen term $N(\cdot)$ comparing one's neighbourhood to other neighbourhoods:[9]

$$U(x,h,\bar{h}) = \Phi \log(x) - \Lambda \exp\left(-\lambda\left[h-\bar{h}\right]\right) + N \log\left(1+\bar{h}/\bar{\bar{h}}\right) \tag{3.1}$$

Under these preferences, neighbourhood benefits accrue relative to a reference level $\bar{\bar{h}}$, which is the average consumption across neighbourhoods. The undesirable neighbourhood externality, on the other hand, comes about through a more local comparison between the neighbourhood standard $\bar{h}$ and the household's own consumption h. Using this form for $N(\cdot)$ is convenient in part because it allows the consideration below of a planner's policy which eliminates all production of the Veblen good[10] and also provides consistency with Section 3.3, to follow.

In choosing its optimal consumption, a household of type w is constrained by the budget

$$w[1-x] \geq h$$

Thus, given the optimality condition

$$x^\star = 1 - h/w \tag{3.2}$$

the household's decision problem may be reduced to a nested choice of an optimal housing purchase $h^\star(\bar{h})$ for each possible neighbourhood $\bar{h}$, followed by a choice of optimal neighbourhood $\bar{h}^\star$. In contrast to other superficially appealing forms for preferences, detailed in the Appendix, the utility function in equation (3.1) embodies bounded benefits to individual consumption of the Veblen good and a large penalty in utility for consuming much less than one's neighbours. Holding $\bar{h}$ fixed, $U(x^\star(h),h)$ is concave and its global optimum must be consistent with the first order condition

[7] Without any benefits to having wealthy neighbours, there cannot be any differentiation of types. See Appendix Section B.2.3 for a discussion of plausible positive consumption externalities in this geographic context.

[8] This form of utility is convenient in that it admits an equilibrium of the desired kind. See Section B.3 for a discussion of the properties of the logarithm and exponential terms and how they relate to past literature exploring utility functions defined over differences — which may be positive or negative — and ratios of quantities of goods.

[9] Also discussed in the Appendix are models incorporating an absolute utility benefit of wealthy neighbours, rather than the relative one posed here. This distinction is unlikely to be important except in as far as it affects analytic tractability and ease of welfare analysis.

[10] For this case, the limit of $1+\bar{h}/\bar{\bar{h}}$ is taken to be 2.

$$F'(1-\frac{h}{w}) = wH_h(h,\bar{h}) \quad \text{or} \quad h=0 \tag{3.3}$$

An explicit form for the optimal consumption choice $h^\star(w,\bar{h})$ for a household placed in a neighbourhood with average consumption $\bar{h}$ can be written in terms of the principal branch of the Lambert W function:[11]

$$h^\star(w,\bar{h}) = \max\left\{0,\, w - \frac{1}{\lambda}L(w,\bar{h})\right\} \tag{3.4}$$

where

$$L(w,\bar{h}) \equiv \text{LambertW}\left(\frac{\Phi}{\Lambda}e^{\lambda\left[w-\bar{h}\right]}\right)$$

Consumption $h^\star(w,\bar{h})$ is increasing (and leisure is decreasing) in $\bar{h}$: households will consume more when their neighbours do.[12] The corner solution, $h^\star = 0$, occurs where $\bar{h} < \frac{1}{\lambda}\log\left(\frac{\Phi}{\Lambda\lambda w}\right)$. The indirect utility $U(w,\bar{h})$ can now be expressed through substitution of $h^\star(w,\bar{h})$ into equation (3.1). Taking derivatives, this indirect utility is seen to be concave in both the interior and corner regions:

$$\frac{d^2U(w,\bar{h})}{d\bar{h}^2} = \begin{cases} -\frac{\lambda^2\Phi}{\left[L(w,\bar{h})+L(w,\bar{h})^2\right]} - \frac{N}{\left[\bar{h}+\bar{\bar{h}}\right]^2} < 0, & \text{for } \bar{h} > \frac{1}{\lambda}\log\left(\frac{\Phi}{\Lambda\lambda w}\right) \\ -\Lambda\lambda^2 e^{\lambda\bar{h}} - \frac{N}{\left[\bar{h}+\bar{\bar{h}}\right]^2} < 0, & \text{for } \bar{h} < \frac{1}{\lambda}\log\left(\frac{\Phi}{\Lambda\lambda w}\right) \end{cases}$$

Because the first derivative $\frac{dU(w,\bar{h})}{d\bar{h}}$ is continuous through $\bar{h} = \frac{1}{\lambda}\log\left(\frac{\Phi}{\Lambda\lambda w}\right)$, concavity ensures that there is a global maximum. Nevertheless, there is no general analytic form for the optimal $\bar{h}$, were a continuous choice available.

Moreover, households are not able to choose an arbitrary $\bar{h}$. Rather, they must choose between one of the two available neighbourhoods whose consumption levels $\bar{h}$ are equilibrium outcomes. For a separating equilibrium[13] in which $h = \bar{h}$ for each type, the equilibrium neigh-

[11] The Lambert W function, also occasionally called the *omega* function or *product-log*, is the inverse function of $f(Z) = Z\exp(Z)$ [*Corless et al.*, 1996]. Although less well known, it is very analogous to the logarithm. The real-valued principal branch is always implied in this work. $\text{LambertW}(x) > 0$ for $x > 0$. It is increasing, concave, and passes through the origin. Two identities used in this work are:

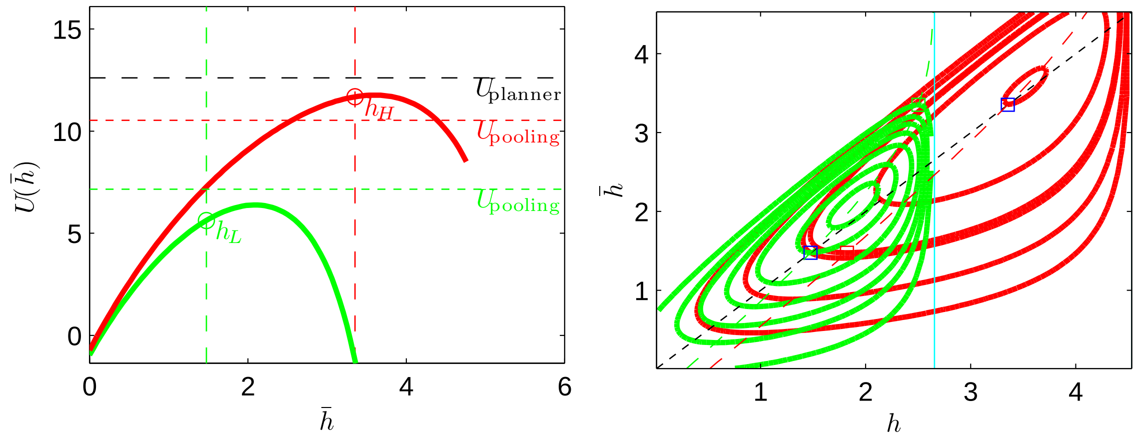
$$\log(\text{LambertW}(Z)) = \log(Z) - \text{LambertW}(Z)$$

and

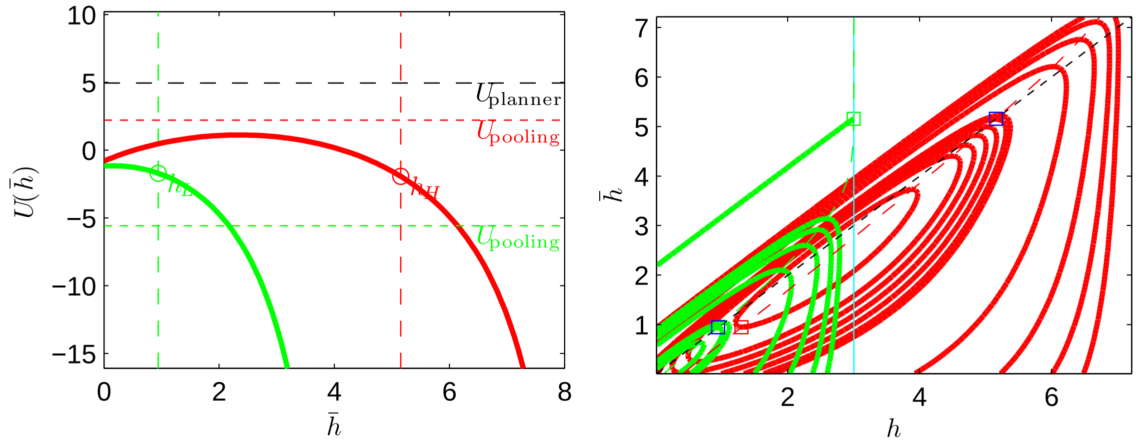
$$\frac{d}{dZ}\text{LambertW}(Z) = \frac{1}{Z}\frac{\text{LambertW}(Z)}{1+\text{LambertW}(Z)}$$

[12] See Equation (B.7) on page 158 for a contrasting case.

[13] A separating equilibrium is one in which neighbourhoods are differentiated according to household type. This



(a) Separating quilibrium for $\Phi \approx 3$, $\Lambda \approx 1$, $\lambda \approx 2$, $N \approx 20$, $w_L \approx 2.7$, and $w_H \approx 4.5$.



(b) No separating equilibrim exists for $\Phi \approx 8$, $\Lambda \approx 1$, $\lambda \approx 3$, $N \approx 9$, $w_L \approx 3$, and $w_H \approx 7$.

Figure 3.1: Contingent existence of separating equilibrium. Separating equilibrium (a) exists for "log-exp-log" preferences given by equation (3.1) but none exists (b) for other parameters in the same functional form. Also shown are utility levels in the pooling equilibrium for each type (U_{pooling}) and under the policy constraint of no Veblen good production (U_{planner}).

bourhoods lie at $\bar{h}_{\text{eq}} = \max\left\{0, w - \frac{\Phi}{\Lambda\lambda}\right\}$. Because for each type w there exists a global optimum $\bar{h} = \bar{h}_{\max U}$, it may be possible for certain fortuitous ranges of parameters to conspire to make $\bar{h}_{\text{eq}} \approx \bar{h}_{\max U}$ for each type. In this case, both types are content in their own neighbourhood and allocations form a separating equilibrium.

Figure 3.1a shows such a situation. By contrast, with different parameter values one or the other of the household types may prefer a deviation from $\bar{h}_{\text{eq}}$, as shown in Figure 3.1b where the high type prefers to move. Marked in the left hand panels of Figure 3.1 are the utility levels for each household type in the alternate, pooling equilibrium, as well as the homogeneous utility level for the case in which Veblen good production is prohibited and leisure is maximised. The pooling outcome is always an equilibrium and in cases such as that of Figure 3.1b it constitutes the unique equilibrium in pure strategies.[14]

For the case shown in Figure 3.1a, the high type is better off in the separating equilibrium, while the low type prefers the pooling equilibrium and could therefore be said to favour policy designed to encourage neighbourhood integration across economic classes. Both types would prefer to have a planner remove the possibility of decentralised decision making about Veblen good production altogether, since the negative externality dominates the benefits even for the high type. This is reminiscent of the findings of *Eaton and Eswaran* [2006].

These qualitative features are not universal, however. In Figure 3.2, panel (a) shows a case when, conversely, the high type rather than the low type prefers an integrated neighbourhood, while in panel (c) both types prefer the pooling equilibrium. Numerous other orderings are possible. Figure 3.3 shows two cases in which the high type prefers to keep Veblen goods in production; that is, the planner's policy of eliminating Veblen goods would not be a Pareto improvement over either unregulated equilibrium. In the second case shown, the high type additionally prefers the integrated neighbourhood with Veblen goods to the one without.

Still other welfare orderings were found for different parameter values. Figure 3.4 shows that different regimes of exogenous parameters result in different welfare implications. Outside the region shown, separating equilibria were not found to exist. The distribution of points shows that endogenous group formation is not possible when within-group comparisons (Λ) receive considerably stronger weight in preferences than the between-group comparisons (N).

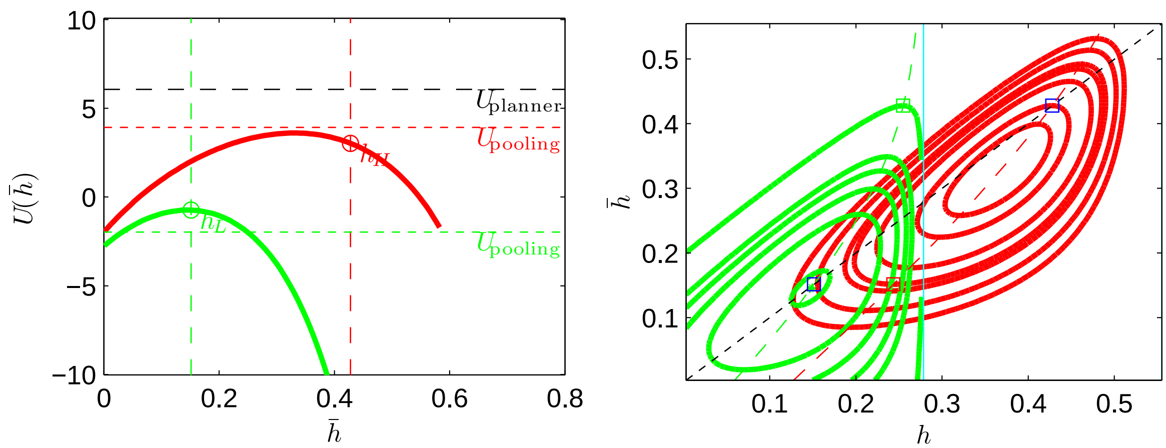
3.2.1 Summary

So far I have analysed the simplest case of a heterogeneous population choosing their own reference groups — the case of two types. Depending on the functional form of the utility, households may prefer to have higher or lower consumption of a Veblen good when they move to a higher consumption neighbourhood.[15] In all cases, there exists a pooling equilibrium con-

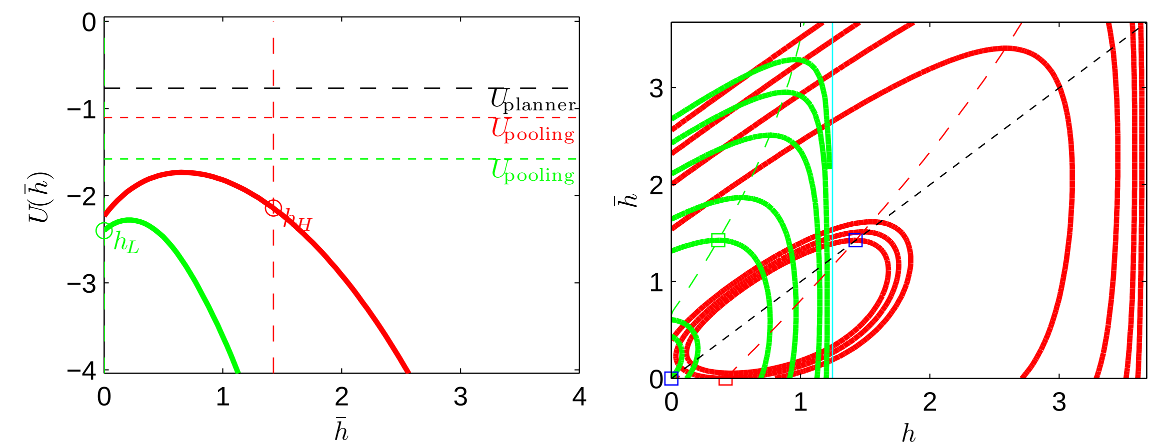
equilibrium is more explicitly defined in the Appendix on page 157. An analogous equilibrium for the continuous case is also defined below in Section 3.3.

[14] See page 160 of the Appendix for a discussion of mixed strategies.

[15] For the latter case, see, for example, Equation (B.7) on page 158.

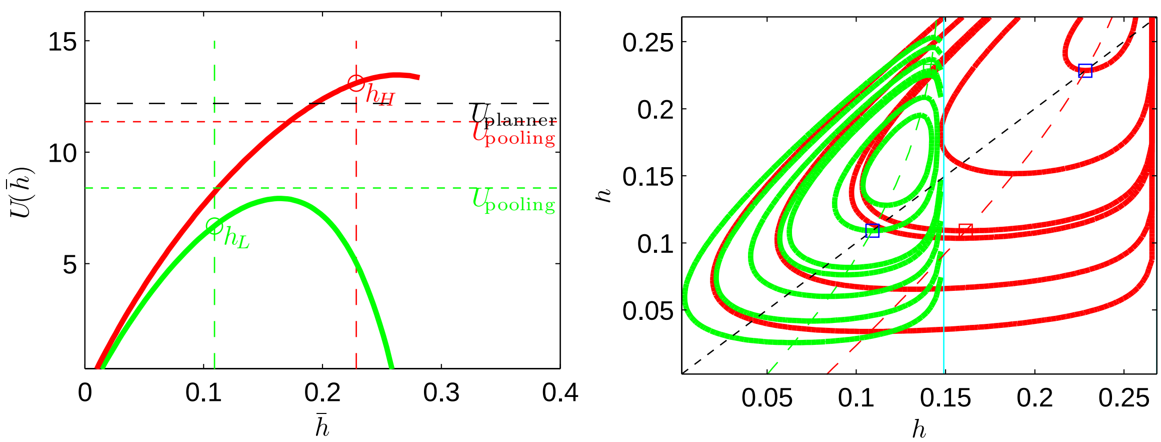


(a) Separating quilibrium for $\Phi \approx 4$, $\Lambda \approx 1$, $\lambda \approx 3$, $N \approx 13$, $w_L \approx 0.3$, and $w_H \approx 0.6$.

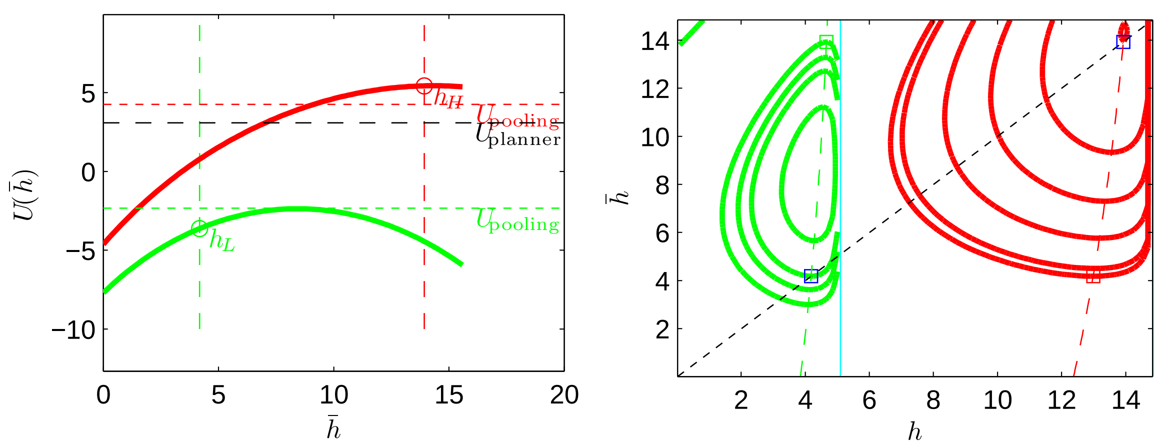


(b) Separating quilibrium for $\Phi \approx 2$, $\Lambda \approx 11$, $\lambda \approx 0.04$, $N \approx 3$, $w_L \approx 6$, and $w_H \approx 12$.

Figure 3.2: Additional cases of equilibrium under "log-exp-log" preferences.



(a) Separating quilibrium for $\Phi \approx 1$, $\Lambda \approx 2$, $\lambda \approx 19$, $N \approx 20$, $w_L \approx 0.15$, and $w_H \approx 0.27$.



(b) Separating quilibrium for $\Phi \approx 0.7$, $\Lambda \approx 9$, $\lambda \approx 0.1$, $N \approx 18$, $w_L \approx 5$, and $w_H \approx 15$.

Figure 3.3: Further cases of equilibrium under "log-exp-log" preferences.

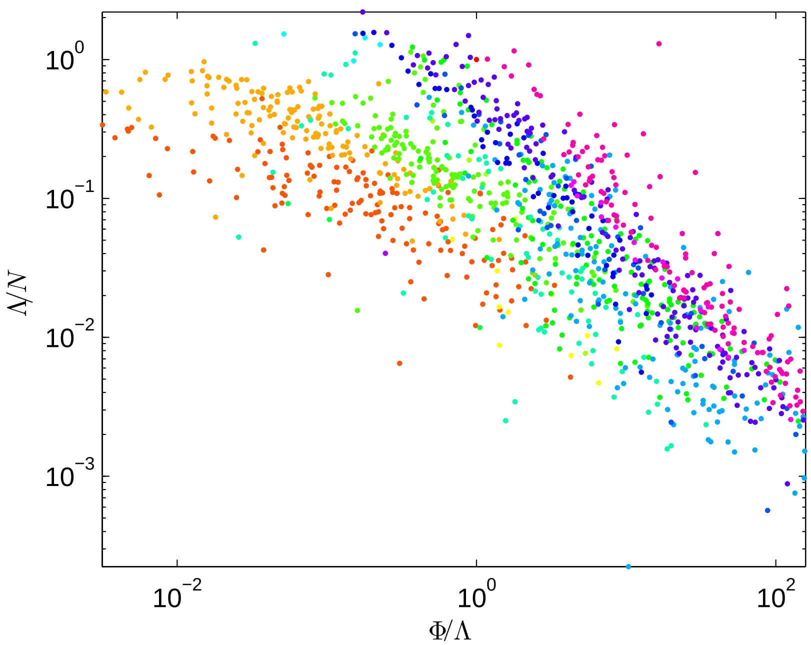


Figure 3.4: Separating equilibrium parameter relationships. Colours indicate different qualitative welfare orderings of pooling, planner, and separating outcomes.

forming to the consistency condition that all households choose each neighbourhood with equal probability. Only for certain cases, on the other hand, does a pure strategy equilibrium exist in which different types prefer to remain segregated in neighbourhoods of internally homogeneous consumption levels. Nevertheless, the discrete nature of the choice amongst neighbourhoods makes it difficult to find closed form solutions or conditions on the existence of such equilibria.

When both pooling and separating equilibria exist, numerical simulation indicates no simple universal welfare implications. Pure Veblen goods may be a desirable feature of the economy for wealthier households, and the freedom to relocate to form one's own reference groups may be desirable for one, both, or neither of the two types. These general features will be recaptured in the more analytical treatment to follow.

One reason for the awkwardness of the household problem and the condition for existence of a separating equilibrium is that there is no price to capture the benefit of a neighbourhood's consumption externalities. A natural way to do this is to allow a price for land, which heretofore has been costless. That is, for the case of a discrete set of neighbourhoods, separating equilibria could more easily be supported if entry to a neighbourhood was competitive and exacted a cost to the household. However, two potential problems present themselves in this regard. First, prices relate to marginal benefits in the real world and are therefore best incorporated into a model with a continuum of neighbourhood consumption levels $\bar{h}$. Secondly, in order to preserve a general equilibrium analysis, revenue from the sale or rental of land must be returned somehow to households.

These two issues are addressed in the following section by extending the endogenous reference group choice set to a continuum and by more realistically pricing land independently from housing.

3.3 A Continuum of types and a market for land

Consider then a framework in which, once again, static consumption reference-setting occurs both within a neighbourhood and between neighbourhoods. In choosing how much to spend on their own dwelling, household make a decision which is framed by the norm in their neighbourhood. In addition, households must choose a neighbourhood in which to position themselves. This affects not only the utility derived from their individual consumption choice but also provides a status payoff since they derive satisfaction from the relative standing of their neighbourhood.[16]

[16] As mentioned previously, there are several possible reasons for neighbourhood status. For the sake of concreteness, I keep as the driver the same conspicuous consumption that drives house choice itself. That is, a neighbourhood's status value is determined by its average level of housing as compared with that of the greater region. This corresponds to the type 3 benefit on page 154. This specification is consistent with the findings of *Barrington-Leigh and Helliwell* [2007b] and provides a coherent interpretation for welfare analysis of the consumption of neighbourhood quality. The drawback of this format is some superficial complexity: the household problem now represents two nested Veblen consumption choices. However, only one incorporates an endogenous choice of reference group,

Therefore, as before, decisionmakers are faced with competing incentives to place themselves in a high or low affluence neighbourhood. In the analysis to follow, however, I introduce an additional direct cost associated with this choice. This comes about by relaxing the assumption of free land. When land is owned and rented, the marginal value to the renter of the reference level embodied by a particular location is captured in the price of land. This market can, as I show below, facilitate a disaggregated choice equilibrium of the kind already treated for discrete types.

In contrast to models such as that of *Rothstein* [2006] in which a small number of school districts confer peer effects to their residents,[17] a reasonable number of consumption reference group choices in the present context is large, since prospective homeowners can typically choose their neighbourhood from a nearly continuous set of affluence levels. Accordingly, I consider the case when there is a continuum of neighbourhoods rather than a discrete set. A crucial feature of the equilibrium to be defined below is that households have the option of moving to a neighbourhood with a marginally greater or lesser average consumption, just as they have the option of marginal changes to the size of their own house. Because households can relocate to their ideal reference group, there is no clustering of different types together in one neighbourhood.

3.3.1 Agents' problem

As before, household agents are exogenously differentiated by their endowed labour productivity $w \in [w_L, w_H]$ in housebuilding, the sole industry. Now, however, types are continuously and uniformly distributed over this range. For each type w, there is a population of measure 1.

Agents maximise the following additively separable utility function through their choice of leisure $0 \leq x \leq 1$; the extravagance $h \geq 0$ of their house, which is the numeraire good; and their choice of a neighbourhood characterised by houses of average value $\bar{h}$:

$$U(x,h,\bar{h}) = F(x) + H(h,\bar{h}) + N(\bar{h},\bar{\bar{h}})$$

The global average level of housing consumption $\bar{\bar{h}}$ is perceived as identical by everyone. Each household is constrained by the budget

$$w[1-x] + r \geq h + p(\bar{h})$$

where r is any land dividend income received, and $p(\bar{h})$ is the competitive price of land in a neighbourhood with mean consumption $\bar{h}$. I will assume that land plots come in parcels that are independent of the size of the house that is built on them, and that this parcel size is uniform across neighbourhoods.

and it is the dynamics of this endogeneity that is the focus of the investigation.

[17] A different Tiebout equilibrium is defined in that case for each exogenously given integer number of discrete districts. In contrast, I consider continua of both household types and neighbourhoods and solve, below, for a unique equilibrium.

3.3.2 Firms' problem

Formally, there are two sectors of competitive firms.

Land management sector

Although the neighbourhood economy considered here is not explicitly spatial in that it abstracts from the arrangement and proximity of different neighbourhoods with respect to each other, the supply side of the land market must nevertheless be modeled in order for land price to be endogenous. Three scenerios present themselves as reasonable model assumptions:

free land: First of all, a simpler case is the one in which land is part of a commons. Then $r = p(\bar{h}) = 0$ and households choose their neighbourhood without any explicit cost, as in the discrete model of Section 3.2. Neighbourhoods are nevertheless mutually segregated.

absentee landowners: In this case, all plots of land are owned by absentee landlords who have no current use for them, and they rent individual plots to the highest bidder. Dividends r are zero for all households. For the purpose of welfare analysis, landowners are considered to be external to the economy.

uniform ownership: This case is similar to that of absentee landowners except that each plot of land is rented by an independent firm whose shares are equally owned by all households.[18] All rental income is profit and is distributed uniformly to shareholders. Thus each household, regardless of type, receives dividends r corresponding to the average price of rented land.

For reasons discussed below, land is assumed in much of what follows to be owned and rented by firms. Each plot of land is owned and managed by a separate price-taking firm whose equity is in turn owned in equal part by all households. Firms have no costs and simply receive rent p from the highest bidder for their land, subject only to the condition of nonnegative profit:

$$p \geq 0 \tag{3.5}$$

Firms then distribute all their profit to their shareholders.

Housebuilding sector

There is also a competitive housing production industry. Agents are endowed with an innate and universally visible productivity. Firms hire workers, pay them according to their productivity, and produce houses (or house maintenance, or conspicuous household consumption goods more generally), making zero profit.

[18] Unequal land ownership may be empirically more appealing and may represent a more acceptable middle ground between the two extremes, but it would constitute a complication at the moment.

3.3.3 Definition of equilibrium

Given a continuous range of types $[w_L, w_H]$, a *separating neighbourhood equilibrium* consists of an average consumption $\bar{h}(n)$ for each neighbourhood n,[19] an overall global average consumption $\bar{\bar{h}}$, market land prices $p(\bar{h})$ in each neighbourhood, rental dividends r, and allocations $\{x(w), h(w), \bar{h}(w)\}$, which

- satisfy consistency and aggregation requirements, in order that the perceived mean $\bar{h}$ is equal to the average consumption in each neighbourhood and that the global mean $\bar{\bar{h}}$ is the average over neighbourhoods,

$$\bar{h} = \int_{\{w|\bar{h}(w)=\bar{h}\}} h(w)dw \quad \forall\bar{h} \tag{3.6}$$

$$\bar{\bar{h}} = \int \bar{h}(w)dw \tag{3.7}$$

- satisfy a non-profit condition on rental income (for the case when dividends are returned to households),

$$r = \int p\left(\bar{h}(w)\right)dw,$$

- satisfy the firms' incentive criterion,

$$p(\bar{h}) \geq 0 \tag{3.8}$$

- satisfy each utility-maximising household who takes the allocations of others as given;

- and for which households are at least partly differentiated by type into different neighbourhood reference groups.

In addition, in order to eliminate degenerate solutions, I constrain the equilibrium to exclude allocations in which a disjoint set of types occupies a neighbourhood. For instance, this allows occupancy by the range $[w_1, w_2]$ but not by the discrete set$\{w_1, w_2\}$ for $w_1 \neq w_2$.

I restrict utility $U(\cdot)$ to be smoothly varying. For such functions, no continuous range of w will find the same (i.e., not varying with w) value of $\bar{h}$ to be optimal for interior allocations. Therefore, the above constraint against disjoint sets implies that equation (3.6) may be simplified to state that neighbourhoods are internally homogeneous:

$$h(w) = \bar{h}(w) \text{ for each } w \tag{3.6'}$$

[19] I will often refer to neighbourhoods, formally indexed by the continuous parameter n, by their equilibrium property, $\bar{h}$.

The case when all occupied neighbourhoods exhibit identical average conspicuous consumption $\bar{h}$ is a *pooling neighbourhood equilibrium.*

3.3.4 Land markets are required for separating equilibria

Not all of the land ownership scenarios listed above admit separating equilibria. I first dispense with the *free land* possibility for a large set of functional forms and later, in Section 3.3.12, show that the absentee landlord case is also incompatible with separating equilibrium. As an additional refinement to Definition 3.3.3, let an *assortative separating neighbourhood equilibrium* be one in which the allocation of household types to neighbourhood types is one to one.

Proposition 3.3.1. (Requirement for land market) *If land is unpriced, there is no assortative separating equilibrium of continuous types. If land is unpriced, $N(\cdot)$ is concave or convex and $H(h,\bar{h})$ is a function of either $h-\bar{h}$ or $h/\bar{h}$, there is no pure strategy separating equilibrium of continuous types.*

Proof. Consider the choice of neighbourhood $\bar{h}$ by agents of type w when a continuum of neighbourhood types exist. The first order condition for the choice of $\bar{h}$, when an optimum exists, is

$$0 = \frac{\partial U(x,h,\bar{h},\bar{\bar{h}})}{\partial \bar{h}} = F_1(x)\frac{\partial x}{\partial \bar{h}} + H_2(h,\bar{h}) + N_1(\bar{h},\bar{\bar{h}}) \tag{3.9}$$

When $p(\bar{h}) = 0$, that is when the choice of neighbourhood has no direct bearing on a household's budget, $\partial x/\partial \bar{h} = 0$. Therefore, when (3.9) is evaluated at the equilibrium condition $h = \bar{h}$, it becomes

$$H_2(\bar{h},\bar{h}) + N_1(\bar{h},\bar{\bar{h}}) = 0 \tag{3.10}$$

which implicitly specifies the same choice(s) of $\bar{h}$ for all agents regardless of type, w. Therefore there is no unique sorting of types into neighbourhoods based on w — that is, no assortative separating neighbourhood equilibrium.

Furthermore, if $H(\cdot)$ takes the special forms $f\left(h-\bar{h}\right)$ or $f\left(\frac{h}{\bar{h}}\right)$, then $H_2(\bar{h},\bar{h})$ has value $-f'(0)$ or $-\frac{1}{\bar{h}}f'(1)$, respectively. In either case $\partial H_2(\bar{h},\bar{h})/\partial \bar{h} = 0$ and since $N_{11} \neq 0$, the left hand side of Equation equation (3.10) is monotonic and thus there is at most a unique solution for $\bar{h}$ and therefore no separating equilibrium.

□

This result may seem unintuitive in the context of the literature on discrete Tiebout equilibria, and it is difficult to find a good conceptual description to complement the proof. The impossibility of a separating equilibrium comes about because agents have two continuous choices to make but equilibrium requires that they align along a single dimension: the assortment of types into neighbourhoods. Without another price to clear the market in neighbourhood choice, the two sets of first order conditions cannot be simultaneously satisfied while meeting the equilibrium condition that $h = \bar{h}$.

3.3.5 Some general properties of equilibrium with a land market

Let $h^{\star}\left(\bar{h}\right)$ be the consumption level chosen optimally in a given neighbourhood with average consumption $\bar{h}$, and consider a utility function for which the indirect utility

$$U(w,\bar{h}) = U\left(w, h^{\star}\left(\bar{h}\right), \bar{h}\right) \tag{3.11}$$

is globally concave and in which x is essential, *i.e.,* $F(x) \to -\infty$ as $x \to 0$. Then the necessary optimality conditions for each household's choice of housing $h \geq 0$ and leisure $0 \leq x \leq 1$ take the following form:

$$F'(x) - wH_h(h,\bar{h}) - w\xi = 0 \quad \text{and} \tag{3.12}$$

$$\xi\left[r - h - p(\bar{h})\right] = 0 \tag{3.13}$$

where ξ is a Lagrange multiplier for the $x \leq 1$ constraint, which is equivalent to $h + p(\bar{h}) \geq r$. Since r and $p(\bar{h})$ are each nonnegative, this condition is stronger than $h \geq 0$, which therefore becomes redundant. For the choice of neighbourhood consumption $\bar{h} \geq 0$, necessary optimality conditions are:

$$\left[F'(x) - w\xi\right] p'(\bar{h}) + wH_{\bar{h}}(h,\bar{h}) - wN_{\bar{h}}(\bar{h},\bar{\bar{h}}) \geq 0 \quad \text{and} \tag{3.14}$$

$$\left[\left[F'(x) - w\xi\right] p'(\bar{h}) + wH_{\bar{h}}(h,\bar{h}) - wN_{\bar{h}}(\bar{h},\bar{\bar{h}})\right]\bar{h} = 0 \tag{3.15}$$

Considering interior values of h and $\bar{h}$, equation (3.12) can be used to eliminate $F'(x)$ in equation (3.15), providing a differential equation in $p'(\bar{h})$, h, $\bar{h}$, and $\bar{\bar{h}}$. Evaluating this at the equilibrium housing choice $h = \bar{h}$ gives:

$$p'(\bar{h}) = \frac{H_2(\bar{h},\bar{h}) + N_1(\bar{h},\bar{\bar{h}})}{H_1(\bar{h},\bar{h})} \tag{3.16}$$

Given a value for $p(0)$, equation equation (3.16) can be integrated to find the price of land for any neighbourhood. This property is used in the sections to follow.

3.3.6 Log-exp-log utility with equitable ownership

In order to find an explicit equilibrium solution for continuous types, I apply the equal ownership land model to the same functional form of utility used for discretely distributed types in

Section 3.2:

$$U(x,h,\bar{h}) = \Phi \log(x) - \Lambda \exp\left(-\lambda \left[h - \bar{h}\right]\right) + N \log\left(1 + \frac{\bar{h}}{\bar{\bar{h}}}\right) \tag{3.17}$$

For this specification, the house choice first order conditions (3.12) and (3.13), evaluated under the equilibrium condition $h = \bar{h}$, determine household leisure:

$$\begin{aligned} x(w) &= \min\left\{\frac{\Phi}{w\Lambda\lambda}, 1\right\} \\ &= \min\left\{\frac{w_0}{w}, 1\right\} \end{aligned} \tag{3.18}$$

where

$$w_0 \equiv \frac{\Phi}{\Lambda\lambda} \tag{3.19}$$

Households with productivity below w_0 choose not to expend any effort on building status symbols or buying into high-status neighbourhoods. Instead, they enjoy leisure $x = 1$ and pool together in a low-status neighbourhood where spending is funded entirely by the universal dividend income, r. Because neighbourhoods in equilibrium are characterised by homogeneous consumption, the marginal value of housing consumption is uniformly equal to

$$\left.\frac{\partial H(h,\bar{h})}{\partial h}\right|_{h=\bar{h}} = \Lambda\lambda$$

As a result, the minimum wealth level for entry into the workforce is independent of the distribution of others' types. Moreover, it does not depend on the household's preference $N(\cdot)$ for neighbourhood status[20] but solely on the relative importance of leisure versus "keeping up with the Jones" in one's own neighbourhood.

Using 3.18 with the condition that no income is wasted generates an equation governing the neighbourhood allocations necessary for equilibrium:

$$\bar{h}(w) + p\left(\bar{h}\right) = r + \max\{0,\ w - w_0\} \tag{3.20}$$

Denote by $\bar{h}_{\min}$ the solution to $\bar{h} + p(\bar{h}) = r$; this neighbourhood is the lowest possible occupied neighbourhood.

The differential equation for price, equation (3.16), becomes[21]

[20] Nor would it depend on the intertemporal elasticity of substitution of leisure, which in the current formulation is fixed to 1.

[21] A closed form of the indirect utility $U(\bar{h})$ based on results to follow shows that the second order condition for the choice of neighbourhood is satisfied for all values of p: $\frac{\partial^2 U\left(h^*(\bar{h}),\bar{h}\right)}{\partial \bar{h}^2} < 0$ for all parameter values. See Section 3.3.8.

$$p'(\bar{h}) = -1 + \frac{N}{\left[\bar{h}+\bar{\bar{h}}\right]\Lambda\lambda} \tag{3.21}$$

First, note that the sign of $p(\cdot)$ is indeterminate. In fact, while $p'(0)$ is positive if $N > \Lambda\lambda\bar{\bar{h}}$, it is negative otherwise; p has a maximum at $\bar{h} = \frac{N}{\Lambda\lambda} - \bar{\bar{h}}$. If $N < \Lambda\lambda\bar{\bar{h}}$, therefore, price is decreasing in neighbourhood affluence for all occupied neighbourhoods. This situation corresponds to preferences in which the neighbourhood status term N is relatively weak compared with consumption comparisons against immediate neighbours, and the average productivity is high.[22]

Land management firms would be willing to lend land for free but, in accordance with equation (3.8), are never willing to pay households to occupy land. Therefore, in equilibrium any continuous range of occupied neighbourhoods must bear a positive land price in order to meet condition equation (3.21).[23] It can be seen that for large $\bar{h}$, $p'(\bar{h}) \to -1$ and therefore p eventually crosses zero. Indeed, for $\bar{h}$ above some $\bar{h}_{\max}$, land price would need to be negative for neighbourhoods to be attractive to any household. The diminishing marginal returns to increasing status through neighbourhood choice are offset by the non-diminishing marginal cost to "keep up with the Jones" within a chosen neighbourhood.

In order to integrate equation (3.21) to find the price of land in any neighbourhood, a boundary condition on $p(\bar{h})$ is required. For the moment, let the integration constant remain unknown as p_0. Then equation (3.21) becomes

$$p(\bar{h}) = \max\left\{0,\, p_0 - \bar{h} + \frac{N}{\Lambda\lambda}\log\left(1+\frac{\bar{h}}{\bar{\bar{h}}}\right)\right\} \tag{3.22}$$

The price can now be eliminated from earlier expressions to find neighbourhood allocations as a function of r and $\bar{\bar{h}}$. Assuming that $p(\bar{h}(w)) > 0\ \forall w$, 3.20 and 3.22 can be combined to find

$$\begin{aligned}
\bar{h} &= r - p(\bar{h}) + \max\{0,\, w-w_0\} \\
&= r - p_0 + \bar{h} - \frac{N}{\Lambda\lambda}\log\left(1+\frac{\bar{h}}{\bar{\bar{h}}}\right) + \max\{0,\, w-w_0\} \\
\to \log\left(1+\frac{\bar{h}}{\bar{\bar{h}}}\right) &= \frac{\Lambda\lambda}{N}\left[r - p_0 + \max\{0,\, w-w_0\}\right] \\
\to \bar{h}(w, r-p_0, \bar{\bar{h}}) &= \bar{\bar{h}}\exp\left(\Lambda\lambda\frac{r-p_0+\max\{0,\, w-w_0\}}{N}\right) - \bar{\bar{h}} \qquad (3.23)
\end{aligned}$$

This states that in equilibrium household consumption choice of the Veblen good increases

[22] The endogenous value $\bar{\bar{h}}$ is expressed in terms of exogenous parameters below.

[23] This logic is the same reason that land pricing is necessary at all. See Proposition 3.3.1 on page 53.

convexly with productivity. Solving for $\bar{h}_{\min}$ gives

$$\bar{h}_{\min} = \bar{\bar{h}} \left[e^{\Lambda\lambda \frac{r-p_0}{N}} - 1 \right]$$

Let $\bar{h}_{\max}$ denote the upper root of $p(\bar{h}) = 0$ in equation (3.22). Below neighbourhood consumption level $\bar{h}_{\min}$, households cannot balance their budget in equilibrium without throwing income away. Above neighbourhood consumption level $\bar{h}_{\max}$, households would need to be compensated for occupying the land.[24]

If this upper limit on neighbourhood affluence is binding — that is, when $\bar{h}(w_H, r, \bar{\bar{h}}) > \bar{h}_{\max}$ — a separating equilibrium cannot exist. However, the next section shows that one can always find some price schedule which avoids this constraint.

3.3.7 General equilibrium averages

Denoting by $\langle \cdot \rangle$ an average over all types, the global average conspicuous consumption level is easily calculated from equation (3.20) as the total labour output in the production of housing:

$$\begin{aligned} \bar{\bar{h}} = \langle \bar{h} \rangle &= \left\langle r - p(\bar{h}(w)) + \max\{0,\ w - w_0\} \right\rangle \\ &= \langle \max\{w_0,\ w\} - w_0 \rangle \\ &= \begin{cases} \frac{w_H + w_L}{2} - w_0 & \text{if } w_L > w_0 \\ \frac{[w_H - w_0]^2}{2[w_H - w_L]} & \text{if } w_L \le w_0 \le w_H \end{cases} \end{aligned} \tag{3.24}$$

where I have used the fact that under uniform land ownership, $\langle r \rangle = \left\langle p(\bar{h}(w)) \right\rangle$.

Recalling that $w_0 = \frac{\Phi}{\Lambda\lambda}$, equation equation (3.24) states that when all households make interior choices, the average consumption of the Veblen good increases with the population average productivity in producing it, increases with the strength of the equilibrium *local* Veblen effect $\Lambda\lambda$ due to comparison with one's immediate neighbours, and decreases with the strength of preferences for leisure.

[24] Equation equation (3.23) shows that, while pooling behaviour amongst the least endowed types is possible at $\bar{h} = \bar{h}_{\min}$, pooling of multiple types is not possible in any neighbourhood with a higher level of affluence. The implication of a downward-sloped price curve and a non-negative land price is that the market may unravel if a sufficiently wealthy type of household exists. For w high enough, the effective marginal cost of neighbourhood membership outweighs the status benefit, and demand for land at non-negative prices is zero in all more affluent neighbourhoods. In order to be induced to settle there, affluent types would need to be subsidised to compensate them for their contribution to the neighbourhood's status. However, once again the land holding firms are unwilling to subsidise (equation equation (3.22)). Households with w greater than some $w_{\max}$ will prefer a neighbourhood $\bar{h}$ in equation (3.23) which will exceed $\bar{h}_{\max}$. Above $\bar{h} = \bar{h}_{\max}$, the land price $p(\bar{h})$ sticks at 0 and there is no way to satisfy wealthy households with pure strategies. The most wealthy with $w > w_{\max}$ would, in the absence of any available neighbourhoods $\bar{h}(w)$, prefer to settle in a community with $\bar{h}_{\max}$, but doing so would raise the average consumption level there, making it unattractive for its original occupants if the rent remains at $p = 0$. Thus those original residents would prefer to move "down" to a less affluent neighbourhood, and so on; the separated neighbourhoods unravel.

Defining $w_m \equiv \max\{w_L, w_0\}$ to be the lowest household type which chooses to work, a constraint on r follows from carrying out the integral over $p(w)$ explicitly. Using equation (3.23),

$$\begin{aligned}
\langle\bar{h}\rangle &= \bar{\bar{h}}\left\langle \exp\left(\Lambda\lambda\frac{r-p_0+\max\{0,\ w-w_0\}}{N}\right)-1\right\rangle \\
1+\frac{\langle\bar{h}\rangle}{\bar{\bar{h}}} &= \frac{1}{w_H-w_L}e^{\frac{\Lambda\lambda}{N}[r-p_0]}\int_{w_L}^{w_H} e^{\frac{\Lambda\lambda}{N}\max\{0,\ w-w_0\}}dw \\
\frac{\Lambda\lambda}{N}[r-p_0] &= \log\left(\left[1+\frac{\langle\bar{h}\rangle}{\bar{\bar{h}}}\right]\frac{w_H-w_L}{\int_{w_L}^{w_H} e^{\frac{\Lambda\lambda}{N}\max\{0,\ w-w_0\}}dw}\right)
\end{aligned}$$

Therefore,

$$\begin{aligned}
r-p_0 &= \frac{1}{\frac{\Lambda\lambda}{N}}\log\left(\frac{\left[1+\langle\bar{h}\rangle/\bar{\bar{h}}\right][w_H-w_L]}{\int_{w_L}^{w_m}1dw+\int_{w_m}^{w_H}e^{\frac{\Lambda\lambda}{N}[w-w_0]}dw}\right) \\
&= \frac{1}{\frac{\Lambda\lambda}{N}}\log\left(\frac{\left[1+\langle\bar{h}\rangle/\bar{\bar{h}}\right][w_H-w_L]}{w_m-w_L+\frac{1}{\frac{\Lambda\lambda}{N}}e^{-\frac{\Lambda\lambda}{N}w_0}\left[e^{\frac{\Lambda\lambda}{N}w_H}-e^{\frac{\Lambda\lambda}{N}w_m}\right]}\right) \\
&= \frac{1}{\frac{\Lambda\lambda}{N}}\log\left(\frac{\left[1+\langle\bar{h}\rangle/\bar{\bar{h}}\right][w_H-w_L]}{w_m-w_L+\frac{1}{\frac{\Lambda\lambda}{N}}e^{\frac{\Lambda\lambda}{N}[w_m-w_0]}\left[e^{\frac{\Lambda\lambda}{N}[w_H-w_m]}-1\right]}\right) \qquad (3.25)
\end{aligned}$$

In equilibrium, $\langle\bar{h}\rangle/\bar{\bar{h}}=1$. If $w_L > w_0$ (that is, for $w_m = w_L$), the above condition takes the form:

$$\begin{aligned}
r-p_0 &= \frac{1}{\frac{\Lambda\lambda}{N}}\log\left(\frac{2\frac{\Lambda\lambda}{N}[w_H-w_L]}{e^{\frac{\Lambda\lambda}{N}[w_L-w_0]}\left[e^{\frac{\Lambda\lambda}{N}[w_H-w_L]}-1\right]}\right) \\
&= w_0-w_L+\frac{1}{\frac{\Lambda\lambda}{N}}\log\left(\frac{2\frac{\Lambda\lambda}{N}[w_H-w_L]}{\left[e^{\frac{\Lambda\lambda}{N}[w_H-w_L]}-1\right]}\right) \qquad (3.26)
\end{aligned}$$

According to equation (3.25) and equation (3.26), r has a fixed relationship to p_0 based on exogenous parameters. Because $\bar{h}(\cdot)$ in 3.23 depends only on the difference $r-p_0$, expressed above, any choice of base price p_0 results in the same consumption allocations amongst separating equilibria. On the other hand, according to equation (3.22) the value of $\bar{h}_{\max}$, where $p(\bar{h})=0$, is monotonically increasing in p_0. Therefore an equilibrium price schedule which accomodates the highest household type always exists. That is, for some p_0 high enough, $p(\bar{h}(w_H)) > 0$ and thus $w_{\max} > w_H$. A higher p_0 simply means higher dividends for all households and a higher base price for land. The insensitivity of equilibrium allocations and utility to the choice of p_0 simplifies welfare analysis somewhat but does not offset the redistributive effect of common land ownership as compared with an absentee land owner model. The slope of the price curve is unaffected by p_0 but is central to the equilibrium distribution of outcomes through the opposing effects of making high-income neighbourhoods exclusive and through more strongly redistributing wealth.

3.3.8 Concavity

As discussed in Section B.4.1 of the Appendix for the case of discrete types, it remains to ensure that the household's problem is characterised by a global maximum. A second order sufficiency condition is that the price schedule presents a concave objective function for the indirect utility $U(w,\bar{h}) = U(w,h^{\star}(\bar{h}),\bar{h})$. Given a neighbourhood choice $\bar{h}$, the optimal household consumption level is

$$h^{\star}(\bar{h}) \quad = \quad \max\left\{r - p(\bar{h}),\ \left[w + r - p(\bar{h})\right] - \frac{1}{\lambda}\mathscr{L}(w,\bar{h})\right\} \tag{3.27}$$

where

$$\mathscr{L}(w,\bar{h}) \equiv \text{LambertW}\left(\frac{\Phi}{\Lambda}e^{\lambda\left[w+r-p(\bar{h})-\bar{h}\right]}\right)$$

Therefore the indirect utility is

$$\begin{aligned} U(w,\bar{h}) \quad = \quad & \Phi\log\left(\min\left\{\frac{w_0}{w},\,1\right\}\right) \\ & -\Lambda\exp\left(-\lambda\left[\max\left\{r - p(\bar{h}),\ \left[w + r - p(\bar{h})\right] - \frac{1}{\lambda}\mathscr{L}(w,\bar{h})\right\} - \bar{h}\right]\right) \\ & +N\log\left(1 + \frac{\bar{h}}{\bar{\bar{h}}}\right) \end{aligned} \tag{3.28}$$

Consider the case of interior equilibria. Then $r - p(\bar{h})$ in the above expression can be eliminated in favour of the constant $[r - p_0]$ using equation (3.22):

$$\begin{aligned} r - p(\bar{h}) &= r - p_0 - [p - p_0] \\ &= [r - p_0] + \bar{h} - \frac{N}{\Lambda\lambda}\log\left(1 + \frac{\bar{h}}{\bar{\bar{h}}}\right) \end{aligned}$$

to find

$$U(w,\bar{h}) = \Phi\log\left(\frac{w_0}{w}\right) + N\log\left(1 + \frac{\bar{h}}{\bar{\bar{h}}}\right)$$

$$
\begin{aligned}
&-\Lambda \exp\left(-\lambda\left[w+[r-p_0]-\frac{N}{\Lambda\lambda}\log\left(1+\frac{\bar{h}}{\bar{\bar{h}}}\right)\right.\right.\\
&\qquad\qquad\left.\left.-\frac{1}{\lambda}\mathrm{LambertW}\left(\frac{\Phi}{\Lambda}e^{\lambda\left[w+[r-p_0]-\frac{N}{\Lambda\lambda}\log\left(1+\frac{\bar{h}}{\bar{\bar{h}}}\right)\right]}\right)\right]\right)\\
&=\Phi\log\left(\frac{w_0}{w}\right)+N\log\left(1+\frac{\bar{h}}{\bar{\bar{h}}}\right)\\
&\quad-\Lambda\left[1+\frac{\bar{h}}{\bar{\bar{h}}}\right]^{-\frac{N}{\Lambda}}e^{-\lambda[w+[r-p_0]]}\\
&\quad\times\exp\left(\mathrm{LambertW}\left(\frac{\Phi}{\Lambda}\left[1+\frac{\bar{h}}{\bar{\bar{h}}}\right]^{-\frac{N}{\Lambda}}e^{\lambda[w+[r-p_0]]}\right)\right)
\end{aligned}
$$

which can be shown to have everywhere a negative second partial derivative with respect to $\bar{h}$.

3.3.9 Existence

The proof of the following existence claim is given in Section B.6 on page 169 of the Appendix and follows by construction from the preceding discussion.

Proposition 3.3.2. (Existence of separating equilibrium) *For preferences of the "LEL" form and with a continuum of types and neighbourhood locations, there is a unique allocation of consumption $x(w)$, $h(w)$, and $\bar{h}(w)$ conforming to the equilibrium of Definition 3.3.3.*

3.3.10 Welfare analysis of interior equilibria

The equilibrium utility can now be written in terms of exogenous parameters,

$$
\begin{aligned}
U(w) \quad = \quad & \Phi\log\left(\min\left\{\frac{w_0}{w},1\right\}\right)-\Lambda+\Lambda\lambda\max\{0,w-w_0\}\\
&+N\log\left(\frac{2[w_H-w_L]}{w_m-w_L+\frac{1}{\frac{\Lambda\lambda}{N}}e^{\frac{\Lambda\lambda}{N}[w_m-w_0]}\left[e^{\frac{\Lambda\lambda}{N}[w_H-w_m]}-1\right]}\right)
\end{aligned}
$$

Note that the last term depends on the distribution of types but not on individual w. Also, the equilibrium welfare does not depend on the choice of base price p_0 in the land market. Using the notation $\Theta \equiv w_H/w_L$, the utility for the interior case, when $w_L > w_0$, takes the form

$$
U(w)=\Phi\log\left(\frac{\Phi}{\Lambda\lambda w}\right)-\Lambda+\Lambda\lambda\left[w-w_L\right]+N\log\left(\frac{2\frac{\Lambda\lambda}{N}[\Theta-1]w_L}{e^{\frac{\Lambda\lambda}{N}[\Theta-1]w_L}-1}\right)
$$

For simplicity, the analysis to follow focuses on interior equilbria. Properties of this equlibrium can now be summarised as follows.

Intra-neighbourhood comparisons Welfare disparity is intensified not by the strength N of preferences over inter-neighbourhood comparisons, but by the strength of the local, intra-neighbourhood Veblen effect, $\Lambda\lambda$:

$$\frac{dU}{dw} = \Lambda\lambda - \frac{\Phi}{w} > 0$$

The negative term reflects the fact that to the extent that non-pecuniary pursuits are important to household utility, i.e. that Φ is large, endowment differences will not be reflected in welfare disparities.

Improvements to productivity As noted by *Eaton and Eswaran* [2006], improvements to productivity in the Veblen good industry can be harmful to welfare. Consider a multiplicative shift in the entire range of household productivities. This corresponds to raising or lowering w_L while holding Θ constant.

To assess the implication of an increase in productivity within a heterogeneous population, two marginal effects must be considered. A given household will experience individual productivity enhancement $dw = \frac{w}{w_L} dw_L$. The household's change in utility will be the sum of a component due to this individual shift within the distribution $U(w)$ and one due to the changing distribution. The latter effect is

$$\left.\frac{\partial U}{\partial w_L}\right|_{\Theta} = -\Lambda\lambda + \frac{N}{w_L} - \Lambda\lambda \frac{[\Theta - 1] e^{\frac{\Lambda\lambda}{N}[\Theta-1]w_L}}{e^{\frac{\Lambda\lambda}{N}[\Theta-1]w_L} - 1} \tag{3.29}$$

$$= -\Lambda\lambda + \frac{N}{w_L} - \Lambda\lambda \frac{\Theta - 1}{1 - e^{-\frac{\Lambda\lambda}{N}[\Theta-1]w_L}} \tag{3.30}$$

which fits a form of the function $\Psi(\cdot)$ defined and characterised in Lemma B.5.3 on page 167 on page 167:

$$\left.\frac{\partial U}{\partial w_L}\right|_{\Theta} = -\Lambda\lambda + \Psi\left(-\Lambda\lambda[\Theta - 1], \frac{w_L}{N}\right) < 0$$

The inequality follows from the property that $\Psi(-a, b) < 0$ for positive a and b. The overall marginal effect on a given household of rescaling productivity is the sum of the individual and distributional effects:

$$dU = \frac{\partial U}{\partial w} dw + \left.\frac{\partial U}{\partial w_L}\right|_{\Theta} dw_L$$

$$
\begin{aligned}
&= \frac{\partial U}{\partial w}\frac{w}{w_L}dw_L + \left.\frac{\partial U}{\partial w_L}\right|_{\Theta} dw_L \\
&= \left[\Lambda\lambda - \frac{\Phi}{w}\right]\frac{w}{w_L}dw_L + \left[-\Lambda\lambda + \Psi\left(-\Lambda\lambda[\Theta-1], \frac{w_L}{N}\right)\right] dw_L \\
&= \left[\Lambda\lambda\left[\frac{w}{w_L} - 1\right] - \frac{\Phi}{w_L} + \Psi\left(-\Lambda\lambda[\Theta-1], \frac{w_L}{N}\right)\right] dw_L \qquad (3.31)
\end{aligned}
$$

Numerical simulations of this function are explored below. The second and third terms are strictly negative for positive dw_L, and for large Φ in this pure Veblen labour economy every individual is worse off when productivities of each participant household are uniformly scaled up.

In general, growth in this context has negative welfare implications for the least wealthy, and may have positive benefits for the wealthiest.

The homogeneous population case from *Eaton and Eswaran* [2006] can be recovered by noting from Lemma B.5.3 on page 167 that

$$\lim_{\Theta\to 1} \left.\frac{dU}{dw_L}\right|_{\Theta} = -\frac{\Phi}{w}$$

That is, for homogeneous populations with sufficient productivity to merit production in the Veblen good industry, any increase in productivity is uniformly bad for welfare.

Helping the poor Indeed, even raising the productivity of only the poorest is bad for everyone else's welfare, a counterintuitive result when thinking is conditioned by non-Veblen goods models:[25]

$$
\begin{aligned}
\left.\frac{\partial U}{\partial w_L}\right|_{w_H} &= -\Lambda\lambda - \frac{N}{w_H - w_L} + \frac{\Lambda\lambda}{e^{\frac{\Lambda\lambda}{N}[w_H - w_L]} - 1} \\
&= -\Lambda\lambda - \Psi\left(\frac{w_H - w_L}{N}, \Lambda\lambda\right) < 0
\end{aligned}
$$

Wealthy and Veblen good productivity Increasing productivity in this model is, however, not bad policy in all cases. Adding wealthy households to the economy is beneficial for everyone

[25] This experiment consists of removing the least productive households from the economy. Therefore, the welfare of the removed households is not included. However, when the loss of dividends r by the removed households outweighs the extra income from their improved w, they too will prefer to remain within the economy. Thus, removing them represents a Pareto decline.

due to the redistributive effects outweighing the comparison externality:

$$\left.\frac{\partial U}{\partial \Theta}\right|_{w_L} = w_L \left.\frac{\partial U}{\partial w_H}\right|_{w_L} = \frac{N}{\Theta - 1} - \frac{\Lambda\lambda w_L}{e^{\frac{\Lambda\lambda}{N}[\Theta-1]w_L} - 1} \tag{3.32}$$
$$= \Psi\left(\Lambda\lambda w_L, \frac{\Theta - 1}{N}\right) > 0$$

Disparity In order to investigate the effect of disparity, consider next a mean-preserving spread in the distribution of w. Rewriting $w_L = \langle w \rangle - \frac{1}{2}\Delta$ and $w_H = \langle w \rangle + \frac{1}{2}\Delta$, the effect of a change in the range Δ is:

$$\begin{aligned}
\left.\frac{\partial U}{\partial \Delta}\right|_{\langle w \rangle} &= \left.\frac{\partial}{\partial \Delta}\right|_{\langle w \rangle}\left(\Phi \log\left(\frac{\Phi}{\Lambda\lambda w}\right) - \Lambda + \Lambda\lambda\left[w - \langle w \rangle + \frac{\Delta}{2}\right] + N \log\left(\frac{2\frac{\Lambda\lambda}{N}\Delta}{e^{\frac{\Lambda\lambda}{N}\Delta} - 1}\right)\right) \\
&= \frac{1}{2}\Lambda\lambda + \frac{N}{\Delta} - \frac{\Lambda\lambda}{1 - e^{-\frac{\Lambda\lambda}{N}\Delta}} \\
&= \frac{1}{2}\Lambda\lambda + \Psi\left(-\Lambda\lambda, \frac{\Delta}{N}\right) < 0
\end{aligned}$$

Increasing exogenous disparity at a constant mean productivity does not affect average consumption $\left(\frac{d\bar{\bar{h}}}{d\Delta} = 0\right)$ nor the price schedule $\left(\frac{dp(\bar{h})}{d\Delta} = 0\right)$ but is uniformly bad for welfare for all extant households. This comes about because when the spread Δ of household productivities increases, the average cost of housing of the new high types and new low types, combined, is less than the old average. In other words, the dividends r decrease:

$$\begin{aligned}
\left.\frac{\partial r}{\partial \Delta}\right|_{\langle w \rangle} &= \left.\frac{\partial}{\partial \Delta}\right|_{\langle w \rangle}\left(-\langle w \rangle - \frac{\Delta}{2} + \frac{\Phi}{\Lambda\lambda} + \frac{N}{\Lambda\lambda}\log\left(\frac{2\Lambda\lambda\Delta}{N\left[1 - e^{-\frac{\Lambda\lambda}{N}\Delta}\right]}\right)\right) \\
&= -\frac{1}{2} + \frac{N}{\Lambda\lambda\Delta} - \frac{e^{-\frac{\Lambda\lambda}{N}\Delta}}{1 - e^{-\frac{\Lambda\lambda}{N}\Delta}} \\
&= -\frac{1}{2} + \frac{N}{\Lambda\lambda\Delta} - \frac{1}{e^{\frac{\Lambda\lambda}{N}\Delta} - 1} \\
&= -\frac{1}{2} + \Psi\left(1, \frac{\Lambda\lambda}{N}\Delta\right) < 0
\end{aligned}$$

The inequality follows, once again, from Lemma B.5.3 which shows that

$$\lim_{b \to 0} \Psi(1, b) = \frac{1}{2}$$

and that $\frac{d}{db}\Psi(1,b) < 0$.

3.3.11 Empirical interpretation

A central feature of the empirical results of *Barrington-Leigh and Helliwell* [2007b] is that the well-being effect of a marginal change in the affluence of one's immediate neighbours is much smaller than the effect of a marginal change in broader consumption averages, which are strongly negative. Accordingly, the separating equilibrium modeled here has the feature that, after households have chosen their neighbourhood reference group, $dU/d\bar{h} = 0$ but $dU/d\bar{\bar{h}}$ is significantly negative.

3.3.12 Log-exp-log utility with absentee landlords

In the previous section, feedback from the aggregate effects of the distribution over types onto the household decision problem comes through both r and $\bar{\bar{h}}$. When land rents are high, the equitable land ownership model significantly redistributes income by returning land rents uniformly to all households, thus narrowing the relative dispersion in wealth and consumption. Because the distribution of consumption is central to household choices and to welfare analysis, the details of how land equity is distributed matters in interpreting equilibrium outcomes.

An alternative extreme is for none of the land rents to be returned to households in the economy; this is the case of absentee landowners. Welfare analysis is also complicated, however, when rents are paid to absentee landowners unless the welfare of those landowners is somehow included in the accounting.

When $r = 0$, households who choose not to work have no outside income with which to pay for land or housing. These households with $w < w_0$ prefer to pool together in a "slum" enjoying leisure $x = 1$, no conspicuous housing consumption, and a reference neighbourhood with zero consumption. That is, $\bar{h}_{\min}$ becomes 0 and the price of land there, p_0, must also be zero. Thus, with $r - p_0 = 0$, equation 3.25 becomes a knife-edge constraint on parameters. Except for certain peculiar parameter sets, there are no separating equilibria when land is owned by absentees and rents leave the economy.

3.3.13 Pooling equilibria

When all neighbourhoods have an equivalent mix of types, $\bar{h} = \bar{\bar{h}}$ for all households. Household choice of consumption h and leisure $x = 1 - h/w$ maximises

$$U(x,h,\bar{h}) = \Phi \log(x) - \Lambda \exp\left(-\lambda\left[h - \bar{h}\right]\right) + N\log(2) \tag{3.33}$$

$$h(w) = \begin{cases} 0, & \text{for } w \leq w_0 e^{-\lambda\bar{h}} \\ w - \frac{1}{\lambda}\text{LambertW}\left(\frac{\Phi e^{\lambda w - \lambda\bar{h}}}{\Lambda}\right), & \text{otherwise} \end{cases}$$

The average consumption $\bar{\bar{h}} = \bar{h}$ can be expressed recursively by computing the average value of $h(w)$:

$$
\begin{aligned}
\bar{h} &= \left\langle \max\left\{0,\ w - \frac{1}{\lambda}\mathrm{LambertW}\left(\frac{\Phi e^{\lambda w - \lambda \bar{h}}}{\Lambda}\right)\right\}\right\rangle \\
&= \frac{w_H^2 - w_m^2}{2\left[w_H - w_L\right]} - \frac{\mathrm{LambertW}\left(\frac{\Phi e^{\lambda w_H - \lambda \bar{h}}}{\Lambda}\right)^2 - \mathrm{LambertW}\left(\frac{\Phi e^{\lambda w_m - \lambda \bar{h}}}{\Lambda}\right)^2}{2\lambda^2\left[w_H - w_L\right]} \\
&\quad - \frac{\mathrm{LambertW}\left(\frac{\Phi e^{\lambda w_H - \lambda \bar{h}}}{\Lambda}\right) - \mathrm{LambertW}\left(\frac{\Phi e^{\lambda w_m - \lambda \bar{h}}}{\Lambda}\right)}{\lambda^2\left[w_H - w_L\right]}
\end{aligned}
\tag{3.34}
$$

where $w_m = w_m(\bar{h}) \equiv \max\left\{w_L,\ w_0 e^{-\lambda \bar{h}}\right\}$. Equation 3.34 may be solved numerically for $\bar{h}$, from which values for $\bar{h}(w)$ and $U(w)$ follow. In this equilibrium, each household randomises its choice of neighbourhood since all neighbourhoods are alike and present the same environment $\bar{h} = \bar{\bar{h}}$. No deviation to another neighbourhood is beneficial and households choose only their individual consumption, h, given the global mean consumption level. This global consumption level is determined by the collective external effects of each household's choice of h. Properties of such pooling equilibria are demonstrated numerically, below.

3.3.14 Planner's problem

In an economy with a pure Veblen good such as the one modeled here, a reasonable policy for a planner is to prevent, for instance through prohibitive taxation,[26] any production of the Veblen good at all. Under this constraint, all households enjoy leisure $x = 1$ and inhabit identical neighbourhoods with $\bar{h} = 0$. The utility in this case is uniformly

$$U = -\Lambda + N\log(2)$$

Below I demonstrate numerically, echoing the earlier results using discrete neighbourhoods, that this outcome does not necessarily Pareto dominate the disaggregated decision equilibrium in which households consume the Veblen good and separate into reference neighbourhoods. This constitutes an important difference from the findings of *Eaton and Eswaran* [2006].

[26]Here a relevant distinction is between a status good valued through a comparison of actual consumption and one which is valued by its cost to the buyer. The former is treated in this paper, while the latter is sometimes referred to as a "snob" good. In the snob good case, taxing the good may not affect net houshold expenditure on it, but it will still decrease its production and redistribute the revenue.

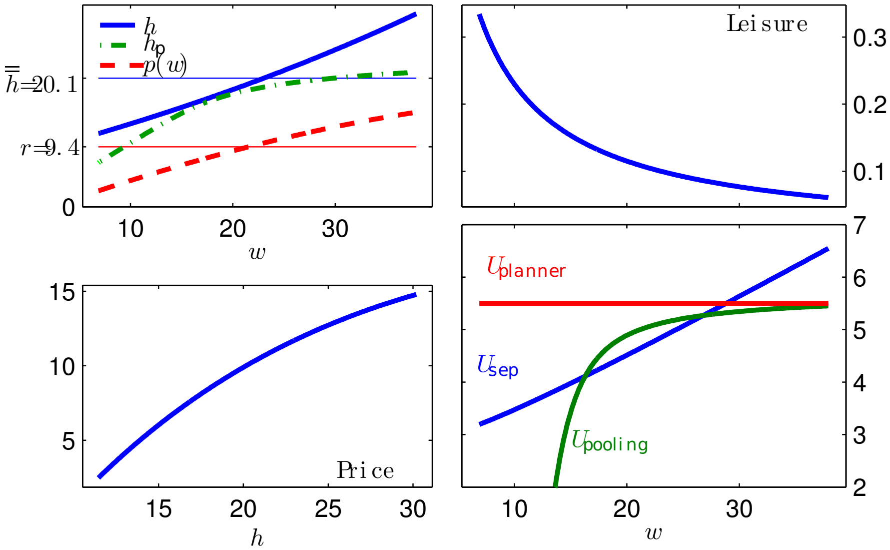


Figure 3.5: An equilibrium with monotonically increasing price amongst occupied neighbourhoods. Parameters: $w \in (6.9, 37.9), \Phi = 0.3, \Lambda = 0.2, \lambda = 0.6, N = 8.2$

3.4 Numerical analysis

This section demonstrates through simulations some of the features that have been described analytically, and emphasises the diversity of possible outcomes given different choices of parameters. Appendix B.6 outlines the method used here for numerically constructing separating equilibria for the problem described in Section 3.3.6. The pooling equilibria could not be characterised in closed form, so these have also been simulated numerically by solving equation 3.34.

Figure 3.5 depicts equilibria for one sample set of parameters. The top left panel shows the separating equilibrium distribution of household consumption $\bar{h}(w) = h(w)$ as well as its mean value $\bar{\bar{h}}$. Also shown is the rent $p(w)$ paid for land by each type w and its mean value r. For this economy, households all spend more on their housing than they do on buying their way into a neighbourhood. Also shown for comparison is the pooling equilibrium outcome $h_p(w)$; in this case all households spend less when mixed in identical heterogenous neighbourhoods than when they are sorted into their preferred reference groups.

The top right panel shows the dependence of leisure on type for the separating equilibrium. In all cases, it is weakly concave and decreasing. The lower left panel shows land price as a function of neighbourhood consumption. For the parameters used in this case, the price is an increasing function of neighbourhood affluence.

The lower right panel shows welfare distributions for three scenarios: the separating equilibrium (U_{sep}), the pooling equilibrium (U_{pooling}), and the planner's economy (U_{planner}, see

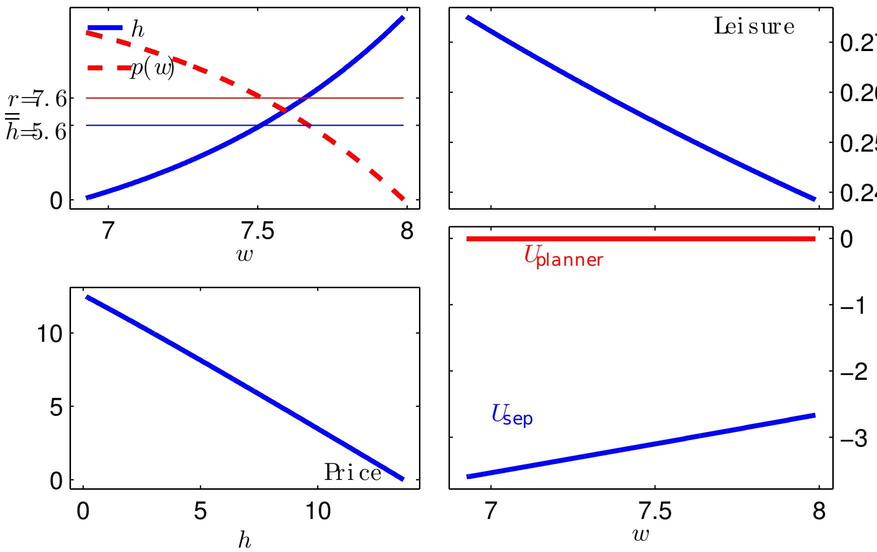


Figure 3.6: An equilibrium for which $r > \bar{\bar{h}}$. Parameters: $w \in (6.9, 8.0), \Phi = 2.3, \Lambda = 0.7, \lambda = 1.6, N = 1.0$

Section Section 3.3.14), in which no one consumes the Veblen good and all households enjoy the maximum amount of leisure. The planner's economy Pareto dominates the pooling equilibrium but is preferred to the separating equilibrium only by the lower types.

Figure 3.6 shows a case with several qualitative differences. For these parameters, the land rent is a *decreasing* function of neighbourhood affluence, indicating that the marginal cost of a higher local reference group outweighs the marginal benefit of a higher-status neighbourhood.

The preferences in this example differ from those in Figure 3.5 in part by having a much higher relative weight on local comparisons as compared with neighbourhood comparisons; that is, $N/\Lambda\lambda$ is much smaller. In this case, the emphasis in preferences on consumption comparison at the local level as compared with leisure is also high, and the Veblen equilibrium is fully Pareto dominated by the planner's outcome with no Veblen good.[27] The significance of allowing for endogenous reference group choice is highlighted in this economy by the fact that households are spending more, on average, on reference group selection — *i.e.*, land — than on the underlying Veblen good itself.

Figure 3.7 shows a land rent schedule that is peaked at $\bar{h} = \frac{N}{\Lambda\lambda} - \bar{\bar{h}}$ (see Section 3.3.6). The equilibrium also includes pooling neighbourhoods in which households with low productivity choose not to work at all. In the case shown in Figure 3.8, every household type consumes more Veblen good and suffers lower utility in the pooling equilibrium than in the separating case. Both outcomes would be unanimously rejected in favour of the planner's allocation.

[27] No pooling equilibrium was found for this set of parameters.

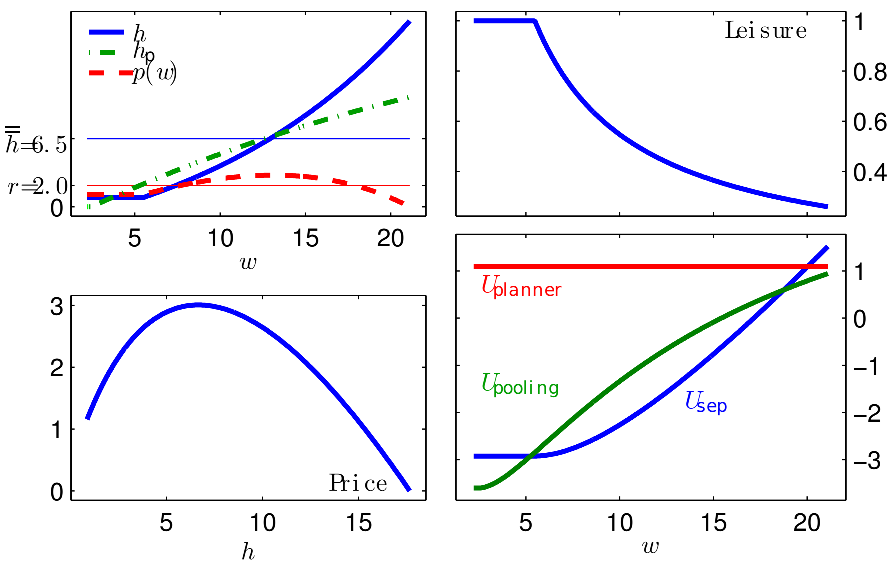


Figure 3.7: An equilibrium with non-monotonic price. Parameters: $w \in (2.2, 21), \Phi = 2.9, \Lambda = 3.8, \lambda = 0.1, N = 7.1$

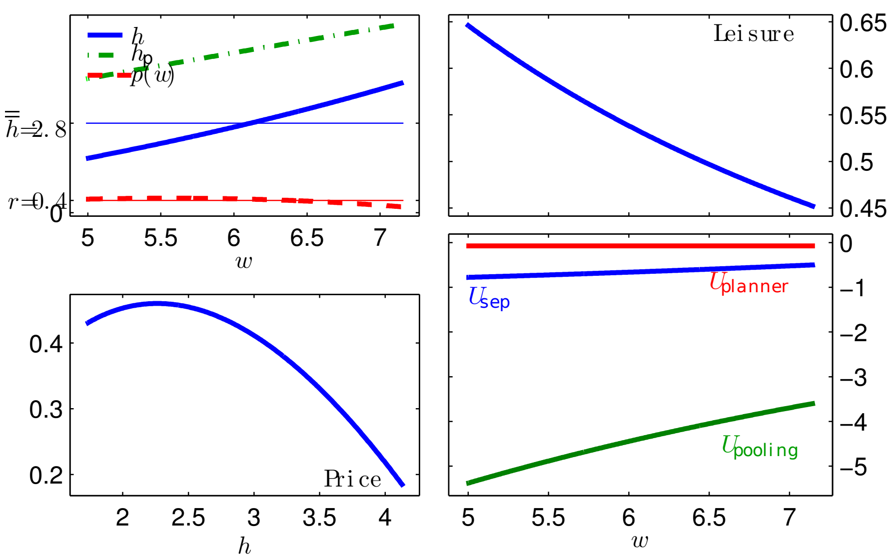


Figure 3.8: An equilibrium in which all households prefer the separating equilibrium to the pooling one. Parameters: $w \in (5.0, 7.2), \Phi = 0.9, \Lambda = 1.1, \lambda = 0.3, N = 1.4$

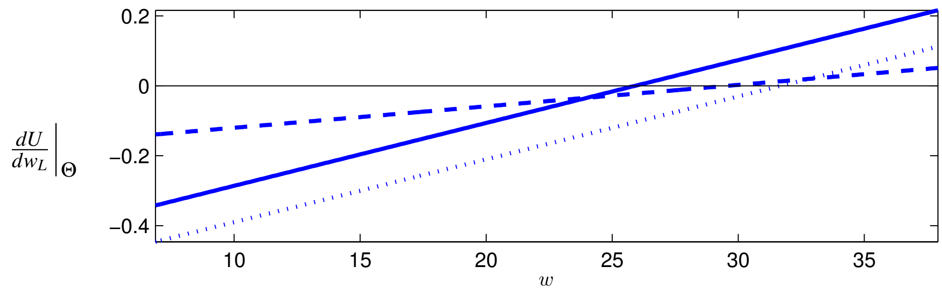


Figure 3.9: Total marginal change to welfare in an economy subject to uniform growth. The solid line represents the case of Figure 3.5 on page 66, the dashed line is the same case except that λ is decreased to 0.2, and the dotted line is the same case except that Φ is increased to 1.

Figure 3.9 shows the effect on households of uniform growth in the economy. As represented in equation (3.31) on page 62, the overall benefits of growth for a given household may be positive or negative. The three cases shown in the figure illustrate that the strength of the dependence of growth effects on w is proportional to λ and that when households place a higher value on leisure, the effect of growth is worse for all.

3.5 Conclusion

Most of economic analysis is still predicated on the plausibility of fixed preferences over absolute consumption. In this paper I take seriously the idea that absolute consumption utility benefits are unlikely to be sensible to humans when modern consumption levels are orders of magnitude higher in real terms than during the vast majority of our evolutionary history. I take a modest step towards exploring some calculus of choice and macroeconomic equilibria when preferences are, indeed, purely relative to what we know and see and when, in addition, we are able to exert some choice over what it is that we know and see.

By considering utility functions with a "pure Veblen" component this work accounts for goods which are consumed conspicuously or "publicly" and therefore are likely preferentially to affect neighbours in close proximity to the consumer. The benefits of "privately" consumed goods are captured in the so-called "leisure" term, $\Phi(x)$, which may encompass not only activities involving social engagement but also other classes of relative preferences for which reference levels are set through means other than the observation of local contemporaries. For instance, expectations about lifestyle and consumption are influenced by advertising and by broad dissemination of cultural norms.

It should above all be kept in mind that the empirical work which motivated this investigation of relative consumption preferences indicates that market-oriented consumption (as proxied by income) benefits are not only relative to others' but are also relatively insignificant for well-being as compared with the contribution from other factors such as positive social engagement. Thus, the importance of social groups in this work might correspond to the lesser of two significant roles: in a broader view, pursuit of social groups is important for the direct social benefits they confer as well as for their influence on emulation behaviour through consumption externalities.

What can be learned from a purely theoretical investigation is limited. Nevertheless, even the extreme models presented here suggest some insights to add to those developed in past work. Firstly, the allowance for heterogeneity and reference group selection significantly modifies the characteristics of general equilibrium. When agents have the tendency to use their own social group rather than a global one as a reference, the ability to differentiate into like groups can lead to a more efficient outcome than that of a heterogeneous mix of types, as evidenced by Figure 3.8. A general interpretation is that the existence of regional diversity can mitigate the extreme Veblen problem described by *Eaton and Eswaran* [2006]. On the other hand, this mitigation is by no means certain. In Figures 3.5 and 3.7, only the highest and lowest types prefer the separating equilibrium to the pooling one.

The most significant findings from this paper and some key differences from those of *Eaton and Eswaran* [2006] are that (1) complete elimination of the Veblen good may not be in everyone's interest and (2) growth in productivity in the Veblen good industry may be beneficial to some households. Even though the economy takes the form of a "rat race" due to the existence of a pure Veblen good, the most wealthy and productive households may actually prefer to have the good permitted on the market and prefer to have policy geared towards increased productivity in producing it. These features are shown, for example, in Figures 3.7 and 3.9. If the high types which benefit from the Veblen economy also have a more than proportional share of influence over policy, they will find it in their interest to promote the production of a completely "useless" good with ever increasing efficiency.

The degree to which such preferences and consequent externalities form an important part of the real economy remains an active empirical question, but the conclusions from this exploration of heterogeneity and autonomous group selection suggest that one should take seriously warnings given by *Eaton and Eswaran* [2006] about economists' canonical assumptions and focus on material consumption growth. If pursuit of Veblen good economies is not *pure* folly, it is only wise from the point of view of the wealthiest consumers.

Bibliography for Chapter 3

Barrington-Leigh, C. P., and J. F. Helliwell, Empathy and emulation: life satisfaction and the urban geography of Veblen effects, 2007b.

Clark, A. E., P. Frijters, and M. A. Shields, Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles, *The Journal of Economic Literature*, *46*, 95–144, 2008, doi:10.1257/jel.46.1.95.

Corless, R., G. Gonnet, D. Hare, D. Jeffrey, and D. Knuth, On the LambertW function, *Advances in Computational Mathematics*, *5*, 329–359, 1996.

Dunn, E., T. Wilson, and D. Gilbert, Location, Location, Location: The Misprediction of Satisfaction in Housing Lotteries, *Personality and Social Psychology Bulletin*, *29*, 1421–1432, 2003.

Eaton, B. C., and M. Eswaran, Well-Being and Affluence in the Presence of a Veblen Good, 2006.

Kingdon, G., and J. Knight, Community, Comparisons and Subjective Well-being in a Divided Society, *Journal of Economic Behaviour and Organization*, 2007.

Loewenstein, G., T. O'Donoghue, and M. Rabin, Projection bias in predicting future utility, *The Quarterly Journal of Economics*, *118*, 1209–1248, 2003.

Luttmer, E. F. P., Neighbors as Negatives: Relative Earnings and Well-Being, *Quarterly Journal of Economics*, *120*, 963–1002, 2005.

Rayo, L., and G. Becker, Evolutionary Efficiency and Mean Reversion in Happiness, *mimeographed, University of Chicago*, 2004.

Rothstein, J., Good principals or good peers? Parental valuation of school characteristics, tiebout equilibrium, and the incentive effects of competition among jurisdictions, *The American economic review*, *96*, 1333–1350, 2006.

Veblen, T., *The theory of the leisure class*, A.M. Kelley, 1899.

Chapter 4

Weather as a transient influence on survey-reported satisfaction with life

4.1 Introduction

Behind the compelling and growing modern evidence about what determines human well-being lie several qualitative claims concerning survey measures of satisfaction with life (SWL).[1] These are that (1) the meaning of standard SWL questions does not vary greatly between respondents from different languages and cultures, that (2) self-reported SWL measures something objective about a person's mental experience which reflects objective circumstances rather than solely individuals' fixed personality types, and that (3) SWL gets at a more lasting or long-term assessment of life quality than just an individual's current mood and its short-term influences. Generally speaking, these claims all have good support [for a brief review, see e.g. *Diener*, 2000] and there are a number of studies showing how the single-question SWL measure compares with other measures of well-being such as positive affect, low levels of negative affect, multi-question indices of life satisfaction and affect, experience sampling methods, and a number of physiological measurements.

Nevertheless, the reliability of life satisfaction data has often been held in low regard by economists on the general grounds that subjective responses may generate large statistical biases. The majority of the studies assessing the reliability and susceptibility to affective influence of reported life satisfaction are based on experiments with relatively low sample sizes. In order to test the robustness of statistical inference concerning the socioeconomic determinants of SWL, it is desirable to have access in a large survey to some random factor which can be expected to affect mood and thus any self-reported values affected by mood. Of primary interest in this regard are the measures of health, trust, and other major established determinants of SWL, as well as SWL itself. If transient influences on mood do not result in large correlated effects between SWL and its ostensible determinants, well-being researchers may rest assured that they are capturing meaningful relationships in ubiquitous econometric models.

Data from two Canada-wide surveys described below include not only the location of each respondent's home but also the precise day of each survey interview, which was conducted

[1] A version of this chapter will be submitted for publication as Barrington-Leigh, C.P., 'Weather as a transient influence on survey-reported satisfaction with life.'

by telephone. Canadian weather archives from the several months during which the surveys were conducted in 2002, 2003, and 2005 are used to determine the local weather conditions experienced by each respondent on the day of their interview. I find that these local weather conditions do indeed serve as a transient influence on both SWL and some of its self-reported determinants, yet I show that the correlations from this influence do not result in a significant bias of estimates for canoncial models of SWL.

The remainder of this section provides an overview of previous investigations into the psychological influences on subjective well-being assessments, the role of climate and weather in well-being and judgement, and the problem of accounting for geographical amenities in cross-sectional studies. Section 4.2 describes the surveys used and the linking of weather data to respondents. Section 4.3 presents the main findings and Section 4.4 concludes.

4.1.1 Reliability: does SWL vary too much?

Bertrand and Mullainathan [2001] discuss and test the reliability and statistical usefulness of survey subjective evaluations.[2] They conclude that subjective responses are unreliable as dependent variables in statistical models because a number of situational and psychological factors are likely to affect both the dependent and independent variables and may therefore cause arbitrarily large biases. Although *Bertrand and Mullainathan* [2001] describe the unwillingness of economists to use subjective data as an "important divide between economists and other social scientists," the role of SWL in economics as a measure of well-being has persisted and grown because regularities of relationships in modeled SWL seem unlikely to be explainable in terms of bias alone. The use in the present work of weather events as an exogenous situational influence makes possible a test for effects on the "right-hand side" variables in typical models for life satisfaction.

Turning more specifically to the central subjective measure of the present study, a considerable literature addresses the degree to which asking people about their SWL elicits meaningful and reproducible responses that are distinct from transient affect. *Krueger and Schkade* [2008] report that the SWL question has a lower consistency amongst individuals re-surveyed after two weeks than do either narrower domain satisfaction questions or measures of net affect.[3] Even though the major known determinants of life satisfaction are circumstances that can be expected not to change much on short time scales, the authors point out that the cognitive process invoked in evaluating SWL is naturally less systematic than and less well circumscribed than those of the

[2]While providing evidence that subjective evaluations *do* have useful explanatory power in predicting outcomes like wage and job turnover, *Bertrand and Mullainathan* [2001] provide only hypothetical problems rather than any statistical evidence for the kind of correlation which they conclude could invalidate the use of subjective measures as dependent variables.

[3]They define net affect as a duration-weighted difference between a composite measure of positive emotions — encompassing happy, affectionate/friendly and calm/relaxed — and one of negative emotions, encompassing tense/stressed, depressed/blue and angry/hostile.

more narrowly defined questions. Thus, while SWL may get at the ultimate outcome measure, it necessarily does so noisily. Despite this susceptibility to context dependence, *Krueger and Schkade* [2008] conclude that the consistency in life satisfaction responses is high enough to justify the typical statistical inferences being made in current research.

The open-endedness of the life satisfaction question means that the cognitive assessment which it elicits is susceptible to variation in focus based on any factor which makes a particular piece of evidence more or less salient, prominent, or subject to immediate attention. In comparison, introspection about mood or about domain satisfaction is a relatively well circumscribed task.

[*Schwarz and Strack*, 1991, p. 37] and others since have shown that making a mood-affecting factor such as weather more explicitly salient reduces its impact on self-reported satisfaction. Their interpretation is that current mood is one piece of evidence used to assess one's own longer-term well-being, but if transient influences on mood are identified or attention is drawn to them, their bias on perceived satisfaction can be cognitively corrected for.

For instance, when phone interviews were conducted on sunny or rainy days, the weather affected reported life satisfaction only when weather was not mentioned either in passing or as a context for the study [*Schwarz and Clore*, 1983]. More generally, when the relevance of momentary affect is drawn into question, subjects cease to let it inform their assessment of their life satisfaction [*Schwarz and Clore*, 1983].

On the other hand *Schkade and Kahneman* [1998] demonstrate how a *focusing illusion* can increase an individual's estimate of the salience of a given factor for SWL when that factor is mentioned or emphasized.[4] In their study, respondents overestimated the importance of climate in determining their life satisfaction when climate was the basis for a comparison with another region. In the present work, weather and climate are not discussed in the survey questions nor did they relate to the original or stated motivation for the surveys.

4.1.2 Meaningfulness: does SWL not vary enough?

Another strand of historical skepticism about subjective well-being studies relates to the opposite concern — that reported SWL does not vary sufficiently in relation to experienced circumstances because it is determined largely by personality. The two strands of objection correspond to two traditions in psychologists' understanding of reported satisfaction with life. These are judgement theories, which look at the momentary influences on the cognitive process of evaluating one's life, and personality theories, which focus on the influence of stable personality type in determining life satisfaction. *Schimmack et al.* [2002] offer an attempt to integrate the two traditions. They provide evidence that, at least amongst their rather uniform sample of students, life satisfaction judgements are made through a deliberate and consciously accessible process. This would help to explain the ability of respondents to discount factors which

[4] *Bertrand and Mullainathan* [2001] give a brief review of this and other possible kinds of biases in subjective responses.

have been deemed uninformative [*Schwarz and Clore*, 1983; *Schwarz and Strack*, 1991]. More generally, *Schimmack et al.* [2002] suggest that while people use readily available introspective evidence in making a life satisfaction assessment, consistency over time comes from the natural fact that accessible sources of information reflect important and repeatably salient aspects of people's lives.

An influence of culture and personality on reported SWL is mediated through the same channel: the perceived importance of different circumstances and domains of success and the strength of memories of emotional experiences reflect the priorities that define an individual's identity. In this sense, the meaning of an open-ended SWL question may not vary between people and cultures as much as the values which inform the answer.

The survey statistical approach typically used by economists studying life satisfaction naturally accounts for influences from both personality and socioeconomic circumstances, where such variables are available. Modern concensus is that reported life satisfaction has both meaningful variation over time and significant reproducibility and consistency over time. In accordance with the description and empirical evidence of *Schimmack et al.* [2002], the latter consistency reflects the information to which a respondent appeals when forming satisfaction assessments. Transient influences such as weather can be thought of as complications to those salient factors, when they are not cognitively compensated for or excluded, and it may be expected that more specific questions than SWL will suffer less from interference simply because the cognitive calculation and relevant pool of introspective information is simpler.

4.1.3 Stock markets and behaviour

The imperfect self-awareness that characterises cognitive assessments has also come up in evidence regarding economic decision making. Influences on mood affect judgement and behaviour through the misattribution of feelings to the wrong source. In this way, for example, mood-enhancing weather may mistakenly become confused with an optimistic assessment of future stock returns, in part by increasing the preceived salience of positive information. There is a small industry of studies on weather, moon phase, and stock returns [*Loughran and Schultz*, 2004; *Cao and Wei*, 2005; *Krämer and Runde*, 1997; *Yuan et al.*, 2006].

For instance, *Hirshleifer and Shumway* [2003] find a highly statistically significant relationship between morning sunshine and stock market performance amongst 26 countries, with cloudiness dominating precipitation as a measure of influence. As mentioned above, drawing attention to a particular influence on mood or explicitly highlighting it as a possible source of bias is likely to diminish the effect of misattribution. A related, preliminary study by *Guven* [2007] analyses the influence of weather, through mood, on household investment and consumption choices. He finds weather to be an appropriate instrument for mood and reports a number of quantifiable behavioural influences which indicate that positive mood has a significant effect on household economic decision making.

4.1.4 Sunlight and depression

Turning now to the specific effects of weather and daylight on well-being, the largest set of evidence relates to seasonality in depressive episodes, which has been recognised for millennia. In modern terminology, seasonal affect disorder (SAD) refers to psychopathologies with distinct seasonal variation for which the patient feels worst in winter [*Magnusson*, 2000, for a review]. Because SAD is thought to be caused primarily by a lack of sunlight, its incidence was expected to vary strongly with latitude as well as with other determinants of sunlight exposure, such as cloudiness. Many studies have addressed this question, however, and found mixed results. *Mersch et al.* [1999] survey the literature and find overall no correlation between latitude and the prevalence of SAD, indicating that seasonality in sunlight may not be the primary factor involved. They suggest that other factors like climate and social-cultural context are instead dominant determinants. They also cite studies suggesting that temperature or even precipitation may be significant factors in explaining differences in SAD incidence between different regions of the world and even the existence of "summer-SAD" in some places.

Furthermore, the incidence of suicide is generally peaked in the summer, when sunlight exposure is at its maximum. This, in conjunction with the relatively high prevalence of suicide in Scandinavia, has led to the proposition that increased sunlight might be associated with suicide risk. As with the contrary hypothesis concerning SAD, the evidence has not painted a simple picture. *Helliwell* [2007] surveys the relevant research and discusses the relationship between suicide and SAD. He then finds limited empirical evidence of a role for latitude in predicting suicide rates. Once again, social-cultural factors appear to be as successful as long or short duration daylight in explaining any correlation between latitude and psychological health.

4.1.5 Climate, geography, and well-being

While the link between long-term sunshine and measures of severely compromised well-being appears to be weak, a related question is how the more central well-being measure of SWL is affected by persistent aspects of climate, physical geography, and other environmental factors. Physical amenities and climate constitute an increasingly significant and marketable factor in migration between cities in the U.S.A. [*Rappaport*, 2007] and the looming task of mitigating the effects of climate change will require an understanding of the welfare implications of climatic factors.

Frijters and Van Praag [1998] construct an estimate of the direct climate costs of global warming using Russian reported satisfaction with life and satisfaction with income. Using geographic variation in mean annual climate, they find that households tend strongly to dislike cold, windy winters and hot, humid summers and that they benefit from higher annual hours of sunlight.

Rehdanz and Maddison [2005] use instead a cross-country comparison of overall happiness in 67 countries to anticipate the direct importance of climate change to the geographic distribution of well-being. Using several national control variables and climate parameters for temperature and precipitation, they find that more moderate temperatures — lower peaks and higher minima — are significantly preferred.

Brereton et al. [2008] use a similar approach to that of *Frijters and Van Praag* [1998] but for a small sample in Ireland and find that windiness and mean annual minimum and maximum temperatures are significant in explaining the geographic variation in SWL. They also find a slightly negative relationship between annual hours of sunshine and SWL but they explain this by appealing to other, unmeasured aspects of geography. In the approach I pursue below, unmeasured geographic variation should not bias results because geographic fixed effects are carefully controlled for. I am also able to compare the magnitude of the influence on SWL from essentially stochastic daily weather events with that due to long-term climatic differences, assuming people have not become strongly geogaphically sorted according to their preferences.

In any attempt to accomplish the just-described task of estimating the effect of regional variation in climate — rather than short-term weather — on SWL, one is confronted with the confounding effect of variation in other geographic amenities. There is a considerable literature treating such "hedonic geography." In addition to the climate studies already discussed, estimates based on SWL have been conducted for aircraft noise near an airport [*van Praag and Baarsma*, 2005], NO_2 air pollution [*Welsch*, 2006], and proximity to the workplace as measured by commuting time [*Stutzer and Frey*, 2004]. *Moro et al.* [2008] use a model of geographic amenities to construct a geographic estimate of SWL by weighting the environmental endowments of each Irish county by the marginal rate of substitution between income and the amenity. They find that this estimate provides a similar ranking to others based more directly on actual reported SWL in each county. In their related work, *Brereton et al.* [2008] conclude that incorporating various geographic factors across Ireland generates a marked increase in the proportion of explained variance in SWL.

Numerous other studies use market outcomes such as house prices rather than SWL to evaluate the well-being contribution of geographic amenities. This hedonic price approach is, however, predicated on a frictionless market in which there are insignificant costs to moving [*Gyourko et al.*, 1999, for a discussion]. Given that in the U.S.A., 57%-79% of Americans reside near where they were born [*Bayer and McMillan*, 2005], this assumption is a poor one. In the opposite case when markets for location are highly frictional and migration is small, correlations between geographic amenities and SWL are more likely to reflect a causal relationship.

4.2 Data and Method

Two surveys in Canada are suited to the current task. The second wave of the Equality, Security, and Community survey (ESC2)[5] includes 5600 respondents interviewed between December 2002 and July 2003. Rather than being uniformly distributed over time, the sampling was

[5]ESC2 is described by *Soroka et al.* [2007] and online at `http://grad.econ.ubc.ca/cpbl/esc2`.

strongly peaked in April to May. Data for Cycle 19 of the General Social Survey (GSS19) were collected in 11 monthly samples from January to November 2005 with data collection for the November sample extending until mid-December. The sampling was evenly distributed over the 11 months.

Both surveys asked respondents to rate their overall life satisfaction on a ten point scale with bipolar verbal descriptions. ESC2 asked:

> On a scale of 1-10 where ONE means dissatisfied and TEN means satisfied, all things considered how satisfied are you with your life as a whole these days?

while in GSS19 the question was phrased:

> Please rate your feelings about them, using a scale of 1 to 10 where 1 means ”Very dissatisfied” and 10 means ”Very satisfied”. ... Using the same scale, how do you feel about your life as a whole right now?

Numerous other questions relevant to social interactions and socioeconomic and cultural backgrounds were posed in these surveys. Of the nearly 20,000 respondents surveyed in GSS19, all were asked the SWL question but just less than half were asked to evaluate their level of trust in neighbours, an important metric for local social capital. Also, nearly 5000 respondents declined to provide an income, half of whom chose “don’t know”. In regressions below where these measures are used, the sample size is accordingly smaller. Household survey weights are available for GSS19 and are used in all estimates.

4.2.1 Assignment of weather stations

Environment Canada offers several kinds of historical weather and climate data via the Internet. Of 2108 weather stations across Canada, a subset recorded daily weather summaries for the years 2002-2005 and a smaller set offer hourly information on sky conditions. These include the cloud fraction and facilitate the calculation of the sunniness of daytime weather for each day.[6] In addition, monthly climatic averages and daily “almanac” averages are available for some stations.

There is no single optimal algorithm for assigning a weather station to each survey respondent. For statistical models which do not include fixed effects for each weather station, the closest suitable station can be used for each respondent irrespective of the number of neighbours assigned to the same station. In some cases, more than one station is used per respondent, such as when the nearest station providing hourly cloud cover data is different from the nearest station providing daily precipitation levels.

On the other hand, for models which involve a constant term for each weather station, there is a tradeoff between minimising the total number of stations used and minimising the distance

[6] Verbal descriptions of fractional cloud cover were coded numerically and averaged over 12 daytime hours.

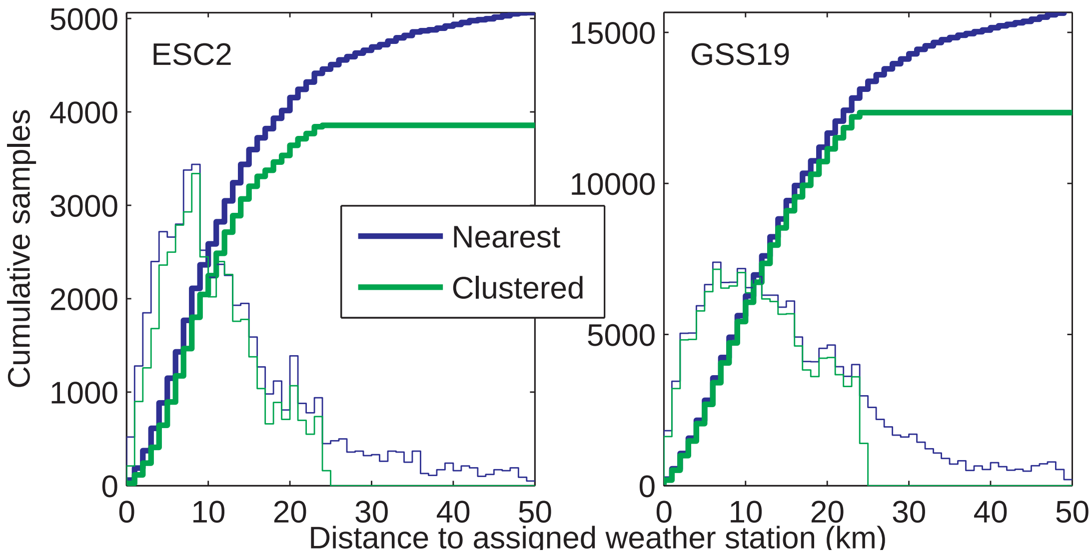


Figure 4.1: Comparison of the "nearest" and "clustered" algorithms for assigning weather stations to respondents. Plots show incremental and cumulative distributions of distance from the assigned station for each of the two surveys, ESC2 and GSS19.

between each respondent and her assigned weather station. For the latter purpose, a multi-step process involving successive reassignment was used to achieve a balance between the two objectives. In each stage, the least populous stations are dropped and respondents are assigned to the nearest station in the remaining set. Respondents who live beyond 20 km from one of the most popular stations are eventually dropped from the analysis. In addition, stations with fewer than ten respondents assigned to them are not included in the regressions to follow.

Altogether, half the GSS19 sample, or ∼10,000 respondents, survive this process when the "clustered" station algorithm is used while ∼12,500 respondents are matched using the "nearest" station algorithm. Of these, only ∼5200 have cloud cover data available from the clustered station algorithm and 5900 from the nearest station method. Figure 4.1 on page 79 shows the coverage of respondents by nearby weather stations for the ESC2 and GSS19 surveys and under the two assignment algorithms. In all cases, approximately half of the respondents are within 10 km of their assigned weather station. Estimates resulting from these two different assignment methods do not differ significantly, and the "cluster"-assigned data are used preferentially in all the results below.

4.3 Evidence and discussion

In this section the main findings are summarised in the form of regression coefficient tables. Because the estimates are primarily made for models of SWL, a proxy for utility itself, there is no structural equation framework motivating the analysis. Reduced form equations estimate the marginal effect of different circumstances on the outcome of interest. Rather than pooling data from two surveys which use different sampling methods, each equation is estimated separately for ESC2 and GSS19. In some tables, mean values of coefficients from the two surveys are reported.

4.3.1 Weather and well-being

Tables 4.1–4.5 report results from an investigation of the influence of weather on responses to several survey questions, including subjective measures of well-being.[7] For discrete dependent variables such as SWL and subjective assessments of trust and health, estimates from a logit or an ordered logit model are reported.[8] The model specifications focus on the average cloudiness over the week prior to the interview as an explanatory variable and show that once this and the same-day cloudiness are controlled for, the temperature and precipitation do not significantly affect outcomes.

Column 1 of Table 4.1 on page 81 shows a significant negative relationship between SWL and the seven-day cloudiness prior to the day of interview for GSS19 respondents when several sociodemographic variables, not including income or self-reported health, are controlled for. These controls encompass the essentially objective measures of sex, a quadratic in age, five dummies for marriage status, and five dummies for workforce status, along with two more subjective measures of religiosity. This set of controls is included[9] in every model throughout the paper but for compactness is generally not shown.

Even after including these important determinants of SWL, the remaining geographic variation in SWL may be correlated with recent weather. Since a sunny climate is likely to serve as a geographic amenity, one might expect to find higher incomes in sunnier locations, given a residential market with high mobility. One might also expect that objective health or at least subjectively reported health would be affected by climate or weather and thus account for some of the correlation between cloudiness and satisfaction with life. In columns 3 and 5, household income and self-reported health along with a subjective measure of trust in neighbours are included in the regression and result in no significant change in coefficients on cloudiness.

[7] The appendix and online supplement contain more complete versions of tables shown in the text.

[8] Raw coefficients are shown in the table. Logit and ordered logit models estimate the marginal change in probability, held uniform across different possible outcome values, of finding a higher dependent variable value for a given marginal change in an explanatory variable. To calculate the probability ratio between successive outcome possibilities, simply exponentiate the raw coefficient shown in the table.

[9] Not all variables are available in both surveys.

	SWL (1)	SWL (2)	SWL ⟨1-2⟩	SWL (3)	SWL (4)	SWL ⟨3-4⟩	SWL (5)	SWL (6)	SWL ⟨5-6⟩	SWL (7)	SWL (8)	SWL ⟨7-8⟩
clouds										−.19	−.12	−.17
										(.15)	(.22)	(.12)
clouds (7 days)	**−.77***	−.43	**−.68***	**−.94***	−.52	**−.81***	**−.78***	−.49	**−.70***	**−.81***	−.58	**−.74***
	(.22)	(.36)	(.19)	(.24)	(.38)	(.20)	(.24)	(.38)	(.20)	(.26)	(.39)	(.21)
T_{high} (°C)										.002	−6e-05	.001
										(.009)	(.011)	(.007)
T_{low} (°C)										−.0006	.007	.002
										(.009)	(.012)	(.007)
rain (mm)										.001	−.007	−.0001
										(.004)	(.010)	(.004)
snow (cm)										−.008	−.003	−.006
										(.016)	(.024)	(.013)
log(HH inc)				**.64***	**.47***	**.59***	**.36***	***.34***	**.35***	**.42***	***.40***	**.41***
				(.11)	(.16)	(.091)	(.11)	(.15)	(.091)	(.12)	(.15)	(.094)
health							**2.81***	**1.66***	**2.55***	**2.85***	**1.70***	**2.58***
							(.15)	(.28)	(.13)	(.16)	(.28)	(.14)
trust-N							**.51***	**.42***	**.46***			
							(.17)	(.14)	(.11)			
controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
survey	G19	E2	⟨2⟩	G19	E2	⟨2⟩	G19	E2	⟨2⟩	G19	E2	⟨2⟩
obs.	6359	1632	7991	5167	1496	6663	5161	1495	6656	4956	1495	6451
pseudo-R^2	.014	.031		.018	.033		.056	.043		.055	.042	

Table 4.1: Weather and satisfaction with life, without geographic controls. Raw ordered logit coefficients and standard errors are shown. A number of other demographic, individual, and household controls are included but not shown; see Table C.1 on page 173 for detailed results behind this and the following five tables. Significance: **1%*** ***5%*** *10%**

Corresponding results for the ESC2 survey, shown in columns 2, 4, and 6, are consistent with those for GSS19 but are based on a much smaller sample and are less significant. Taken together, the two surveys produce a significant negative coefficient for cloudiness, as shown in the greyed columns following each pair. These report weighted mean coefficients for the two surveys, using the reciprocal squared standard errors as weights.

The final two columns in Table 4.1 confirm that the additional same-day weather effects of temperature, precipitation, and cloudiness are insignificant. Further tests of these findings are shown in the Appendix.

In order to control for any seasonal variation in life satisfaction due to length of daylight or other annual cycles, monthly fixed effects were included and the findings are reported in Table 4.2. Adding these controls uniformly strengthens the estimated influence of recent cloudiness, possibliy indicating the importance of expectations in moderating the effect of weather on satisfaction with life. This possibility is revisited further on but the present interest is in isolating the effect of short term weather.

In Table 4.4 the estimated models include a dummy variable for each of 22 (for ESC2) or 49 (for GSS19) weather stations used in matching weather data to respondents with a minimal set of locations, i.e. via the "clustered" method. These stations are the ones with ten or more respondents nearby. Controlling for weather station fixed effects removes the confounding influence of most geographic variations in climate as well as other geographical amenities and local contextual effects. The coefficient estimated for cloudiness is only slightly diminished in this case and as an interesting side note, the effects of health and own trust in neighbours remain unchanged in this specification. The calculation of standard errors is performed with clustering at the same level as the fixed effect controls.

An account of the effect of short-term weather on SWL is only credible when the influence of climatic norms, which vary over both season and geography, is fully controlled for. Accordingly, the central result is presented in Table 4.4 which includes fixed effects for every possible combination of calendar month and weather station. Such clusters containing less than ten respondents are again dropped, diminishing the sample size somewhat. By including this generous set of controls, all aspects of the climate are accounted for and the seven-day cloudiness measure represents a highly exogenous event determined through the fully randomized algorithm of the survey sampling method, which for GSS19 was stratified by month and by geographic region. The estimates indicate a strong effect of recent cloudiness on SWL that is consistent between the two surveys, marginally significant for ESC2, and strongly significant within the larger sample of GSS19. The probability ratio corresponding to the recent cloudiness coefficient in column ⟨61–62⟩ of Table 4.4 is 0.53, indicating that a run of completely sunny weather increases the chance of an individual reporting an extra point higher on the ten-point SWL scale by over 20%, as compared with a completely overcast week.[10]

[10]Odds of reporting a higher score are even under sunny weather (i.e. with cloudiness=0) and the odds ratio is $\exp(-0.64) \approx 0.53$ under cloudy weather, meaning the odds of a higher SWL score are only half those of a lower

	SWL (19)	SWL (20)	SWL ⟨19-20⟩	SWL (21)	SWL (22)	SWL ⟨21-22⟩	SWL (23)	SWL (24)	SWL ⟨23-24⟩	SWL (25)	SWL (26)	SWL ⟨25-26⟩
clouds										−.23	−.17	−.20
										(.21)	(.22)	(.15)
clouds (7 days)	−**.83***	−.57	−**.75***	−**.99***	−*.69**	−**.89***	−**.91***	−*.68**	−**.84***	−**.87***	−.69	−**.82***
	(.31)	(.47)	(.26)	(.29)	(.40)	(.23)	(.26)	(.39)	(.22)	(.27)	(.44)	(.23)
T_{high} (°C)										−.004	.0006	−.003
										(.007)	(.012)	(.006)
T_{low} (°C)										−.009	.001	−.005
										(.007)	(.009)	(.006)
rain (mm)										.0002	−.008	−.0008
										(.003)	(.008)	(.003)
snow (cm)										−.010	−.003	−.009
										(.012)	(.030)	(.011)
log(HH inc)				**.64***	**.47***	**.54***	**.36***	*.33*	**.34***	**.42***	**.38***	**.40***
				(.15)	(.12)	(.094)	(.14)	(.13)	(.094)	(.13)	(.14)	(.096)
health							**2.81***	**1.66***	**2.56***	**2.84***	**1.70***	**2.58***
							(.14)	(.26)	(.12)	(.14)	(.25)	(.12)
trust-N							**.51***	**.44***	**.46***			
							(.19)	(.12)	(.098)			
controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
mnth f.e.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
clustering	mnth	mnth	mnth	mnth	mnth	mnth	mnth	mnth	mnth	mnth	mnth	mnth
survey	G19	E2	⟨2⟩	G19	E2	⟨2⟩	G19	E2	⟨2⟩	G19	E2	⟨2⟩
obs.	6359	1632	7991	5167	1496	6663	5161	1495	6656	4956	1495	6451
pseudo-R^2	.015	.033		.020	.035		.057	.045		.057	.044	
$N_{clusters}$	12	8		12	8		12	8		12	8	

Table 4.2: Weather and satisfaction with life, allowing for monthly fixed effects. Significance: **1%*** ***5%*** *10%**

	SWL (37)	SWL (38)	SWL ⟨37-38⟩	SWL (39)	SWL (40)	SWL ⟨39-40⟩	SWL (41)	SWL (42)	SWL ⟨41-42⟩	SWL (43)	SWL (44)	SWL ⟨43-44⟩
clouds										−.14	−.013	−.12
										(.10)	(.22)	(.094)
clouds (7 days)	**−.71***	−.23	***−.50****	**−.84***	−.18	**−.58***	***−.65***	−.20	*−.42**	***−.68***	−.25	***−.49****
	(.26)	(.30)	(.20)	(.26)	(.32)	(.20)	(.31)	(.31)	(.22)	(.28)	(.32)	(.21)
T_{high} (°C)										−.003	−.007	−.004
										(.009)	(.013)	(.007)
T_{low} (°C)										.007	.015	.009
										(.008)	(.013)	(.007)
rain (mm)										.0006	−.010	−.0006
										(.004)	(.011)	(.004)
snow (cm)										−.012	−.003	−.008
										(.020)	(.024)	(.015)
log(HH inc)				**.67***	**.51***	**.61***	**.39***	**.38***	**.38***	**.45***	***.41***	**.44***
				(.13)	(.17)	(.10)	(.12)	(.15)	(.092)	(.13)	(.17)	(.10)
health							**2.84***	**1.74***	**2.64***	**2.89***	**1.76***	**2.70***
							(.12)	(.26)	(.11)	(.12)	(.26)	(.10)
trust-N							**.50***	***.38***	**.44***			
							(.16)	(.17)	(.12)			
controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
stn f.e.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
clustering	stn	stn	stn	stn	stn	stn	stn	stn	stn	stn	stn	stn
survey	G19	E2	⟨2⟩	G19	E2	⟨2⟩	G19	E2	⟨2⟩	G19	E2	⟨2⟩
obs.	6334	1594	7928	5147	1461	6608	5141	1460	6601	4928	1460	6388
pseudo-R^2	.020	.036		.025	.039		.062	.049		.063	.048	
$N_{clusters}$	50	22		50	22		50	22		49	22	

Table 4.3: Weather and satisfaction with life, allowing for local fixed effects. Significance: **1%*** ***5%*** *10%**

	SWL (55)	SWL (56)	SWL ⟨55-56⟩	SWL (57)	SWL (58)	SWL ⟨57-58⟩	SWL (59)	SWL (60)	SWL ⟨59-60⟩	SWL (61)	SWL (62)	SWL ⟨61-62⟩
clouds										−.23	−.35	***−.29****
										(.19)	(.22)	(.14)
clouds (7 days)	−.47	−.65	*−.52**	***−.71***	−.56	***−.67****	*−.67**	−.58	***−.64****	*−.67**	−.58	***−.64****
	(.34)	(.54)	(.29)	(.35)	(.52)	(.29)	(.37)	(.55)	(.31)	(.38)	(.53)	(.31)
T_{high} (°C)										−.004	−.006	−.005
										(.012)	(.014)	(.009)
T_{low} (°C)										−.011	.009	−.001
										(.013)	(.013)	(.009)
rain (mm)										.004	−.011	3e-05
										(.006)	(.010)	(.005)
snow (cm)										−.010	−.037	−.021
										(.035)	(.041)	(.027)
log(HH inc)				**.67***	**.72***	**.68***	**.35***	**.56***	**.41***	**.42***	**.61***	**.47***
				(.13)	(.20)	(.11)	(.12)	(.20)	(.10)	(.13)	(.21)	(.11)
health							**2.95***	**1.53***	**2.44***	**2.99***	**1.58***	**2.47***
							(.17)	(.23)	(.13)	(.17)	(.23)	(.14)
trust-N							**.62***	***.48***	**.56***			
							(.20)	(.23)	(.15)			
controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
mnthStn f.e.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
clustering	mnthStn	mnthStn	mnthStn	mnthStn	mnthStn	mnthStn	mnthStn	mnthStn	mnthStn	mnthStn	mnthStn	mnthStn
survey	G19	E2	⟨2⟩	G19	E2	⟨2⟩	G19	E2	⟨2⟩	G19	E2	⟨2⟩
obs.	5144	1245	6389	4040	1122	5162	4017	1122	5139	3833	1122	4955
pseudo-R^2	.027	.033		.033	.036		.073	.045		.074	.044	
$N_{clusters}$	169	44		152	42		150	42		143	42	

Table 4.4: Weather and satisfaction with life, controlling for local climate. Significance: **1%*** ***5%*** *10%**

4.3.2 Weather and other determinants of well-being

Ascertaining a large effect of purely exogenous weather shocks on SWL does not directly elucidate the mechanism of influence. Two possible channels are (1) a sun-associated shift towards optimism when conducting the life satisfaction assessment and (2) a weather-mediated effect on time use over the week preceeding the interview. For instance, sunny weather may be conducive to socialising with family, friends, or community out of the home or pursuing other rewarding activities, in particular those that are outdoors or require outdoor travel. Recent enjoyment of such weather-modulated activities may promote the salience of the respondent's social connectedness or access to chosen leisure activities.

The subsequent two tables may shed some preliminary light on these possibilities. Firstly, the first four columns of Table 4.5 contain the surprising result that when a conventional measure of affect, or mood, is substituted in place of the more cognitive and reflective SWL, the influence from weather nearly disappears. The coefficients come from the GSS19 survey which asked the question "Presently, would you describe yourself as: very happy, somewhat happy, somewhat unhappy, or very unhappy?" to all respondents (stating "no opinion" was also an option). ESC2 had no similar question about mood. When complete controls for climate and other geographic effects are included,[11] the estimated effect of recent and current cloudiness on self-reported happiness is not statistically distinguishible from zero. There is the weak suggestion that cooler nighttime temperatures promote higher happiness, and it is also worthy of note that self-reported health is almost as strongly related to short-term happiness as to the longer-term report of SWL.

The compressed, four-point scale of the happiness question can be expected to elicit numerically smaller marginal effects than the ten-point SWL question, simply on the basis of its coarser resolution. Thus, comparable effects from recent cloudiness cannot be altogether statistically ruled out by the results of Table 4.5, but they nevertheless strongly suggest that the first postulated channel described above, in which cloudiness affects mood which in turn affects the calculation of SWL, is not a good description. One way to check this implication is to convert SWL into a comparable four-point scale to see whether the reduced resolution itself is to blame for the insignificant coefficients. This is carried out in Table 4.6. The ten-point responses given in GSS19 for SWL are mapped into four points in order to match as closely as possible the distribution of the happiness response. The result is clearly no decrease in the significance of the effect, confirming the surprising result that the SWL question is more sensitive than happiness to the influence of transient weather.

While self-reported health is a strong predictor of both SWL and happiness, like happiness it does not appear to be significantly driven by the degree of recent cloudiness nor by daily temperatures. Columns (5) – (8) of Table 4.5 show means of coefficients from both surveys with health as the dependent variable and with local climate fixed effects fully accounted for. These are extracted from the more detailed set of estimates which include regressions without

---

SWL score — ie $\frac{1}{3}$ and $\frac{2}{3}$, respectively.

[11] Once again, the more complete set of tests carried out can be seen in Table C.1 on page 173.

	happy (1)	happy (2)	happy (3)	happy (4)	health (5)	health (6)	health (7)	health (8)	trust-N (9)	trust-N (10)	trust-N (11)	trust-N (12)	trust-G (13)	trust-G (14)	trust-G (15)	trust-G (16)	log(HH inc) (17)	log(HH inc) (18)	log(HH inc) (19)
clouds				−.11				.013				.15				.068			*−.049**
				(.17)				(.15)				(.18)				(.25)			(.027)
clouds (7 days)	−.28	−.31	−.27	−.15	−.16	−.095	−.088	−.015	***−.73***	−.35	−.42	−.66	.24	.62	.74	.72	*−.082**	−.074	*−.087**
	(.38)	(.43)	(.43)	(.48)	(.26)	(.28)	(.28)	(.31)	(.34)	(.38)	(.39)	(.41)	(.48)	(.50)	(.52)	(.56)	(.047)	(.047)	(.049)
T_{high} (°C)				.016				−.005				.013				−.020			.002
				(.015)				(.009)				(.012)				(.016)			(.001)
T_{low} (°C)				***−.040***				−.005				−.009				−.009			−.001
				(.020)				(.011)				(.013)				(.018)			(.001)
rain (mm)				.005				.006				−.005				*−.014**			−.0008
				(.007)				(.004)				(.006)				(.008)			(.0009)
snow (cm)				−.041				**.053***				.045				−.017			**−.010***
				(.029)				(.014)				(.030)				(.034)			(.003)
log(HH inc)		**.63***	***.35***	***.31***		**.78***	**.75***	**.82***		**.76***	**.71***	**.75***		**1.02***	**.75***	**.98***			
		(.14)	(.14)	(.16)		(.10)	(.10)	(.11)		(.15)	(.15)	(.14)		(.15)	(.16)	(.16)			
health			**2.63***	**2.67***							**.78***	**.77***			**.90***	**1.20***		**.17***	**.19***
			(.19)	(.19)							(.18)	(.18)			(.23)	(.25)		(.023)	(.023)
trust-N			.37				**.49***								**1.70***			**.11***	
			(.23)				(.11)								(.17)			(.019)	
controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
mnthStn f.e./clust	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
survey	G19	G19	G19	G19	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩
obs.	5169	4052	4029	3846	6447	5195	5195	5009	3390	2683	2682	2603	3753	3067	3059	2967	5350	5335	5141

Table 4.5: Weather and other covariates of satisfaction with life. Mean coefficients, calculated as weighted averages over estimates carried out separately for each available survey, are shown. Significance: **1%*** ***5%*** *10%**

the fixed effects.

Corresponding findings for weather effects on two measures of trust and on self-reported household income are also summarised in Table 4.5. Because income is a continuous variable, an ordinary least-squares (OLS) model is used in the final three columns. Only weighted averages from the two surveys are displayed in the table. The appendix shows that in general the effect of precipitation is not consistent between the two surveys, while those of temperature and cloudiness are. Trust in neighbours is negatively but marginally dependent on recent cloudiness while reported income is negatively — but more significantly for GSS19 than for ESC2 — associated with snowfall. Because only half of the GSS19 respondents were asked trust questions, the sample sizes are smaller for these than for other questions.

The possibility that some of the major self-reported covariates of life satisfaction are also strongly affected by weather conditions is important. If spurious influences on mood can be shown simultaneously to affect both satisfaction with life and the "right hand side" variables typically portrayed as causative, the consistency of estimates in individual level regressions for life satisfaction could be put gravely in doubt. Correlations between SWL and trust and even between SWL and self-reported income that are due to separate but simultaneous influence from transient factors like weather may be indistinguishable from correlations that are due to a causal channel running only through more long-term effects. This amounts to the central critique made by *Bertrand and Mullainathan* [2001] and is also the classic endogeneity problem.

To lay out some possibilities explicitly for the three-way relationship between weather, SWL, and other subjective measures like trust, consider the following causal relationships corresponding to the case of spurious correlation:

$$
\begin{array}{ccccc}
 & & & & \text{trust} \\
 & & & \nearrow & \\
\text{weather} & \longrightarrow & \text{mood, judgement} & & \\
 & & & \searrow & \\
 & & & & \text{life satisfaction}
\end{array}
$$

There need be no effect at all of trust on life satisfaction in order to observe a statistical correlation between the two. In this case weather conditions influence an individual's assessment of others' trustworthiness through some affective bias in judgement. For instance, sunny weather may generate a good mood and good moods may tend to promote the salience of positive rather than negative attributes of remembered experience. Parallel biases may then influence responses to the trust question and the SWL question.

Another possibility is that the relationship between trust and life satisfaction is more or less causal in the way generally portrayed in the social capital and well-being literature, and that weather is correlated with SWL largely through its influence on the measured and well-recognized principal determinants of SWL, such as trust:

	SWL (4-point) (1)	SWL (4-point) (2)	SWL (4-point) (3)	SWL (4-point) (4)
clouds				***−.52***
				(.26)
clouds (7 days)	***−.91***	**−1.25***	**−1.22***	**−1.22***
	(.39)	(.39)	(.44)	(.46)
T_{high} (°C)				−.002
				(.017)
T_{low} (°C)				−.026
				(.020)
rain (mm)				.003
				(.008)
snow (cm)				−.046
				(.038)
log(HH inc)		**.59***	*.28**	*.28**
		(.16)	(.15)	(.16)
health			**2.97***	**2.96***
			(.22)	(.22)
trust-N			***.47***	
			(.22)	
controls	✓	✓	✓	✓
mnthStn f.e.	✓	✓	✓	✓
clustering	mnthStn	mnthStn	mnthStn	mnthStn
survey	G19	G19	G19	G19
obs.	5144	4040	4017	3833
pseudo-R^2	.055	.065	.124	.126
$N_{clusters}$	169	152	150	143

Table 4.6: Weather and a compressed measure of life satisfaction. The dependent variable is the 10-point satisfaction with life response compressed into four categories for better comparability with happiness in GSS19. Significance: **1%*** ***5%*** *10%**

$$
\begin{array}{ccccccc}
 & & \text{mood, judgement} & \longrightarrow & \text{trust} & & \\
 & \nearrow & & & & \searrow & \\
\text{weather} & & & & & & \text{life satisfaction} \\
 & \searrow & & & & \nearrow & \\
 & & \text{activities, encounters} & \longrightarrow & \text{trust} & &
\end{array}
$$

Two examples are shown of how this influence on trust could come about. The top one works through the same judgement bias channel discussed above, while the bottom is that described previously in which recent activities that are influenced by weather may change the salience or freshness of memories, in this case relating specifically to the familiarity and trustworthiness of neighbours or others. In each of these two interpretations, short-term weather conditions act like a natural experiment in which the independent variable, trust, is modulated randomly around its longer average without directly affecting SWL. Under this assumption the importance of trust in determining SWL could be correctly estimated by using the projection of reported trust onto current weather conditions in a two-stage regression for SWL. The randomness of recent weather, controlling for climatic norms, would eliminate other endogenous factors linking trust and SWL. However, given that weather is highly correlated with SWL even after trust and other subjective responses are controlled for suggests that weather is not a reasonable instrument for trust when predicting SWL.

The lack of an effect of weather on happiness may be an argument against the mood-mediated channels, while the significant coefficient on weather in explaining SWL even when trust, health, and income are controlled for (column ⟨59-60⟩ in Table 4.4) suggests that the introspective judgement leading to SWL responses is being affected by weather in some other way.

In order to test for the validity of standard inferences about the subjective (health and trust) and ostensibly objective (income) determinants of SWL in the presence of an influence on mood and judgement, Table 4.7 compares regression results with and without controls for weather. Columns 1, 4, 7, and 10 control for current weather conditions. The subsequent columns to each of these — 2, 5, 8, and 11 — estimate a version of the equation which is naïve to weather but uses precisely the same sample as the first specification. The remaining columns estimate the naïve equation using the entire available sample — that is, including samples which are missing one of the weather condition variables and therefore excluded in the earlier estimates. In all cases, fixed effects are included for every combination of month and geography.

Reassuringly, despite the significant influence already shown of weather on both SWL and some of its explanatory variables, the inclusion and exclusion of weather conditions result in indistinguishible coefficients on each of those explanatory variables.

4.3.3 Climate and well-being

The foregoing analysis addresses the question of how much is missing when a transient influence like weather is absent from an empirical model for SWL. I now turn to the analogous question

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
clouds	***−.28***			***−.31***			*−.25**			***−.30***		
	(.14)			(.13)			(.13)			(.14)		
clouds (7 days)	***−.64***	***−.68***		−.45	−.47		−.44	−.46		***−.61***	***−.64***	
	(.30)	(.29)		(.29)	(.29)		(.29)	(.29)		(.31)	(.31)	
T_{high} (°C)	−.005			−.007			−.007			−.005		
	(.009)			(.008)			(.008)			(.009)		
T_{low} (°C)	−.003			−.003			−.001			−.0005		
	(.009)			(.009)			(.009)			(.009)		
rain (mm)	.003			−.0004			−.002			−.002		
	(.004)			(.004)			(.005)			(.005)		
snow (cm)	−.022			*−.040**			−.035			−.024		
	(.025)			(.022)			(.021)			(.026)		
log(HH inc)	**.71***	**.72***	**.72***							**.43***	**.41***	**.44***
	(.11)	(.11)	(.11)							(.11)	(.10)	(.11)
trust-N				**.80***	**.76***	**.79***				**.61***	**.56***	**.59***
				(.13)	(.13)	(.13)				(.15)	(.15)	(.15)
health							**2.62***	**2.64***	**2.63***	**2.43***	**2.44***	**2.42***
							(.13)	(.12)	(.13)	(.14)	(.13)	(.14)
controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
clustering	mnthStn	mnthStn	mnthStn	mnthStn	mnthStn	mnthStn	mnthStn	mnthStn	mnthStn	mnthStn	mnthStn	mnthStn
survey	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩
obs.	4978	4978	4978	6160	6389	6160	6146	6375	6146	4955	5139	4955

Significance: **1%*** ***5%*** *10%**

Table 4.7: Comparison between naïve and weather-aware models of SWL. Raw ordered logit coefficients and standard errors are shown.

regarding climate. When geographic or seasonal differences in climate are ignored across a sample population, one might expect significant variation in SWL to go unexplained due to this missing variable. In sections 4.3.1 and 4.3.2 these differences have been controlled for using fixed effects for month, location, or the combination of the two in order to focus on the relatively unexpected, short-term component of weather. In place of these all-encompassing climate fixed effects, I now use some measures of long-term climate averages available from Environment Canada to investigate climate as an amenity. Such efforts have also been made for Russia and Ireland by *Frijters and Van Praag* [1998] and *Brereton et al.* [2008].

Table 4.8 summarises the results, presented in more detail in the Appendix.[12] Climate parameters are grouped into three categories: those that describe annual, monthly, and daily averages at each weather station. The first column of the table shows an ordered logit estimate for SWL which includes month fixed effects, the standard suite of socioeconomic controls along with household income, and three measures of annual average climate. These are the average maximum temperature of the warmest month, the average minimum of the coldest month, and the average number of days of sunshine per year. The second and third columns bring in local monthly averages and local daily averages for each station, including the probability of receiving more than 5 mm of precipitation and the average amount of precipitation received. Because these climate measures are not available for all stations, sample sizes are relatively small.

Generally, the climatic variables do not appear to have a significant effect on SWL once the season and demographic controls are accounted for.[13] The next three columns show the same specifications with the omission of household income, in order to test for the possibility that people with greater financial means of choosing their location are more likely to experience a favourable climate. This turns out not to be the case. Columns (7) to (10) repeat the specifications allowing for a fixed effect for each weather station rather than for each month. Thus the month-level climate averages now represent climate features that are special for the interview month at a given location rather than those that are special to the location for a given month.

The estimates shown in the remaining three columns of Table 4.8 include the detailed set of controls for local and seasonal climate. Once again, expectations for the day's weather do not appear to play significantly into SWL responses yet — as shown in the final column — the actual cloudiness experienced has a very significant impact on SWL.

4.3.4 Cyclic temporal effects

The date of the interview itself represents another possible contextual effect that is usually ignored in large survey analysis. *Csikszentmihalyi and Hunter* [2003] use an experience sampling method to investigate the correlates of reported momentary happiness. For their sample of teenagers, significant though slight differences in happiness were found as a function of time of day and the day of the week, with times free of school constraints being favoured. To check

[12] See Table C.2 on page 177.

[13] The significant coefficients on precipitation-related variables only occur when collinear variables are present.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
YEAR: $\langle T_{max} \rangle$ (°C)	*.068**	.052	***.085***	***.063***	*.070**	***.069***							
	(.036)	(.044)	(.041)	(.032)	(.040)	(.035)							
YEAR: $\langle T_{min} \rangle$ (°C)	−.011	−.019	−.011	−.011	−.021	−.008							
	(.012)	(.016)	(.018)	(.010)	(.016)	(.018)							
YEAR: days sun	***−.004***	−.002	−.002	−.004	−.003	−.004							
	(.002)	(.003)	(.004)	(.003)	(.003)	(.004)							
MONTH: days sun		.053			.075		−.020		−.007				
		(.11)			(.094)		(.032)		(.026)				
MONTH: sun fraction		−.010			−.017		.003		.007				
		(.026)			(.023)		(.010)		(.008)				
MONTH: $\langle T \rangle$ (°C)		.029			.031		−.002		−.010				
		(.029)			(.027)		(.012)		(.012)				
MONTH: rain>5mm		.033			.030		.032		.029				
		(.051)			(.047)		(.033)		(.029)				
MONTH: snow>5cm		−.059			.010		−.038		−.018				
		(.10)			(.091)		(.055)		(.046)				
DAY: precipitation			**.013***			**.012***		.003		.0002	−.005	.007	***−.030***
			(.005)			(.004)		(.003)		(.003)	(.009)	(.007)	(.014)
DAY: $\langle T_{max} \rangle$ (°C)			.060			.046		**−.061***		***−.045***	***−.27***	**−.26***	**−.41***
			(.046)			(.048)		(.021)		(.020)	(.12)	(.095)	(.13)
DAY: $\langle T_{min} \rangle$ (°C)			−.004			−.014		**.069***		***.053***	***.31***	**.31***	**.47***
			(.048)			(.052)		(.024)		(.022)	(.13)	(.11)	(.15)
clouds (7 days)													*−.56**
													(.29)
log(HH inc)	**.57***	**.59***	**.57***				**.54***	**.59***			**.70***		**.69***
	(.14)	(.14)	(.14)				(.100)	(.074)			(.086)		(.11)
controls	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
f.e./clustering	mnth	mnth	mnth	mnth	mnth	mnth	stn	stn	stn	stn	mnthStn	mnthStn	mnthStn
survey	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩	⟨2⟩
obs.	2285	2285	2285	2774	2774	2774	4538	12216	5453	14753	8090	10252	5162

Table 4.8: Climate and satisfaction with life. Covariates include local climatic expectations in the form of probabilities and means for each station's overall climate (YEAR) and for its averages for the month (MONTH) and day (DAY) of the interview. Standard errors are calculated with clustering at the level of the fixed effects (f.e.) indicated. Results in this table are all weighted averages of coefficients determined separately for each of the two surveys; see Table C.2 on page 177 for details. Significance: **1%*** ***5%*** *10%**

whether the social structure of time also affects life satisfaction reported by adults, I estimate the standard SWL equation with fixed effects for the days of the week and for the months of the year. To provide more constrained alternatives, a weekend dummy variable and an annual-cycle sinusoid peaking on summer solstice are also tested.

Tables 4.9 and 4.10 summarise the results. There is no significant pattern throughout the week, but there is a significant seasonal variation, with a sharp mid winter or holiday peak in SWL. Because the ESC2 survey did not span an entire year, it is not possible to corroborate the pattern properly between surveys.

4.4 Conclusions

The perspective underpinning this work is to recognise subjective responses as the result of a cognitive evaluation that is likely to be imperfect yet which contains useful information. In principle there is no alternative to reliance on subjective assessments to evaluate a population's well-being or at least to learn or elucidate the importance of various factors in promoting this ultimate social goal. Since SWL data are characterised by a high degree of variability, both between individuals and for a given individual over time, understanding what influences and biases lie in this variation is an ongoing task. Given the importance of large survey data for modern inference about subjective well-being and its judgement-based explanatory factors, for instance measures of trust that proxy for social capital, being able to quantify or put constraints on psychological bias in survey responses remains an important component of analysis.

I find that after controlling for local climate expectations, an average of recent cloud cover levels has a large and significant effect on SWL responses. The magnitude of the modeled effect of a change in weather circumstances from half-cloudy to completely sunny is comparable to that associated with more than a factor of ten increase in household income, more than a full-spectrum shift in perceived trust in neighbours, and nearly twice the entire benefit of being married as compared with being single. In addition, there is an effect of weather on responses to some of the questions typically used in explaining variation in SWL. In particular, trust in neighbours shows a large effect significant at the 5% level and self-reported income may also be subject to a bias related to current weather conditions.

Nevertheless, the findings in this work do not support the hypothesis that the impact of weather on respondents' reported SWL acts through a broad affective bias which would cause correlated mistakes in explanatory and explained variables. There is no evidence of a strong weather effect on reported happiness, the best available measure of affective state at the time of interview, nor is there any evidence that weather causes a spurious correlation between SWL and standard explanatory variables. Statistical estimates which are not informed about the state of weather produce the same inferences regarding the determinants of SWL as those which take weather's influence into account.

To the extent that this work can be taken to be an applied test of the concerns laid out by

	(9)	(10)	⟨9-10⟩	(11)	(12)	⟨11-12⟩
Monday	−.075	.095	−.012			
	(.11)	(.15)	(.090)			
Tuesday	.038	.082	.050			
	(.096)	(.15)	(.081)			
Wednesday	−.15	−.009	−.095			
	(.10)	(.14)	(.083)			
Thursday	−.084	−.055	−.074			
	(.10)	(.14)	(.082)			
Friday	***−.25***	.31	−.12			
	(.12)	(.22)	(.11)			
Saturday	−.035	−.049	−.040			
	(.13)	(.17)	(.10)			
weekend				.082	−.074	.019
				(.077)	(.094)	(.060)
log(HH inc)	**.71***	**.52***	**.65***	**.71***	**.52***	**.65***
	(.10)	(.16)	(.087)	(.10)	(.16)	(.086)
trust-N	**.86***	**.59***	**.73***	**.87***	**.58***	**.73***
	(.15)	(.15)	(.11)	(.15)	(.15)	(.10)
controls	✓	✓	✓	✓	✓	✓
mnthStn f.e.	✓	✓	✓	✓	✓	✓
clustering	mnthStn	mnthStn	mnthStn	mnthStn	mnthStn	mnthStn
survey	G19	E2	⟨2⟩	G19	E2	⟨2⟩
obs.	6309	1780	8089	6309	1780	8089
pseudo-R^2	.037	.033		.036	.032	
$N_{clusters}$	254	62		254	62	

Table 4.9: Days of the week and satisfaction with life. Significance: **1%*** ***5%*** *10%**

	(5)	(6)	⟨5-6⟩	(7)	(8)	⟨7-8⟩
February	*−.37**	***−.45***	**−.41***			
	(.20)	(.21)	(.15)			
March	***−.34***	***−.48***	**−.38***			
	(.17)	(.24)	(.14)			
April	***−.43***	−.39	**−.41***			
	(.21)	(.25)	(.16)			
May	*−.34**	*−.44**	**−.37***			
	(.18)	(.24)	(.14)			
June	**−.53***	−.31	**−.47***			
	(.15)	(.27)	(.13)			
July	***−.36***	−.30	***−.34****			
	(.18)	(.32)	(.16)			
August	−.20	−.15	−.19			
	(.17)	(.30)	(.15)			
September	**−.45***		**−.45***			
	(.17)		(.17)			
October	***−.38***		***−.38****			
	(.17)		(.17)			
November	−.24		−.24			
	(.17)		(.17)			
December	−.28		−.28			
	(.20)		(.20)			
sun cycle				*−.048**	.052	−.037
				(.028)	(.081)	(.026)
log(HH inc)	**.59***	**.47***	**.55***	**.59***	**.47***	**.55***
	(.090)	(.12)	(.072)	(.089)	(.12)	(.071)
trust-N	**.84***	**.57***	**.71***	**.84***	**.57***	**.72***
	(.12)	(.13)	(.089)	(.12)	(.13)	(.088)
controls	✓	✓	✓	✓	✓	✓
stn f.e.	✓	✓	✓	✓	✓	✓
clustering	stn	stn	stn	stn	stn	stn
survey	G19	E2	⟨2⟩	G19	E2	⟨2⟩
obs.	9710	2561	12271	9710	2561	12271
pseudo-R^2	.028	.037		.027	.037	
$N_{clusters}$	137	49		137	49	

Table 4.10: Calendar months and satisfaction with life. Significance: **1%*** ***5%*** *10%**

Bertrand and Mullainathan [2001], their objections appear to be pessimistic in that they do not gain support in the expected way. At least for the case of weather and SWL, it appears that the effects of transient influences can be significant yet not overwhelm the underlying relationships evident through large statistical inferences.

The lack of a strong correlation between reported happiness and the aspects of weather which influence SWL and other subjective variables is surprising and remains mysterious. On the other hand there is a plausible explanation for the positive effect of sunniness on SWL and trust in neighbours. The influence could come as a result of modified behaviour, for instance the promotion of outdoor activity or social gathering, rather than directly from sunlight. Tests of this hypothesis will be carried out in future work.

Bibliography for Chapter 4

Bayer, P., and R. McMillan, Racial Sorting and Neighborhood Quality, *Working Paper 11813*, National Bureau of Economic Research, 2005.

Bertrand, M., and S. Mullainathan, Do People Mean What They Say? Implications for Subjective Survey Data, *The American Economic Review*, *91*, 67–72, 2001.

Brereton, F., J. Clinch, and S. Ferreira, Happiness, geography and the environment, *Ecological Economics*, *65*, 386–396, 2008.

Cao, M., and J. Wei, Stock market returns: A note on temperature anomaly, *Journal of Banking and Finance*, *29*, 1559–1573, 2005.

Csikszentmihalyi, M., and J. Hunter, Happiness in Everyday Life: The Uses of Experience Sampling, *Journal of Happiness Studies*, *4*, 185–199, 2003.

Diener, E., Subjective well-being. The science of happiness and a proposal for a national index., *Am Psychol*, *55*, 34–43, 2000.

Frijters, P., and B. Van Praag, The Effects of Climate on Welfare and Well-Being in Russia, *Climatic Change*, *39*, 61–81, 1998.

Guven, C., Reversing the Question. Does Happiness Affect Individual Economic Behavior? Evidence from Survey Data from Netherlands INCOMPLETE AND PRELIMINARY DRAFT, 2007.

Gyourko, J., M. Kahn, and J. Tracy, *Quality of life and environmental comparisons*, vol. 3 of *Handbook of Regional and Urban Economics*, chap. 37, pp. 1413–1454, Elsevier, 1999.

Helliwell, J., Well-Being and Social Capital: Does Suicide Pose a Puzzle?, *Social Indicators Research*, *81*, 455–496, 2007.

Hirshleifer, D., and T. Shumway, Good Day Sunshine: Stock Returns and the Weather, *Journal of Finance*, *58*, 2003.

Krämer, W., and R. Runde, Stocks and the weather: An exercise in data mining or yet another capital market anomaly?, *Empirical Economics*, *22*, 637–641, 1997.

Krueger, A. B., and D. A. Schkade, The reliability of subjective well-being measures, *Journal of Public Economics*, *92*, 1833–1845, 2008.

Loughran, T., and P. Schultz, Weather, stock returns, and the impact of localized trading behavior, *Journal of Financial and Quantitative Analysis*, *39*, 343–64, 2004.

Magnusson, A., An overview of epidemiological studies on seasonal affective disorder, *Acta Psychiatrica Scandinavica*, *101*, 176–184, 2000.

Mersch, P., H. Middendorp, A. Bouhuys, D. Beersma, and R. van den Hoofdakker, Seasonal affective disorder and latitude: a review of the literature., *J Affect Disord*, *53*, 35–48, 1999.

Moro, M., F. Brereton, S. Ferreira, and J. P. Clinch, Ranking quality of life using subjective well-being data, *Ecological Economics,*, *65*, 448–460, 2008.

Rappaport, J., Title Moving to High Quality of Life, *CEA 41st Annual Meetings*, 2007.

Rehdanz, K., and D. Maddison, Climate and happiness, *Ecological Economics*, *52*, 111–125, 2005.

Schimmack, U., E. Diener, and S. Oishi, Life-satisfaction is a momentary judgment and a stable personality characteristic: The use of chronically accessible and stable sources, *Journal of personality*, *70*, 345–384, 2002.

Schkade, D., and D. Kahneman, Does living in California make people happy?: A focusing illusion in judgments of life satisfaction, *Psychological Science*, *9*, 340–46, 1998.

Schwarz, N., and G. Clore, Mood, misattribution, and judgments of well-being: informative and directive functions of affective states, *Journal of personality and social psychology*, *45*, 513–523, 1983.

Schwarz, N., and F. Strack, Evaluating one's life: a judgment model of subjective well-being, in *Subjective well-being: an interdisciplinary perspective*, edited by F. Strack, M. Argyle, and N. Schwarz, Pergamon Press, Oxford [England]; New York, 1991.

Soroka, S., R. Johnston, and J. F. Helliwell, *Diversity, Social Capital and the Welfare State*, chap. Measuring and Modelling Interpersonal Trust, UBC Press, 2007.

Stutzer, A., and B. S. Frey, Stress That Doesn't Pay Off: The Commuting Paradox, *SSRN eLibrary*, 2004.

van Praag, B., and B. Baarsma, Using Happiness Surveys to Value Intangibles: The Case of Airport Noise*, *The Economic Journal*, *115*, 224–246, 2005.

Welsch, H., Environmental welfare analysis: A life satisfaction approach, *Ecological Economics*, 2006.

Yuan, K., L. Zheng, and Q. Zhu, Are investors moonstruck? Lunar phases and stock returns, *Journal of Empirical Finance*, *13*, 1–23, 2006.

Chapter 5

Conclusions and further work

In this dissertation I have tackled difficulties that arise in empirical and theoretical work when certain common, simplifying assumptions of microeconomics are relaxed. Along with a growing volume of very recent literature, it represents a modern continuation of the ideas of economists like Thorstein Veblen and, earlier, Karl Marx and Adam Smith, who all realised that basic motivations, even those for wealth and material gain, were primarily social in nature. While this fact has been actively forgotten in economics textbooks of the last half century, its importance in modern society is no longer ignorable. In all sectors of policy making and academic pursuit and throughout public discourse, attention has turned towards the social and economic imperatives presented to us by massive overconsumption, at least of carbon-releasing fuels, since the first industrial revolution. In order to contribute to guiding our way towards a more sustainable form of prosperity, economics must restore some of its emphasis on social externalities which tend to dominate over private returns.

Fortuitously, two of these revisions may be synergistic for informing policy. The unregulated emission of greenhouse gases and the general lack of accounting of actions which degrade Earth's natural capital have put those systems into a state of general collapse. Were this all in the name of individual absolute material consumption benefits of current and past generations, the problem might look very dire without the presence of future generations to stake their claim. However, if the problem is characterised by a second major externality from consumption, as described in this dissertation, in which a rat-race driven by Veblen preferences tends to expand production efforts without net welfare benefits, then a new industrial revolution aimed at shifting the pattern of consumption may not represent a traumatic tradeoff in well-being.

Instead, a further externality of the positive kind is becoming clear from analysis of life satisfaction data [*Helliwell and Putnam*, 2004]. If the true determinants of human experienced utility are dominated by the quality of inherently mutual social interactions, then there is hope indeed for a happier outcome for societies which succeed in realigning their social metrics and policy priorities. If this is a fair description, then economics must adapt quickly to stay both relevant and ethical in light of the evidence it faces.

It is worth pointing out two aspects of utility treated in the present work. In contemporary economics, the wholly separate concepts of decision-making utility and experienced utility may easily be confounded. The latter corresponds to well-being, that thing by which we wish to judge the success of our policies and predict the comparative effects of alternative scenarios in our theories. The former is a theoretical device used to encapsulate rational behaviour succinctly, under the assumption that the postulates of rational behaviour are justified.

In Chapter 2 I set aside this question of decision making to look at *ex post* outcomes in terms of empirical well-being. Chapter 3, on the other hand, represents primarily a thought experiment in what decision making utility dominated by Veblen comparisons means for group formation and sorting. In discussing the social desirability of outcomes, however, I have used the same functional form of utility as I used in describing decision making.

There is the implicit implication in these two chapters (explicit in the welfare analysis of Chapter 3) that the two species of utility — decision making and experienced — are closely related, just as they are silently and nonfalsifiably assumed to be in much of classical economics. The extent to which people do tend to make choices which maximise their own life satisfaction given the independence of the choices of others is a separate, empirical, and tractable question of major importance. If our (material) consumption-oriented society is driven not only by huge consumption externalities of the kind analysed here, but in addition by a self-perpetuating delusion about whether individual consumption is likely to improve our lot, then the reasons for such a systematic bias must also be found. Obvious places to look for that are the advertising industry [*Bertrand et al.*, 2006], which now forms one of the largest components of capitalist economies, and the ideas that trickle out of economics acadaemia itself.

The three works comprising this dissertation provide complementary contributions in three senses. Firstly, they offer both empirical and theoretical insights to the topic of consumption externalities. Chapter 3 provides a theoretical framework with which to think about the endogeneity of local reference groups, due to the mobility of households, which might play a role in the findings of Chapter 2. Secondly, I provide a confident interpretation of life satisfaction data to address an important policy question (Chapter 2), on the one hand, and a skeptical assessment of the reliability of life satisfaction data in general for any econometric inference (Chapter 4), on the other. Thirdly, I address both the behavioural implications (Chapter 3) of consumption emulation, assuming sophisticated agents with fully rational utility functions, and the welfare implications of world in which we gain consumption benefits by comparison, regardless of our rationality, decision making algorithms, or susceptibility to marketing influences (Chapter 2).

Some of the policy implications of work focusing on relativities in consumption benefits are undoubtedly difficult, or unsettling, at this point. While Chapter 4 suggests that fears about the unreliability of subjective data may continue to dissolve as more experience builds confidence, Chapters 2 and 3 highlight the difficulty of robust inference using a noisy empirical measure like life satisfaction or complex, interdependent utility functions in economic models. To anyone convinced that one must not pass up the opportunity to measure well-being by a means more direct than the often-tautologous rhetoric of "revealed preference," one clear policy implication of work such as the present one is to do more measuring of the things which appear to be important. Life satisfaction is top among them, followed by those factors it may reveal to be most important, including measures of health, opportunity for self-determination, and positive social engagement. The resulting world may not look very different from the current structure of society, in most ways, but in any engineering or design problem one must start with the right objective function. After that, heuristics and trial and error will be needed to make progress, but each may always be justified or rejected by the available metric.

The present work has focused on well-being within one country, Canada, and particularly on residents of urban areas. A major aspect of the larger picture of our understanding of the determinants of well-being and in particular of the role of own and others' income lies beyond this scope. Much current attention is focused on international differences in well-being, and an all-important debate on the importance of income is alive and well [*e.g., Easterlin, 1995; Stevenson and Wolfers, 2008a; Easterlin, 2008]*. Comparing respondents between countries has the advantage that the benefit of all federally-funded public goods is included in average effects. This is lacking in the within-country analysis I have conducted, and as a result any objection about the sum of income coefficients estimated in Chapter 2 being negative overall is misguided. The federal public benefits that likely contribute a great deal to Canadian well-being are likely, could they be measured, to tip the balance in favour of higher incomes benefiting well-being overall. Between countries, of course, the variation is as much in *how* federal funds are spent as in the economic cost.

One drawback of cross-country well-being analysis, on the other hand, is that samples within any country are small and thus the geographic structure of local social effects treated in this dissertation is hard to take into account. The likely and necessary future trend of the field, then, is clear, and coincides with one of the easiest policy implications of the work at this stage. Larger survey samples are needed within countries and across all countries. In this regard, recent initiatives by Gallup Corporation are an exciting sign. Gallup now conducts an annual survey in over 130 countries and a daily survey in the United States. Both of these contain questions on well-being and are presently objects of intensive study.

One further complication relating to geographic scales of consumption reference groups is that their scope may be changing rapidly. Further research is needed to determine whether influences like national television or global media, entertainment, and electronic networks are bringing everyone closer to having a global perspective on what level of affluence they should consider normal. Naturally, such "normality" will be defined by whatever group broadcasts the loudest. If people the world over are comparing themselves increasingly to one standard, then variation in reference groups will be less available for studies like this one which aim to infer their effects.

I conclude by pointing out a few of the many directions indicated for future research.

- By focusing on income and consumption in this dissertation, I have set aside what appear to be the most important measures for empirical experienced utility. To further the broader research programme, similar multi-level geographic contextual effects can be probed for measures of social capital and other determinants of well-being, thus emphasising the geography of empathy rather than that of emulation.

- Although in Canada the ten-point life satisfaction question has not been used in any large panel surveys, there now exist some newer cross-sectional surveys which present the opportunity of simple geographic time series analysis of factors affecting social connectedness and life satisfaction. GSS Cycle 22 will duplicate the format and many of the questions from GSS17 and will thus be especially suitable. In addition there are some large Canadian health surveys which include fine resolution geographic identifiers and a question on general satisfaction. These will provide a much larger sample and therefore may allow better city-level and community-level comparisons, in addition to further validation tests on the patterns described in this dissertation.

- The theoretical work of Chapter 3 leaves some questions unanswered. In particular, finding a functional form of utility which admits algebraic equilibrium solutions for both the single-neighbourhood and the differentiated neighbourhood cases would make possible a more pleasing welfare analysis. Furthermore, the work would relate better to the literature on Tiebout economies with a finite number of neighbourhoods if an analytic solution for N neighbourhoods, or at least for the limiting case of $N \to \infty$, were found.

Bibliography for Chapter 5

Bertrand, M., S. Mullainathan, and E. Shafir, Behavioral Economics and Marketing in Aid of Decision-Making among the Poor, *Journal of Public Policy & Marketing*, *25*, 8–23, 2006.

Easterlin, R. A., Will raising the incomes of all increase the happiness of all?, *Journal of Economic Behavior & Organization*, *27*, 35–47, 1995.

Easterlin, R. A., Lost in Transition: Life Satisfaction on the Road to Capitalism, *SSRN eLibrary*, 2008.

Helliwell, J., and R. Putnam, The social context of well-being., *Philos Trans R Soc Lond B Biol Sci*, *359*, 1435–46, 2004.

Stevenson, B., and J. Wolfers, Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox, 2008a.

Appendix A

Appendix to Chapter 2

A.1 Supplementary tables for urban geography of life satisfaction

Some more detailed tabulated regression results are collected here.

Table A.1: Detailed regressions for baseline model and subpopulation estimates. These results are summarised in Table 2.3 on page 14.
Significance: 1% 5% 10%

	log(HH inc)	DA: log(HH inc)	CT: log(HH inc)	CSD: log(HH inc)	CMA: log(HH inc)	$\sum\beta_{\text{inc}}$	health	trust-N	married	asmarried	separated	divorced	widowed	male	noReligion	godImportance	student	employed	domestic	unemployed	retired	age	$(\text{age}/100)^2$	controls	f.e./clustering	$\tau_{\text{neigh}}\geq 10$yr	$\tau_{\text{city}}\geq 10$yr	foreign born	own house	survey	obs.	pseudo-R^2	N_{clusters}
(1)	.26	.093	.14	−1.46	−.56	−1.53	1.04	.49	.63	.37		−.24	.10	−.018		.36				−1.17		−.030	4.94	✓						E2	1141	.038	
	(.20)	(.50)	(.61)	(.99)	(1.21)	(1.18)	(.32)	(.17)	(.17)	(.21)		(.24)	(.32)	(.11)		(.18)				(.31)		(.021)	(2.06)										
(2)	.23	.14	−.33	.11	−1.13	−.98		1.73	.45	.46				−.078	.009	.56	1.27	1.20	1.07		1.41	−.059	5.95	✓						ED	23589	.054	
	(.048)	(.20)	(.24)	(.32)	(.38)	(.41)		(.097)	(.061)	(.092)				(.041)	(.064)	(.068)	(.14)	(.13)	(.15)		(.15)	(.009)	(.97)										
(3)	.32	.41	−.58	−.43	−1.08	−1.36	2.74	1.05	.44	.35	−.45	−.098	−.029	−.16	−.21	.38	.65	.58	.69	−.19	.84	−.091	8.99	✓						G17	12201	.064	
	(.059)	(.18)	(.22)	(.30)	(.37)	(.40)	(.095)	(.087)	(.068)	(.077)	(.10)	(.089)	(.13)	(.040)	(.048)	(.061)	(.18)	(.17)	(.18)	(.22)	(.18)	(.009)	(.95)										
⟨1-3⟩	.27	.28	−.43	−.24	−1.08	−1.19	2.61	1.25	.45	.40	−.45	−.11	−.009	−.11	−.13	.45	1.03	.98	.92	−.51	1.17	−.071	7.26	✓						⟨3⟩	36931		
	(.037)	(.13)	(.16)	(.22)	(.26)	(.28)	(.091)	(.061)	(.044)	(.057)	(.10)	(.084)	(.12)	(.028)	(.038)	(.044)	(.11)	(.10)	(.12)	(.18)	(.12)	(.006)	(.65)										
(4)	.29	.24	.031	−.72		−.15	.87	.40	.69	.27		−.15	.084	−.11		.44				−.99		−.027	4.37	✓	CMA					E2	1035	.041	23
	(.18)	(.42)	(.57)	(1.48)		(100.0)	(.32)	(.13)	(.21)	(.27)		(.43)	(.23)	(.079)		(.14)				(.39)		(.016)	(1.21)										
(5)	.23	.15	−.33	.21		.26		1.80	.47	.42				−.082	.083	.59	1.26	1.20	1.08		1.40	−.061	6.20	✓	CMA					ED	23589	.058	42
	(.043)	(.14)	(.20)	(.24)		(.21)		(.088)	(.057)	(.068)				(.040)	(.072)	(.065)	(.15)	(.14)	(.16)		(.13)	(.009)	(.92)										
(6)	.31	.41	−.60	−.36		−.25	2.75	1.05	.45	.34	−.44	−.086	−.037	−.16	−.18	.40	.64	.57	.69	−.21	.84	−.093	9.12	✓	CMA					G17	12201	.066	46
	(.051)	(.13)	(.16)	(.21)		(100.0)	(.10)	(.13)	(.047)	(.059)	(.079)	(.065)	(.13)	(.036)	(.036)	(.072)	(.17)	(.16)	(.15)	(.24)	(.14)	(.006)	(.66)										
⟨4-6⟩	.26	.28	−.47	−.12		.26	2.59	1.27	.46	.37	−.44	−.087	−.008	−.12	−.13	.50	.99	.93	.88	−.42	1.13	−.078	7.50	✓	CMA					⟨3⟩	36825		
	(.032)	(.093)	(.12)	(.16)		(.21)	(.096)	(.063)	(.036)	(.044)	(.079)	(.064)	(.11)	(.025)	(.032)	(.046)	(.12)	(.11)	(.11)	(.20)	(.095)	(.005)	(.49)										
(7)	.41	.81	−.91			.31	.76	.47	.77	.47		−.17	.49	−.11		.45				−.94		−.035	5.30	✓	CSD					E2	806	.050	23
	(.20)	(.54)	(.54)			(.31)	(.46)	(.14)	(.22)	(.40)		(.38)	(.25)	(.12)		(.18)				(.34)		(.020)	(1.86)										
(8)	.25	.13	−.29			.093		1.84	.46	.45				−.076	.11	.62	1.24	1.18	1.08		1.40	−.065	6.62	✓	CSD					ED	22955	.069	220
	(.050)	(.18)	(.20)			(.067)		(.092)	(.057)	(.091)				(.040)	(.071)	(.062)	(.17)	(.17)	(.18)		(.19)	(.009)	(.99)										
(9)	.30	.42	−.61			.12	2.78	1.09	.42	.31	−.47	−.10	−.072	−.17	−.15	.41	.53	.47	.58	−.29	.71	−.093	9.20	✓	CSD					G17	11429	.070	185
	(.060)	(.17)	(.15)			(.060)	(.11)	(.074)	(.064)	(.081)	(.096)	(.078)	(.14)	(.039)	(.046)	(.058)	(.19)	(.17)	(.17)	(.23)	(.18)	(.009)	(.96)										
⟨7-9⟩	.27	.32	−.51			.11	2.66	1.26	.46	.38	−.47	−.11	.072	−.12	−.074	.51	.93	.84	.82	−.50	1.03	−.075	7.62	✓	CSD					⟨3⟩	35190		
	(.038)	(.12)	(.12)			(.044)	(.11)	(.054)	(.042)	(.060)	(.096)	(.076)	(.12)	(.027)	(.039)	(.041)	(.13)	(.12)	(.13)	(.19)	(.13)	(.006)	(.65)										
(10)																								✓	CT					E2	0		
						(0)																											
(11)	.28	.58				.86		2.22	.73	.45				−.17	.21	.83	1.57	1.38	1.20		1.65	−.070	7.34	✓	CT					ED	8257	.166	747
	(.099)	(.38)				(.31)		(.18)	(.12)	(.20)				(.077)	(.13)	(.14)	(.30)	(.27)	(.29)		(.31)	(.017)	(1.83)										
(12)	.70	−.18				.52	3.29	1.04	.30	.57	−1.07	.048	.026	−.21	−.18	.62	.43	.23	.056	−.24	.089	−.084	7.49	✓	CT					G17	1363	.099	109
	(.21)	(.47)				(.26)	(.37)	(.25)	(.28)	(.28)	(.39)	(.36)	(.51)	(.18)	(.19)	(.20)	(.70)	(.62)	(.65)	(.67)	(.56)	(.035)	(3.82)										
⟨10-12⟩	.35	.27				.66	3.29	1.81	.67	.49	−1.07	.048	.026	−.17	.089	.76	1.39	1.20	1.00	−.24	1.28	−.073	7.37	✓	CT					⟨2⟩	9620		
	(.090)	(.29)				(.20)	(.37)	(.15)	(.11)	(.16)	(.39)	(.36)	(.51)	(.071)	(.11)	(.11)	(.27)	(.25)	(.27)	(.67)	(.27)	(.015)	(1.65)										

Continued on next page

	log(HH inc)	DA: log(HH inc)	CT: log(HH inc)	CSD: log(HH inc)	CMA: log(HH inc)	$\sum \beta_{\text{inc}}$	health	trust-N	married	asmarried	separated	divorced	widowed	male	noReligion	godImportance	student	employed	domestic	unemployed	retired	age	$(\text{age}/100)^2$	controls	f.e./clustering	$\tau_{\text{neigh}} \geq 10\text{yr}$	$\tau_{\text{city}} \geq 10\text{yr}$	foreign born	own house	survey	obs.	pseudo-R^2	N_{clusters}
(13)	.28	.23	−.68	−.31	−.91	−1.40	2.81	1.15	.53	.48	−.33	−.12	.043	−.10	−.12	.49	.80	.59	.72	−.045	.90	−.089	8.21	✓		✓				G17	7087	.067	
	(.073)	(.24)	(.29)	(.39)	(.48)	(.49)	(.12)	(.12)	(.10)	(.12)	(.15)	(.12)	(.16)	(.052)	(.067)	(.082)	(.23)	(.20)	(.22)	(.28)	(.22)	(.012)	(1.19)										
(14)	.28	.24	−.72	−.28		−.48	2.83	1.17	.55	.46	−.31	−.095	.035	−.099	−.079	.51	.78	.58	.71	−.070	.89	−.091	8.38	✓	CMA	✓				G17	7087	.069	46
	(.060)	(.26)	(.27)	(.35)		(100.0)	(.12)	(.20)	(.075)	(.074)	(.11)	(.090)	(.18)	(.053)	(.057)	(.076)	(.19)	(.19)	(.17)	(.25)	(.14)	(.008)	(.81)										
(15)	.28	.27	−.73			−.19	2.83	1.24	.49	.40	−.35	−.14	.042	−.097	−.032	.53	.57	.41	.51	−.22	.69	−.086	8.03	✓	CSD	✓				G17	6312	.074	141
	(.072)	(.24)	(.26)			(100.0)	(.14)	(.12)	(.12)	(.13)	(.17)	(.13)	(.22)	(.056)	(.069)	(.082)	(.24)	(.21)	(.23)	(.29)	(.21)	(.012)	(1.24)										
(16)																								✓	CT	✓				G17	0		
						(0)																											
(17)	.35	.53	−.42	−.68	−.96	−1.17	2.67	.88	.45	.34	−.47	−.006	−.096	−.23	−.28	.26	.46	.54	.64	−.31	.69	−.11	11.7	✓		×				G17	5114	.061	
	(.095)	(.26)	(.32)	(.48)	(.58)	(.70)	(.15)	(.13)	(.097)	(.10)	(.14)	(.14)	(.29)	(.063)	(.068)	(.094)	(.32)	(.31)	(.33)	(.36)	(.35)	(.016)	(1.93)										
(18)	.35	.54	−.41	−.67		−.18	2.71	.87	.46	.33	−.48	−.032	−.11	−.23	−.28	.28	.49	.56	.65	−.29	.66	−.12	12.1	✓	CMA	×				G17	5108	.066	45
	(.076)	(.25)	(.24)	(.66)		(100.0)	(.13)	(.12)	(.052)	(.080)	(.14)	(.15)	(.31)	(.055)	(.069)	(.10)	(.26)	(.27)	(.26)	(.33)	(.30)	(.016)	(1.95)										
(19)	.33	.47	−.39			.41	2.80	.91	.42	.38	−.54	−.085	−.27	−.24	−.29	.31	.46	.55	.68	−.34	.61	−.12	12.1	✓	CSD	×				G17	4481	.074	105
	(.11)	(.25)	(.25)			(.19)	(.16)	(.12)	(.071)	(.083)	(.15)	(.14)	(.36)	(.060)	(.080)	(.099)	(.34)	(.36)	(.36)	(.40)	(.37)	(.018)	(2.20)										
(20)																								✓	CT	×				G17	0		
						(0)																											
(21)	.28	.30	−.72	−.27	−1.03	−1.44	2.83	1.10	.50	.36	−.40	−.12	.061	−.11	−.15	.42	.74	.54	.65	−.29	.81	−.087	8.26	✓			✓			G17	9001	.067	
	(.067)	(.21)	(.26)	(.35)	(.43)	(.43)	(.11)	(.10)	(.084)	(.10)	(.12)	(.10)	(.14)	(.046)	(.057)	(.072)	(.21)	(.19)	(.21)	(.25)	(.20)	(.010)	(1.06)										
(22)	.29	.32	−.76	−.24		−.40	2.84	1.12	.52	.35	−.39	−.10	.058	−.11	−.11	.45	.72	.53	.66	−.33	.81	−.089	8.42	✓	CMA		✓			G17	9001	.069	46
	(.048)	(.18)	(.19)	(.30)		(100.0)	(.12)	(.16)	(.065)	(.065)	(.10)	(.068)	(.15)	(.037)	(.044)	(.070)	(.19)	(.18)	(.18)	(.28)	(.15)	(.008)	(.81)										
(23)	.27	.27	−.71			−.17	2.85	1.18	.50	.29	−.41	−.12	.10	−.093	−.074	.46	.55	.39	.50	−.43	.65	−.087	8.23	✓	CSD		✓			G17	8218	.074	152
	(.062)	(.20)	(.20)			(100.0)	(.13)	(.096)	(.094)	(.10)	(.14)	(.093)	(.17)	(.043)	(.047)	(.063)	(.22)	(.19)	(.21)	(.30)	(.21)	(.011)	(1.15)										
(24)																								✓	CT		✓			G17	0		
						(0)																											
(25)	.36	.58	−.21	−1.04	−.64	−.95	2.53	.90	.41	.43	−.52	.036	−.35	−.28	−.32	.30	.60	.73	.84	.060	.93	−.12	12.8	✓			×			G17	3200	.059	
	(.11)	(.33)	(.40)	(.59)	(.73)	(1.12)	(.19)	(.17)	(.12)	(.13)	(.20)	(.18)	(.39)	(.080)	(.087)	(.12)	(.37)	(.37)	(.39)	(.44)	(.45)	(.023)	(2.93)										
(26)	.35	.54	−.16	−.90		−.17	2.64	.88	.47	.46	−.48	.070	−.37	−.28	−.32	.30	.59	.67	.80	−.035	.87	−.12	13.0	✓	CMA		×			G17	3154	.064	38
	(.14)	(.25)	(.34)	(.90)		(100.0)	(.15)	(.13)	(.096)	(.11)	(.21)	(.19)	(.38)	(.076)	(.071)	(.10)	(.43)	(.44)	(.47)	(.46)	(.53)	(.019)	(2.57)										
(27)	.38	.56	−.44			.51	2.64	.98	.42	.52	−.65	.11	−.90	−.29	−.30	.33	.71	.76	.94	.030	.90	−.13	13.8	✓	CSD		×			G17	2564	.073	79
	(.15)	(.31)	(.34)			(.23)	(.20)	(.16)	(.12)	(.14)	(.24)	(.19)	(.48)	(.081)	(.087)	(.11)	(.48)	(.51)	(.53)	(.60)	(.57)	(.025)	(3.13)										
(28)																								✓	CT		×			G17	0		
						(0)																											
(37)	.23	−.10	.040	.26	−.93	−.51		1.93	.39	.32				−.088	−.043	.69	.79	1.03	.85		1.13	−.046	5.08	✓				✓		ED	6689	.061	
	(.090)	(.29)	(.36)	(.51)	(.63)	(100.0)		(.15)	(.091)	(.19)				(.064)	(.10)	(.11)	(.23)	(.20)	(.22)		(.22)	(.014)	(1.49)										
(38)	.45	.76	−.51	−.98	−.044	−.32	2.72	1.05	.15	.44	−.49	−.10	−.17	−.21	−.14	.42	.53	.65	.73	−.22	.73	−.092	9.39	✓				✓		G17	2627	.061	
	(.12)	(.37)	(.42)	(.60)	(.83)	(100.0)	(.20)	(.18)	(.13)	(.21)	(.20)	(.19)	(.34)	(.084)	(.096)	(.13)	(.40)	(.38)	(.41)	(.42)	(.41)	(.019)	(2.00)										

Continued on next page

	log(HH inc)	DA: log(HH inc)	CT: log(HH inc)	CSD: log(HH inc)	CMA: log(HH inc)	$\sum\beta_{\text{inc}}$	health	trust-N	married	asmarried	separated	divorced	widowed	male	noReligion	godImportance	student	employed	domestic	unemployed	retired	age	$(\text{age}/100)^2$	controls	f.e./clustering	$\tau_{\text{neigh}} \geq 10\text{yr}$	$\tau_{\text{city}} \geq 10\text{yr}$	foreign born	own house	survey	obs.	pseudo-R^2	N_{clusters}
⟨37-38⟩	.31	.22	−.19	−.26	−.61	−.42	2.72	1.58	.32	.37	−.49	−.10	−.17	−.13	−.098	.58	.72	.95	.82	−.22	1.04	−.062	6.63	✓				✓		⟨2⟩	9316		
	(.072)	(.23)	(.27)	(.39)	(.50)	(70.7)	(.20)	(.11)	(.075)	(.14)	(.20)	(.19)	(.34)	(.051)	(.070)	(.081)	(.20)	(.17)	(.20)	(.42)	(.19)	(.011)	(1.20)										
(39)	.25	−.16	.17	.085		.35		1.93	.42	.33				−.091	−.010	.71	.82	1.05	.88		1.13	−.047	5.18	✓	CMA			✓		ED	6659	.065	34
	(.080)	(.14)	(.14)	(.18)		(.15)		(.16)	(.090)	(.17)				(.045)	(.091)	(.13)	(.26)	(.19)	(.13)		(.10)	(.014)	(1.42)										
(40)	.43	.68	−.40	−1.09		−.37	2.72	1.03	.15	.43	−.51	−.056	−.31	−.20	−.12	.46	.60	.77	.88	−.20	.90	−.096	9.68	✓	CMA			✓		G17	2558	.064	25
	(.092)	(.25)	(.20)	(.55)		(100.0)	(.23)	(.16)	(.11)	(.37)	(.12)	(.17)	(.33)	(.093)	(.086)	(.097)	(.23)	(.23)	(.25)	(.30)	(.34)	(.016)	(1.69)										
⟨39-40⟩	.33	.049	−.010	−.024		.35	2.72	1.47	.31	.34	−.51	−.056	−.31	−.11	−.068	.55	.70	.94	.88	−.20	1.11	−.068	7.04	✓	CMA			✓		⟨2⟩	9217		
	(.060)	(.12)	(.11)	(.17)		(.15)	(.23)	(.11)	(.070)	(.15)	(.12)	(.17)	(.33)	(.041)	(.063)	(.077)	(.17)	(.15)	(.11)	(.30)	(.099)	(.011)	(1.09)										
(41)	.22	−.061	.071			.23		1.94	.40	.28				−.097	.062	.76	.78	1.00	.87		1.10	−.048	5.23	✓	CSD			✓		ED	6267	.072	93
	(.066)	(.26)	(.24)			(.11)		(.13)	(.095)	(.18)				(.058)	(.12)	(.12)	(.20)	(.18)	(.16)		(.17)	(.015)	(1.53)										
(42)	.34	.76	−.25			.85	2.55	1.05	.22	.52	−.51	.15	−.50	−.23	−.10	.50	.67	.88	.87	−.20	1.08	−.10	10.4	✓	CSD			✓		G17	2257	.065	52
	(.087)	(.36)	(.33)			(.27)	(.24)	(.23)	(.11)	(.28)	(.12)	(.13)	(.35)	(.093)	(.11)	(.11)	(.43)	(.40)	(.43)	(.43)	(.43)	(.020)	(2.11)										
⟨41-42⟩	.26	.22	−.039			.31	2.55	1.72	.33	.35	−.51	.15	−.50	−.13	−.025	.62	.76	.98	.87	−.20	1.10	−.068	7.00	✓	CSD			✓		⟨2⟩	8524		
	(.053)	(.21)	(.19)			(.10)	(.24)	(.11)	(.072)	(.15)	(.12)	(.13)	(.35)	(.049)	(.082)	(.084)	(.18)	(.16)	(.15)	(.43)	(.16)	(.012)	(1.24)										
(43)																								✓	CT			✓		ED	0		
						(0)																											
(44)																								✓	CT			✓		G17	0		
						(0)																											
⟨43-44⟩																								✓	CT			✓		⟨0⟩	0		
(45)	.26	.22	−.46	.055	−1.40	−1.32		1.67	.47	.49				−.083	.028	.49	1.45	1.27	1.17		1.56	−.060	5.83	✓				×		ED	16900	.053	
	(.058)	(.26)	(.31)	(.40)	(.48)	(.46)		(.12)	(.079)	(.10)				(.051)	(.078)	(.086)	(.18)	(.16)	(.19)		(.20)	(.011)	(1.25)										
(46)	.23	.23	−.60	−.28	−.82	−1.24	2.77	1.04	.58	.39	−.43	−.11	.025	−.14	−.22	.40	.77	.59	.71	−.027	.90	−.089	8.61	✓				×		G17	9574	.066	
	(.069)	(.20)	(.25)	(.35)	(.43)	(.43)	(.11)	(.099)	(.079)	(.085)	(.12)	(.10)	(.14)	(.045)	(.055)	(.071)	(.21)	(.20)	(.21)	(.27)	(.21)	(.010)	(1.10)										
⟨45-46⟩	.24	.23	−.55	−.13	−1.07	−1.28	2.77	1.29	.53	.43	−.43	−.11	.025	−.11	−.14	.44	1.17	1.00	.96	−.027	1.25	−.076	7.39	✓				×		⟨2⟩	26474		
	(.044)	(.16)	(.20)	(.26)	(.32)	(.32)	(.11)	(.077)	(.056)	(.066)	(.12)	(.10)	(.14)	(.034)	(.045)	(.055)	(.14)	(.12)	(.14)	(.27)	(.14)	(.007)	(.82)										
(47)	.24	.23	−.48	.23		.22		1.77	.51	.45				−.088	.13	.54	1.45	1.28	1.18		1.56	−.063	6.11	✓	CMA			×		ED	16900	.060	42
	(.058)	(.19)	(.26)	(.34)		(.27)		(.10)	(.082)	(.081)				(.047)	(.10)	(.092)	(.13)	(.15)	(.18)		(.17)	(.009)	(1.05)										
(48)	.23	.23	−.63	−.15		−.32	2.77	1.05	.61	.37	−.40	−.082	.022	−.14	−.19	.43	.73	.57	.68	−.066	.89	−.092	8.84	✓	CMA			×		G17	9574	.068	46
	(.068)	(.16)	(.24)	(.23)		(100.0)	(.10)	(.17)	(.052)	(.076)	(.12)	(.078)	(.17)	(.035)	(.050)	(.079)	(.22)	(.20)	(.20)	(.27)	(.20)	(.008)	(.82)										
⟨47-48⟩	.24	.23	−.56	−.032		.22	2.77	1.57	.58	.41	−.40	−.082	.022	−.12	−.13	.48	1.27	1.02	.96	−.066	1.27	−.079	7.81	✓	CMA			×		⟨2⟩	26474		
	(.044)	(.12)	(.18)	(.19)		(.27)	(.10)	(.088)	(.044)	(.055)	(.12)	(.078)	(.17)	(.028)	(.045)	(.060)	(.11)	(.12)	(.14)	(.27)	(.13)	(.006)	(.65)										
(49)	.28	.24	−.45			.062		1.85	.50	.49				−.075	.15	.55	1.44	1.27	1.18		1.58	−.070	6.80	✓	CSD			×		ED	16279	.076	213
	(.067)	(.25)	(.28)			(.058)		(.13)	(.084)	(.12)				(.046)	(.099)	(.098)	(.20)	(.20)	(.24)		(.26)	(.011)	(1.31)										
(50)	.21	.29	−.67			−.17	2.84	1.11	.58	.35	−.42	−.16	.009	−.16	−.16	.43	.63	.46	.55	−.089	.71	−.090	8.82	✓	CSD			×		G17	8769	.075	171
	(.077)	(.20)	(.21)			(100.0)	(.12)	(.10)	(.080)	(.083)	(.14)	(.10)	(.18)	(.043)	(.054)	(.074)	(.22)	(.20)	(.22)	(.28)	(.23)	(.011)	(1.16)										

Continued on next page

	log(HH inc)	DA: log(HH inc)	CT: log(HH inc)	CSD: log(HH inc)	CMA: log(HH inc)	$\sum\beta_{inc}$	health	trust-N	married	asmarried	separated	divorced	widowed	male	noReligion	godImportance	student	employed	domestic	unemployed	retired	age	$(age/100)^2$	controls	f.e./clustering	$\tau_{neigh} \geq 10yr$	$\tau_{city} \geq 10yr$	foreign born	own house	survey	obs.	pseudo-R^2	$N_{clusters}$
⟨49-50⟩	.25	.27	−.59			.062	2.84	1.41	.54	.39	−.42	−.16	.009	−.12	−.091	.48	1.07	.87	.84	−.089	1.10	−.080	7.94	✓	CSD			×		⟨2⟩	25048		
	(.051)	(.16)	(.17)			(.058)	(.12)	(.081)	(.058)	(.068)	(.14)	(.10)	(.18)	(.031)	(.047)	(.059)	(.15)	(.14)	(.16)	(.28)	(.17)	(.008)	(.87)										
(51)	.18	.79				.96		2.62	.96	.73				−.19	.45	1.02	2.80	2.22	2.14		2.71	−.045	3.85	✓	CT			×		ED	2425	.190	231
	(.18)	(.70)				(.76)		(.31)	(.26)	(.42)				(.14)	(.24)	(.28)	(.48)	(.41)	(.47)		(.51)	(.033)	(3.84)										
(52)	.47	−.29				.17	3.13	1.17	.21	.58	−1.89	−.10	.015	−.21	−.22	.42	.73	.13	.25	−.24	−.14	−.047	4.70	✓	CT			×		G17	936	.094	73
	(.28)	(.50)				(.27)	(.54)	(.36)	(.28)	(.27)	(.42)	(.49)	(.64)	(.21)	(.26)	(.35)	(1.08)	(1.03)	(1.00)	(1.05)	(.93)	(.060)	(5.99)										
⟨51-52⟩	.27	.068				.26	3.13	1.99	.61	.63	−1.89	−.10	.015	−.20	.14	.79	2.47	1.93	1.79	−.24	2.05	−.045	4.10	✓	CT			×		⟨2⟩	3361		
	(.15)	(.41)				(.25)	(.54)	(.24)	(.19)	(.23)	(.42)	(.49)	(.64)	(.12)	(.17)	(.22)	(.44)	(.38)	(.43)	(1.05)	(.45)	(.029)	(3.23)										
(57)	.15	.36	−.54	.30	−.94	−.68		1.85	.57	.57				−.096	.005	.61	1.25	1.07	1.05		1.46	−.042	3.55	✓					✓	ED	17619	.053	
	(.054)	(.25)	(.30)	(.36)	(.43)	(.52)		(.12)	(.075)	(.11)				(.047)	(.074)	(.078)	(.18)	(.16)	(.19)		(.19)	(.011)	(1.12)										
(58)	.29	.34	−.71	−.83	−.39	−1.30	2.81	1.22	.56	.45	−.36	−.19	−.009	−.13	−.20	.46	.52	.44	.67	−.11	.81	−.097	8.92	✓					✓	G17	8248	.064	
	(.070)	(.22)	(.28)	(.37)	(.44)	(.45)	(.12)	(.11)	(.088)	(.10)	(.15)	(.13)	(.17)	(.048)	(.061)	(.075)	(.25)	(.23)	(.25)	(.32)	(.25)	(.011)	(1.21)										
⟨57-58⟩	.20	.35	−.63	−.26	−.67	−1.03	2.81	1.52	.56	.50	−.36	−.19	−.009	−.11	−.12	.53	1.00	.87	.91	−.11	1.23	−.070	6.04	✓					✓	⟨2⟩	25867		
	(.043)	(.17)	(.20)	(.26)	(.31)	(.34)	(.12)	(.079)	(.057)	(.077)	(.15)	(.13)	(.17)	(.034)	(.047)	(.054)	(.15)	(.13)	(.15)	(.32)	(.15)	(.008)	(.82)										
(59)	.15	.34	−.52	.25		.22		1.91	.59	.52				−.10	.059	.63	1.27	1.09	1.08		1.47	−.044	3.78	✓	CMA				✓	ED	17619	.057	42
	(.054)	(.35)	(.43)	(.32)		(.27)		(.10)	(.075)	(.099)				(.045)	(.10)	(.085)	(.16)	(.15)	(.15)		(.13)	(.009)	(.94)										
(60)	.29	.36	−.76	−.74		−.85	2.82	1.24	.58	.42	−.34	−.17	−.007	−.12	−.17	.48	.49	.41	.65	−.16	.79	−.10	9.15	✓	CMA				✓	G17	8248	.066	46
	(.061)	(.18)	(.21)	(.34)		(.65)	(.089)	(.15)	(.087)	(.085)	(.11)	(.095)	(.20)	(.041)	(.043)	(.082)	(.24)	(.21)	(.21)	(.28)	(.19)	(.009)	(.92)										
⟨59-60⟩	.21	.35	−.71	−.20		.063	2.82	1.70	.58	.46	−.34	−.17	−.007	−.12	−.13	.55	1.03	.87	.93	−.16	1.26	−.072	6.53	✓	CMA				✓	⟨2⟩	25867		
	(.040)	(.16)	(.19)	(.23)		(.25)	(.089)	(.084)	(.057)	(.064)	(.11)	(.095)	(.20)	(.030)	(.040)	(.059)	(.13)	(.12)	(.12)	(.28)	(.11)	(.006)	(.66)										
(61)	.15	.36	−.47			.035		1.98	.57	.59				−.082	.11	.69	1.28	1.09	1.12		1.53	−.046	3.99	✓	CSD				✓	ED	17029	.071	212
	(.054)	(.27)	(.33)			(.062)		(.11)	(.086)	(.13)				(.043)	(.097)	(.090)	(.19)	(.18)	(.18)		(.21)	(.010)	(1.11)										
(62)	.28	.32	−.66			−.061	2.85	1.32	.54	.37	−.42	−.20	−.042	−.15	−.14	.48	.27	.17	.38	−.36	.48	−.096	8.86	✓	CSD				✓	G17	7457	.071	161
	(.061)	(.23)	(.23)			(.049)	(.13)	(.11)	(.096)	(.13)	(.17)	(.13)	(.22)	(.047)	(.058)	(.077)	(.32)	(.30)	(.31)	(.36)	(.30)	(.012)	(1.23)										
⟨61-62⟩	.21	.34	−.60			−.024	2.85	1.64	.56	.49	−.42	−.19	−.042	−.11	−.076	.57	1.02	.85	.93	−.36	1.19	−.066	6.16	✓	CSD				✓	⟨2⟩	24486		
	(.040)	(.17)	(.19)			(.038)	(.13)	(.078)	(.064)	(.092)	(.17)	(.13)	(.22)	(.032)	(.050)	(.059)	(.16)	(.15)	(.16)	(.36)	(.17)	(.008)	(.82)										
(63)	.14	.74				.88		2.22	.78	.70				−.26	.13	.74	1.52	1.08	1.04		1.46	−.028	2.64	✓	CT				✓	ED	4424	.165	403
	(.13)	(.63)				(.70)		(.26)	(.19)	(.32)				(.10)	(.17)	(.19)	(.55)	(.51)	(.53)		(.56)	(.024)	(2.53)										
(64)																								✓	CT				✓	G17	0		
						(0)																											
⟨63-64⟩	.14	.74				.88		2.22	.78	.70				−.26	.13	.74	1.52	1.08	1.04		1.46	−.028	2.64	✓	CT				✓	⟨1⟩	4424		
	(.13)	(.63)				(.70)		(.26)	(.19)	(.32)				(.10)	(.17)	(.19)	(.55)	(.51)	(.53)		(.56)	(.024)	(2.53)										
(65)	.42	−.34	−.090	−.32	−1.25	−1.58		1.49	.21	.28				−.018	.049	.49	1.26	1.35	1.01		1.06	−.097	11.1	✓					×	ED	5970	.053	
	(.11)	(.34)	(.39)	(.65)	(.79)	(.80)		(.17)	(.10)	(.15)				(.080)	(.12)	(.14)	(.22)	(.20)	(.24)		(.27)	(.016)	(1.83)										
(66)	.32	.34	−.43	.57	−2.56	−1.76	2.62	.67	.21	.29	−.54	−.047	−.14	−.23	−.20	.23	.89	.82	.60	−.10	.74	−.088	9.79	✓					×	G17	3953	.063	
	(.11)	(.29)	(.35)	(.54)	(.71)	(.67)	(.16)	(.14)	(.11)	(.11)	(.14)	(.12)	(.21)	(.072)	(.076)	(.11)	(.24)	(.23)	(.27)	(.28)	(.26)	(.015)	(1.62)										

Continued on next page

	log(HH inc)	DA: log(HH inc)	CT: log(HH inc)	CSD: log(HH inc)	CMA: log(HH inc)	$\sum\beta_{inc}$	health	trust-N	married	asmarried	separated	divorced	widowed	male	noReligion	godImportance	student	employed	domestic	unemployed	retired	age	$(age/100)^2$	controls	f.e./clustering	$\tau_{neigh} \geq 10yr$	$\tau_{city} \geq 10yr$	foreign born	own house	survey	obs.	pseudo-R^2	$N_{clusters}$
⟨65-66⟩	.37	.050	−.28	.21	−1.98	−1.68	2.62	1.00	.21	.29	−.54	−.047	−.14	−.14	−.13	.33	1.10	1.13	.83	−.10	.89	−.092	10.4	✓					×	⟨2⟩	9923		
	(.079)	(.22)	(.26)	(.41)	(.53)	(.51)	(.16)	(.11)	(.077)	(.091)	(.14)	(.12)	(.21)	(.054)	(.065)	(.084)	(.16)	(.15)	(.18)	(.28)	(.19)	(.011)	(1.21)										
(67)	.40	−.37	−.15	.32		.21		1.61	.24	.28				−.018	.17	.54	1.23	1.34	1.00		1.04	−.10	11.5	✓	CMA				×	ED	5970	.064	42
	(.085)	(.22)	(.39)	(.47)		(.21)		(.20)	(.073)	(.11)				(.068)	(.093)	(.18)	(.24)	(.20)	(.25)		(.24)	(.015)	(1.70)										
(68)	.33	.39	−.36	.35		.72	2.65	.70	.21	.28	−.58	−.052	−.14	−.23	−.17	.27	.98	.87	.64	−.075	.76	−.089	10.0	✓	CMA				×	G17	3940	.068	44
	(.22)	(.25)	(.29)	(.31)		(.27)	(.20)	(.12)	(.14)	(.14)	(.14)	(.11)	(.18)	(.057)	(.051)	(.087)	(.19)	(.26)	(.23)	(.26)	(.21)	(.016)	(1.94)										
⟨67-68⟩	.39	−.028	−.28	.34		.40	2.65	.94	.23	.28	−.58	−.052	−.14	−.14	−.091	.32	1.08	1.16	.80	−.075	.88	−.095	10.8	✓	CMA				×	⟨2⟩	9910		
	(.079)	(.17)	(.23)	(.26)		(.16)	(.20)	(.10)	(.065)	(.085)	(.14)	(.11)	(.18)	(.044)	(.045)	(.079)	(.15)	(.16)	(.17)	(.26)	(.16)	(.011)	(1.28)										
(69)	.42	−.28	−.16			−.016		1.53	.20	.23				−.037	.15	.46	1.05	1.20	.92		.86	−.11	12.2	✓	CSD				×	ED	5524	.070	99
	(.10)	(.28)	(.35)			(.005)		(.22)	(.087)	(.13)				(.085)	(.092)	(.15)	(.24)	(.23)	(.28)		(.27)	(.015)	(1.85)										
(70)	.30	.47	−.47			.30	2.62	.72	.20	.24	−.63	−.049	−.094	−.20	−.17	.23	1.12	1.03	.86	.076	.86	−.097	11.0	✓	CSD				×	G17	3524	.073	84
	(.17)	(.25)	(.29)			(.24)	(.20)	(.14)	(.12)	(.13)	(.15)	(.11)	(.26)	(.077)	(.076)	(.096)	(.22)	(.26)	(.29)	(.29)	(.28)	(.015)	(1.76)										
⟨69-70⟩	.39	.13	−.34			−.016	2.62	.94	.20	.23	−.63	−.049	−.094	−.12	−.041	.30	1.09	1.13	.89	.076	.86	−.10	11.6	✓	CSD				×	⟨2⟩	9048		
	(.089)	(.19)	(.22)			(.005)	(.20)	(.12)	(.070)	(.091)	(.15)	(.11)	(.26)	(.057)	(.059)	(.080)	(.16)	(.17)	(.20)	(.29)	(.19)	(.011)	(1.28)										
(71)																								✓	CT				×	ED	0		
						(0)																											
(72)																								✓	CT				×	G17	0		
						(0)																											
⟨71-72⟩																								✓	CT				×	⟨0⟩	0		

	log(own inc)	log(HH inc/$\sqrt{hh}$)	DA: log(HH inc)	CT: log(HH inc)	CSD: log(HH inc)	CMA: log(HH inc)	$\sum\beta_{\rm inc}$	mortgagePayment	log(houseValue)	DA: log(houseValue)	health	trust-N	trust-G
(1)	.019 (.28)	.38 (.32)	−.66 (.70)	−.42 (.82)	−.25 (1.36)	−2.62 (1.84)	−3.56 (1.72)			.84 (.47)	1.26 (.41)	.65 (.22)	.025 (.17)
(2)	.23 (.082)	.051 (.072)	.44 (.28)	−.56 (.32)	.13 (.38)	−.98 (.49)	−.69 (.82)	−.034 (.041)	.13 (.12)	−.13 (.20)		1.88 (.13)	−.039 (.050)
(3)	−.083 (.071)	.32 (.075)	.43 (.21)	−.57 (.24)	−.32 (.33)	−.69 (.44)	−.91 (.70)			−.26 (.15)	2.80 (.10)	1.04 (.099)	.029 (.045)
⟨1-3⟩	.049 (.053)	.18 (.051)	.37 (.16)	−.56 (.19)	−.13 (.24)	−.87 (.32)	−1.06 (.51)	−.034 (.041)	.13 (.12)	−.15 (.12)	2.71 (.100)	1.27 (.074)	−.0004 (.033)
(4)							(0)						
(5)	.24 (.067)	.035 (.084)	.41 (.47)	−.55 (.35)	.073 (.33)		.21 (.14)	−.030 (.053)	.15 (.17)	−.13 (.38)		1.92 (.12)	−.003 (.043)
(6)	−.083 (.052)	.31 (.058)	.35 (.17)	−.63 (.23)	−.22 (.19)		−.26 (100.0)			−.10 (.14)	2.81 (.10)	1.03 (.13)	.039 (.042)
⟨4-6⟩	.038 (.041)	.22 (.048)	.36 (.16)	−.60 (.19)	−.14 (.17)		.21 (.14)	−.030 (.053)	.15 (.17)	−.10 (.13)	2.81 (.10)	1.52 (.085)	.018 (.030)
(7)							(0)						
(8)	.29 (.087)	.006 (.083)	.20 (.29)	−.64 (.32)			−.14 (100.0)	−.024 (.046)	.095 (.12)	.40 (.25)		2.00 (.12)	−.020 (.049)
(9)	−.044 (.082)	.27 (.079)	.36 (.22)	−.68 (.19)			−.092 (100.0)			−.018 (.20)	2.84 (.12)	1.07 (.085)	.055 (.045)
⟨7-9⟩	.11 (.060)	.15 (.057)	.30 (.18)	−.67 (.16)			−.12 (70.7)	−.024 (.046)	.095 (.12)	.14 (.15)	2.84 (.12)	1.37 (.070)	.021 (.033)
(10)							(0)						
(11)	.19 (.23)	−.074 (.20)	.18 (.95)				.29 (100.0)	.050 (.094)	−.006 (.30)	.54 (1.11)		2.28 (.33)	−.073 (.13)
(12)	.56 (.26)	.35 (.29)	.052 (.61)				.96 (.48)			−.10 (.79)	3.19 (.46)	1.36 (.33)	−.26 (.18)
⟨10-12⟩	.35 (.17)	.059 (.16)	.089 (.52)				.96 (.48)	.050 (.094)	−.006 (.30)	.11 (.64)	3.19 (.46)	1.82 (.23)	−.14 (.11)

	married	asmarried	separated	divorced	widowed	male	noReligion	godImportance	student	employed	domestic	unemployed	retired	age	(age/100)^2	f.e./clustering	survey	obs.	pseudo-R^2	$N_{\rm clusters}$
(1)	.38 (.21)	.26 (.26)		−.78 (.39)	−.40 (.92)	.068 (.15)		.51 (.23)				−1.74 (.45)		.019 (.006)	.024 (.007)		E2	733	.043	
(2)	.57 (.077)	.58 (.12)				−.13 (.050)	.040 (.078)	.64 (.082)	1.28 (.20)	.99 (.18)	.98 (.21)		1.37 (.20)	−.051 (.011)	4.49 (1.19)		ED	16121	.054	
(3)	.44 (.072)	.36 (.082)	−.48 (.11)	−.12 (.094)	−.052 (.14)	−.15 (.043)	−.19 (.052)	.34 (.066)	.69 (.20)	.65 (.18)	.59 (.20)	−.067 (.24)	.88 (.20)	−.091 (.010)	9.15 (1.04)		G17	10780	.064	
⟨1-3⟩	.49 (.051)	.42 (.065)	−.48 (.11)	−.16 (.091)	−.060 (.14)	−.13 (.032)	−.12 (.043)	.46 (.050)	.98 (.14)	.82 (.13)	.77 (.14)	−.43 (.21)	1.12 (.14)	−.017 (.005)	.025 (.007)		⟨3⟩	27634		
(4)																CMA	E2	0		
(5)	.58 (.076)	.53 (.10)				−.14 (.045)	.093 (.10)	.66 (.095)	1.30 (.16)	1.01 (.15)	1.02 (.17)		1.38 (.16)	−.054 (.010)	4.77 (1.02)	CMA	ED	16121	.059	42
(6)	.46 (.045)	.35 (.056)	−.47 (.084)	−.10 (.066)	−.053 (.14)	−.15 (.031)	−.17 (.043)	.36 (.079)	.66 (.15)	.64 (.13)	.57 (.15)	−.098 (.21)	.87 (.14)	−.093 (.007)	9.28 (.73)	CMA	G17	10780	.066	46
⟨4-6⟩	.49 (.039)	.39 (.049)	−.47 (.084)	−.10 (.066)	−.053 (.14)	−.15 (.026)	−.13 (.040)	.48 (.061)	.95 (.11)	.79 (.098)	.77 (.11)	−.098 (.21)	1.09 (.10)	−.080 (.006)	7.76 (.59)	CMA	⟨2⟩	26901		
(7)																CSD	E2	0		
(8)	.56 (.090)	.58 (.13)				−.13 (.049)	.14 (.096)	.70 (.096)	1.33 (.18)	1.00 (.19)	1.02 (.20)		1.41 (.21)	−.056 (.011)	5.05 (1.21)	CSD	ED	15516	.073	203
(9)	.41 (.070)	.33 (.085)	−.54 (.098)	−.12 (.085)	−.083 (.15)	−.17 (.042)	−.14 (.051)	.38 (.065)	.53 (.19)	.51 (.19)	.41 (.19)	−.21 (.29)	.71 (.20)	−.091 (.009)	9.25 (1.00)	CSD	G17	9970	.071	173
⟨7-9⟩	.47 (.055)	.41 (.070)	−.54 (.098)	−.12 (.085)	−.083 (.15)	−.15 (.032)	−.074 (.045)	.48 (.054)	.95 (.13)	.75 (.13)	.70 (.14)	−.21 (.29)	1.04 (.14)	−.077 (.007)	7.53 (.77)	CSD	⟨2⟩	25486		
(10)																CT	E2	0		
(11)	.72 (.21)	.86 (.36)				−.20 (.13)	.14 (.21)	.71 (.22)	.83 (.63)	.16 (.61)	.19 (.62)		.54 (.65)	−.021 (.029)	1.88 (3.13)	CT	ED	3192	.146	292
(12)	.61 (.35)	.63 (.33)	−1.07 (.54)	.081 (.39)	−.075 (.67)	−.32 (.20)	−.30 (.18)	.66 (.26)	1.85 (.51)	.88 (.46)	.77 (.53)	.53 (.58)	.70 (.59)	−.096 (.039)	8.84 (4.41)	CT	G17	950	.106	73
⟨10-12⟩	.69 (.18)	.74 (.24)	−1.07 (.54)	.081 (.39)	−.075 (.67)	−.23 (.11)	−.12 (.14)	.69 (.17)	1.45 (.40)	.62 (.37)	.53 (.40)	.53 (.58)	.63 (.44)	−.048 (.023)	4.21 (2.55)	CT	⟨2⟩	4142		

Significance: 1% 5% 10%

Table A.2: Detailed regressions with alternate measures of wealth and income. These results are summarised in Table 2.5 on page 18.

	log(HH inc)	DA: log(HH inc)	CT: log(HH inc)	CSD: log(HH inc)	CMA: log(HH inc)	$\sum\beta_{\text{inc}}$	houseRooms	DA: houseRooms	CT: houseRooms	health	trust-N	trust-G	married	asmarried	separated	divorced	widowed	male	noReligion	godImportance	student	employed	domestic	unemployed	retired	age	$(\text{age}/100)^2$	f.e./clustering	survey	obs.	pseudo-R^2	N_{clusters}
(1)	.19	.35	−.99	−2.00	.71	−1.74		−.040	.17	1.06	.49	.026	.54	.25		−.37	.22	.006		.36				−1.28		.018	.023		E2	1123	.039	
	(.20)	(.70)	(.89)	(1.03)	(1.38)	(1.05)		(.088)	(.11)	(.33)	(.18)	(.12)	(.17)	(.21)		(.24)	(.31)	(.11)		(.18)				(.32)		(.004)	(.005)					
(2)	.15	.53	−.61	.31	−.99	−.61	.002	−.027	.011		1.86	−.022	.57	.57				−.094	.006	.61	1.26	1.07	1.05		1.46	−.041	3.53		ED	17619	.053	
	(.054)	(.35)	(.45)	(.36)	(.46)	(.47)	(.011)	(.039)	(.048)		(.12)	(.048)	(.075)	(.11)				(.047)	(.074)	(.078)	(.18)	(.16)	(.19)		(.19)	(.011)	(1.12)					
(3)	.29	.039	−.50	−.86	−.32	−1.35		.048	−.032	2.81	1.21	.012	.56	.45	−.36	−.19	−.009	−.13	−.20	.45	.53	.44	.68	−.11	.81	−.097	8.93		G17	8248	.064	
	(.070)	(.33)	(.43)	(.37)	(.47)	(.47)		(.038)	(.050)	(.12)	(.12)	(.052)	(.089)	(.10)	(.15)	(.13)	(.17)	(.048)	(.061)	(.075)	(.25)	(.24)	(.25)	(.32)	(.25)	(.011)	(1.20)					
⟨1-3⟩	.20	.28	−.60	−.37	−.59	−1.05	.002	.007	.006	2.61	1.33	−.004	.56	.47	−.36	−.23	.046	−.10	−.12	.51	1.01	.87	.91	−.69	1.23	−.0002	.023		⟨3⟩	26990		
	(.042)	(.23)	(.29)	(.25)	(.32)	(.32)	(.011)	(.026)	(.033)	(.11)	(.076)	(.034)	(.054)	(.072)	(.15)	(.12)	(.15)	(.032)	(.047)	(.052)	(.15)	(.13)	(.15)	(.22)	(.15)	(.004)	(.005)					
(4)	.23	.31	−1.10	−1.18		−1.74		.001	.17	.89	.44	.033	.61	.17		−.26	.22	−.095		.45				−1.11		.015	.019	CMA	E2	1017	.042	23
	(.18)	(.61)	(.58)	(1.59)		(2.06)		(.078)	(.099)	(.31)	(.14)	(.19)	(.19)	(.27)		(.41)	(.25)	(.078)		(.14)				(.39)		(.006)	(.007)					
(5)	.15	.46	−.83	.21		−.022	.004	−.017	.046		1.90	.014	.59	.52				−.10	.061	.63	1.27	1.09	1.08		1.47	−.044	3.79	CMA	ED	17619	.057	42
	(.056)	(.35)	(.48)	(.31)		(100.0)	(.011)	(.035)	(.046)		(.11)	(.035)	(.075)	(.10)				(.045)	(.10)	(.084)	(.16)	(.15)	(.16)		(.13)	(.009)	(.94)					
(6)	.29	.036	−.63	−.81		−1.12		.052	−.018	2.82	1.22	.028	.58	.42	−.34	−.17	−.007	−.13	−.17	.48	.49	.41	.66	−.16	.79	−.10	9.19	CMA	G17	8248	.066	46
	(.061)	(.27)	(.37)	(.39)		(.67)		(.023)	(.036)	(.093)	(.13)	(.054)	(.087)	(.085)	(.11)	(.096)	(.19)	(.042)	(.043)	(.082)	(.24)	(.21)	(.21)	(.28)	(.20)	(.009)	(.91)					
⟨4-6⟩	.21	.21	−.79	−.22		−1.18	.004	.029	.019	2.66	1.30	.018	.59	.45	−.34	−.17	.078	−.11	−.13	.54	1.03	.86	.93	−.48	1.26	−.026	.020	CMA	⟨3⟩	26884		
	(.040)	(.20)	(.26)	(.24)		(.64)	(.011)	(.019)	(.027)	(.089)	(.072)	(.029)	(.054)	(.063)	(.11)	(.093)	(.15)	(.029)	(.040)	(.054)	(.13)	(.12)	(.13)	(.23)	(.11)	(.004)	(.007)					
(7)																												CSD	E2	0		
						(0)																										
(8)	.15	.49	−.56			.088	.006	−.024	.011		1.99	−.002	.56	.59				−.082	.11	.69	1.28	1.09	1.12		1.53	−.046	3.97	CSD	ED	17029	.071	212
	(.055)	(.35)	(.42)			(.14)	(.012)	(.036)	(.051)		(.12)	(.045)	(.086)	(.13)				(.043)	(.097)	(.090)	(.18)	(.18)	(.18)		(.21)	(.010)	(1.10)					
(9)	.28	.11	−.61			−.23		.034	−.008	2.85	1.29	.042	.55	.38	−.42	−.19	−.040	−.15	−.14	.47	.26	.17	.39	−.36	.49	−.097	8.90	CSD	G17	7457	.071	161
	(.061)	(.34)	(.36)			(100.0)		(.033)	(.049)	(.13)	(.12)	(.052)	(.097)	(.13)	(.17)	(.13)	(.22)	(.047)	(.058)	(.077)	(.32)	(.30)	(.31)	(.36)	(.30)	(.011)	(1.23)					
⟨7-9⟩	.21	.29	−.59			.088	.006	.008	.001	2.85	1.63	.017	.56	.49	−.42	−.19	−.040	−.11	−.075	.57	1.02	.85	.93	−.36	1.19	−.069	6.18	CSD	⟨2⟩	24486		
	(.041)	(.24)	(.28)			(.14)	(.012)	(.024)	(.035)	(.13)	(.083)	(.034)	(.064)	(.092)	(.17)	(.13)	(.22)	(.032)	(.050)	(.059)	(.16)	(.15)	(.16)	(.36)	(.17)	(.007)	(.82)					
(10)																												CT	E2	0		
						(0)																										
(11)	.14	.35				.49	.003	.053			2.21	.009	.78	.70				−.26	.13	.74	1.51	1.07	1.03		1.45	−.029	2.67	CT	ED	4424	.165	403
	(.13)	(.95)				(.74)	(.025)	(.092)			(.28)	(.12)	(.19)	(.32)				(.10)	(.17)	(.19)	(.55)	(.51)	(.53)		(.56)	(.024)	(2.53)					
(12)																												CT	G17	0		
						(0)																										
⟨10-12⟩	.14	.35				.49	.003	.053			2.21	.009	.78	.70				−.26	.13	.74	1.52	1.07	1.03		1.45	−.029	2.67	CT	⟨1⟩	4424		
	(.13)	(.95)				(.74)	(.025)	(.092)			(.28)	(.12)	(.19)	(.32)				(.10)	(.17)	(.19)	(.55)	(.51)	(.53)		(.56)	(.024)	(2.53)					

Significance: 1% 5% 10%

Table A.3: Detailed regressions with dwelling size. These results are summarised in Table 2.6 on page 19.

Table A.4: Detailed regressions with demographic comparison groups. See Table 2.11 on page 29 for explanation and summary.
Significance: 1% 5% 10%

		log(HH inc)	CT: log(HH inc)	CSD: log(HH inc)	CMA: log(HH inc)	$\sum\beta_{inc}$	health	trust-N	trust-G	married	asmarried	separated	divorced	widowed	male	noReligion	godImportance	student	employed	domestic	unemployed	retired	age	$(age/100)^2$	f.e./clustering	survey	obs.	pseudo-R^2	$N_{clusters}$
(37)	All	.33	−.40	−1.00	−1.59	−2.67	1.03	.52	.041	.65	.36		−.26	.095	−.060		.28				−1.23		−.034	5.36		E2	1131	.040	
		(.20)	(.51)	(1.40)	(1.81)	(1.61)	(.33)	(.17)	(.12)	(.17)	(.21)		(.24)	(.32)	(.11)		(.18)				(.31)		(.021)	(2.08)					
(38)	All	.24	−.16	.056	−1.76	−1.62		1.75	−.014	.45	.47				−.071	.016	.56	1.27	1.19	1.07		1.41	−.056	5.69		ED	24113	.053	
		(.047)	(.20)	(.41)	(.58)	(.57)		(.10)	(.041)	(.060)	(.090)				(.040)	(.063)	(.068)	(.14)	(.13)	(.15)		(.15)	(.009)	(.97)					
(39)	All	.33	−.39	−.40	−1.59	−2.05	2.77	1.04	.026	.44	.36	−.45	−.071	−.023	−.17	−.20	.37	.64	.55	.66	−.20	.81	−.091	9.03		G17	12457	.064	
		(.058)	(.19)	(.39)	(.55)	(.61)	(.095)	(.090)	(.042)	(.066)	(.075)	(.10)	(.089)	(.13)	(.039)	(.047)	(.061)	(.18)	(.17)	(.18)	(.21)	(.18)	(.009)	(.93)					
⟨37-39⟩	All	.28	−.29	−.21	−1.67	−1.88	2.63	1.24	.007	.46	.40	−.45	−.094	−.006	−.12	−.12	.44	1.03	.95	.90	−.52	1.16	−.070	7.23		⟨3⟩	37701		
		(.036)	(.13)	(.28)	(.39)	(.40)	(.091)	(.063)	(.029)	(.043)	(.056)	(.10)	(.083)	(.12)	(.027)	(.038)	(.044)	(.11)	(.10)	(.11)	(.18)	(.12)	(.006)	(.64)					
(40)	All	.38	−.18	−.48		−.27	.85	.47	.038	.74	.30		−.16	.13	−.13		.38				−1.07		−.034	4.98	CMA	E2	1031	.043	24
		(.18)	(.26)	(2.76)		(100.0)	(.33)	(.12)	(.19)	(.24)	(.28)		(.44)	(.21)	(.072)		(.13)				(.39)		(.020)	(1.61)					
(41)	All	.23	−.17	.14		.19		1.80	.018	.47	.42				−.075	.093	.59	1.26	1.19	1.07		1.41	−.059	5.96	CMA	ED	24113	.058	42
		(.041)	(.19)	(.39)		(.24)		(.095)	(.024)	(.057)	(.067)				(.038)	(.079)	(.071)	(.14)	(.13)	(.14)		(.13)	(.009)	(.94)					
(42)	All	.33	−.42	−.39		−.48	2.77	1.03	.038	.45	.33	−.45	−.063	−.031	−.17	−.16	.40	.63	.55	.67	−.21	.81	−.093	9.19	CMA	G17	12457	.066	46
		(.052)	(.13)	(.35)		(100.0)	(.10)	(.12)	(.039)	(.046)	(.057)	(.079)	(.061)	(.13)	(.033)	(.037)	(.068)	(.18)	(.17)	(.15)	(.23)	(.14)	(.006)	(.67)					
⟨40-42⟩	All	.27	−.32	−.16		.19	2.60	1.24	.024	.46	.37	−.45	−.065	.012	−.13	−.11	.48	1.02	.94	.88	−.44	1.13	−.080	7.79	CMA	⟨3⟩	37601		
		(.032)	(.097)	(.26)		(.24)	(.099)	(.064)	(.020)	(.035)	(.043)	(.079)	(.060)	(.11)	(.024)	(.034)	(.046)	(.11)	(.11)	(.11)	(.20)	(.095)	(.005)	(.52)					
(43)	All	.51	−.63			−.12	.71	.53	.16	.78	.45		−.17	.56	−.12		.44				−1.04		−.039	5.69	CSD	E2	804	.052	24
		(.16)	(.44)			(.17)	(.47)	(.16)	(.22)	(.23)	(.39)		(.39)	(.21)	(.11)		(.16)				(.33)		(.022)	(2.11)					
(44)	All	.25	−.16			.089		1.84	.014	.46	.45				−.068	.12	.62	1.25	1.19	1.10		1.42	−.063	6.40	CSD	ED	23468	.069	221
		(.050)	(.17)			(.064)		(.098)	(.034)	(.057)	(.088)				(.039)	(.077)	(.068)	(.16)	(.16)	(.17)		(.18)	(.009)	(.99)					
(45)	All	.32	−.43			−.11	2.81	1.07	.056	.43	.31	−.47	−.081	−.060	−.18	−.14	.41	.52	.44	.55	−.30	.69	−.093	9.23	CSD	G17	11665	.070	185
		(.057)	(.11)			(.075)	(.11)	(.080)	(.042)	(.061)	(.080)	(.093)	(.076)	(.14)	(.039)	(.047)	(.058)	(.18)	(.17)	(.17)	(.23)	(.17)	(.009)	(.93)					
⟨43-45⟩	All	.29	−.36			−.004	2.69	1.26	.033	.46	.37	−.47	−.084	.13	−.12	−.069	.49	.94	.84	.82	−.54	1.03	−.075	7.70	CSD	⟨3⟩	35937		
		(.037)	(.091)			(.047)	(.11)	(.058)	(.026)	(.041)	(.059)	(.093)	(.075)	(.12)	(.027)	(.040)	(.043)	(.12)	(.11)	(.12)	(.19)	(.13)	(.006)	(.64)					
(46)	All																								CT	E2	0		
						(0)																							
(47)	All	.30				.30		2.18	.11	.74	.45				−.18	.20	.82	1.63	1.43	1.25		1.75	−.073	7.59	CT	ED	8454	.167	762
		(.096)				(.096)		(.19)	(.083)	(.12)	(.20)				(.077)	(.13)	(.14)	(.29)	(.26)	(.29)		(.31)	(.016)	(1.83)					
(48)	All	.71				.71	3.26	1.07	−.11	.31	.56	−1.14	.073	.15	−.18	−.22	.61	.61	.37	.20	−.33	.27	−.086	7.58	CT	G17	1397	.100	111
		(.21)				(.21)	(.39)	(.23)	(.14)	(.27)	(.26)	(.35)	(.36)	(.48)	(.18)	(.18)	(.20)	(.68)	(.60)	(.63)	(.69)	(.53)	(.034)	(3.79)					
⟨46-48⟩	All	.37				.37	3.26	1.73	.049	.67	.49	−1.13	.073	.15	−.18	.060	.76	1.47	1.26	1.07	−.33	1.38	−.075	7.59	CT	⟨2⟩	9851		
		(.087)				(.087)	(.39)	(.15)	(.071)	(.11)	(.16)	(.35)	(.36)	(.48)	(.071)	(.10)	(.11)	(.27)	(.24)	(.26)	(.69)	(.27)	(.014)	(1.64)					

Continued on next page

	log(HH inc)	CT: log(HH inc)	CSD: log(HH inc)	CMA: log(HH inc)	$\sum \beta_{\text{inc}}$	health	trust-N	trust-G	married	asmarried	separated	divorced	widowed	male	noReligion	godImportance	student	employed	domestic	unemployed	retired	age	$(\text{age}/100)^2$	f.e./clustering	survey	obs.	pseudo-R^2	N_{clusters}
(49) Age groups	.40	−.68	−1.11	−.63	−2.03	1.01	.50	.069	.76	.46		−.27	.18	−.063		.28				−1.20		.053	−2.71		E2	1129	.045	
	(.20)	(.50)	(1.36)	(1.36)	(.86)	(.32)	(.17)	(.12)	(.17)	(.21)		(.24)	(.31)	(.11)		(.18)				(.30)		(.028)	(2.67)					
(50) Age groups	.23	−.039	.16	−.51	−.17		1.74	−.023	.46	.52				−.072	.0005	.55	1.20	1.20	1.08		1.43	−.034	3.53		ED	24085	.053	
	(.047)	(.19)	(.40)	(.39)	(100.0)		(.10)	(.041)	(.060)	(.090)				(.041)	(.063)	(.068)	(.14)	(.13)	(.15)		(.15)	(.012)	(1.25)					
(51) Age groups	.33	−.32	−.27	−.23	−.49	2.78	1.02	.021	.49	.45	−.41	−.051	.062	−.17	−.20	.36	.53	.58	.72	−.17	.85	−.046	4.65		G17	12433	.064	
	(.058)	(.17)	(.37)	(.37)	(.58)	(.095)	(.090)	(.042)	(.067)	(.075)	(.10)	(.089)	(.13)	(.040)	(.047)	(.061)	(.18)	(.17)	(.18)	(.21)	(.18)	(.012)	(1.21)					
⟨49-51⟩ Age groups	.27	−.22	−.11	−.38	−.97	2.64	1.23	.002	.49	.48	−.41	−.078	.080	−.12	−.13	.43	.94	.97	.94	−.52	1.18	−.032	3.46		⟨3⟩	37647		
	(.036)	(.12)	(.26)	(.26)	(.48)	(.091)	(.063)	(.029)	(.043)	(.056)	(.10)	(.083)	(.12)	(.028)	(.038)	(.044)	(.11)	(.10)	(.11)	(.17)	(.12)	(.008)	(.83)					
(52) Age groups	.48	−.59	−1.54		−1.65	.85	.45	.066	.85	.40		−.18	.23	−.14		.38				−1.05		.043	−2.09	CMA	E2	1029	.048	24
	(.18)	(.55)	(.67)		(.63)	(.33)	(.13)	(.19)	(.25)	(.28)		(.43)	(.22)	(.073)		(.13)				(.35)		(.023)	(1.75)					
(53) Age groups	.23	−.081	−.087		.065		1.80	.018	.47	.43				−.078	.095	.60	1.23	1.19	1.08		1.41	−.050	5.10	CMA	ED	24085	.058	42
	(.045)	(.14)	(.12)		(.042)		(.097)	(.025)	(.058)	(.065)				(.038)	(.080)	(.072)	(.16)	(.13)	(.14)		(.13)	(.010)	(1.10)					
(54) Age groups	.34	−.36	−.27		−.29	2.79	1.00	.041	.49	.38	−.42	−.056	.028	−.17	−.15	.41	.53	.57	.71	−.19	.84	−.059	5.86	CMA	G17	12433	.067	46
	(.053)	(.18)	(.22)		(100.0)	(.11)	(.12)	(.037)	(.051)	(.057)	(.081)	(.059)	(.13)	(.033)	(.037)	(.065)	(.17)	(.17)	(.16)	(.24)	(.14)	(.008)	(.81)					
⟨52-54⟩ Age groups	.29	−.20	−.17		.058	2.61	1.22	.026	.49	.41	−.42	−.058	.079	−.13	−.11	.48	.91	.96	.91	−.46	1.16	−.049	4.66	CMA	⟨3⟩	37547		
	(.034)	(.11)	(.11)		(.042)	(.10)	(.064)	(.021)	(.038)	(.042)	(.081)	(.058)	(.11)	(.024)	(.034)	(.045)	(.12)	(.11)	(.11)	(.20)	(.096)	(.006)	(.61)					
(55) Age groups	.65	−1.47			−.82	.73	.52	.18	.86	.52		−.16	.68	−.13		.41				−1.02		.013	.92	CSD	E2	802	.057	24
	(.15)	(.21)			(100.0)	(.48)	(.17)	(.22)	(.24)	(.39)		(.39)	(.21)	(.11)		(.15)				(.31)		(.024)	(2.24)					
(56) Age groups	.25	−.082			.17		1.84	.014	.46	.46				−.069	.12	.62	1.24	1.19	1.10		1.42	−.059	5.98	CSD	ED	23448	.069	221
	(.054)	(.17)			(.076)		(.098)	(.034)	(.059)	(.090)				(.040)	(.077)	(.068)	(.18)	(.15)	(.17)		(.18)	(.012)	(1.25)					
(57) Age groups	.34	−.52			−.18	2.82	1.05	.059	.46	.35	−.45	−.077	−.012	−.18	−.13	.41	.44	.46	.59	−.29	.72	−.066	6.56	CSD	G17	11646	.071	185
	(.059)	(.12)			(.16)	(.12)	(.081)	(.041)	(.062)	(.077)	(.092)	(.074)	(.14)	(.040)	(.046)	(.058)	(.18)	(.17)	(.17)	(.23)	(.18)	(.011)	(1.13)					
⟨55-57⟩ Age groups	.32	−.58			.11	2.70	1.27	.035	.47	.40	−.45	−.080	.20	−.12	−.067	.49	.84	.86	.85	−.55	1.06	−.055	5.64	CSD	⟨3⟩	35896		
	(.039)	(.088)			(.069)	(.11)	(.059)	(.026)	(.042)	(.058)	(.092)	(.073)	(.12)	(.027)	(.039)	(.042)	(.13)	(.11)	(.12)	(.19)	(.13)	(.008)	(.78)					
(58) Age groups																								CT	E2	0		
					(0)																							
(59) Age groups	.30				.30		2.18	.11	.74	.45				−.18	.20	.82	1.63	1.43	1.25		1.75	−.073	7.59	CT	ED	8454	.167	762
	(.096)				(.096)		(.19)	(.083)	(.12)	(.20)				(.077)	(.13)	(.14)	(.29)	(.26)	(.29)		(.31)	(.016)	(1.83)					
(60) Age groups	.71				.71	3.26	1.07	−.11	.31	.56	−1.14	.073	.15	−.18	−.22	.61	.61	.37	.20	−.33	.27	−.086	7.58	CT	G17	1397	.100	111
	(.21)				(.21)	(.39)	(.23)	(.14)	(.27)	(.26)	(.35)	(.36)	(.48)	(.18)	(.18)	(.20)	(.68)	(.60)	(.63)	(.69)	(.53)	(.034)	(3.79)					
⟨58-60⟩ Age groups	.37				.37	3.26	1.73	.049	.67	.49	−1.13	.073	.15	−.18	.060	.76	1.47	1.26	1.07	−.33	1.38	−.075	7.59	CT	⟨2⟩	9851		
	(.087)				(.087)	(.39)	(.15)	(.071)	(.11)	(.16)	(.35)	(.36)	(.48)	(.071)	(.10)	(.11)	(.27)	(.24)	(.26)	(.69)	(.27)	(.014)	(1.64)					
(61) Vismin groups	.069	.35	−.28	−1.40	−1.26		1.57	.092	.34	.51				−.035	.31	1.17	1.10	1.15	1.09		1.55	−.038	4.30		ED	4581	.057	
	(.10)	(.32)	(.91)	(1.19)	(.98)		(.22)	(.088)	(.13)	(.37)				(.087)	(.18)	(.19)	(.30)	(.28)	(.33)		(.35)	(.022)	(2.37)					
(62) Vismin groups	.091	.33	−1.25		−.83		1.59	.10	.36	.56				−.035	.35	1.18	1.12	1.15	1.06		1.55	−.039	4.46	CMA	ED	4541	.061	18
	(.061)	(.25)	(.52)		(.79)		(.22)	(.036)	(.11)	(.55)				(.051)	(.15)	(.17)	(.33)	(.26)	(.19)		(.19)	(.015)	(1.25)					

Continued on next page

	log(HH inc)	CT: log(HH inc)	CSD: log(HH inc)	CMA: log(HH inc)	$\sum \beta_{\text{inc}}$	health	trust-N	trust-G	married	asmarried	separated	divorced	widowed	male	noReligion	godImportance	student	employed	domestic	unemployed	retired	age	$(\text{age}/100)^2$	f.e./clustering	survey	obs.	pseudo-R^2	N_{clusters}
(63) Vismin groups	.075	.082			.16		1.57	.12	.36	.60				−.057	.41	1.20	1.09	1.11	1.03		1.53	−.042	4.77	CSD	ED	4425	.066	50
	(.10)	(.29)			(100.0)		(.24)	(.073)	(.14)	(.37)				(.089)	(.15)	(.19)	(.32)	(.24)	(.25)		(.25)	(.026)	(2.72)					
(64) Vismin groups																								CT	ED	0		
					(0)																							

Table A.5: Detailed regressions for income effects, sex, and marriage. Summary of estimates in the format described on page 14. These results are summarised in Table 2.9 on page 24.

Significance: 1% 5% 10%

		log(HH inc)	DA: log(HH inc)	CT: log(HH inc)	CSD: log(HH inc)	CMA: log(HH inc)	$\sum \beta_{\text{inc}}$	trust-N	married	asmarried	separated	divorced	widowed	male	noReligion	godImportance	student	employed	domestic	unemployed	retired	age	$(\text{age}/100)^2$	CMA f.e.	CSD f.e.	CT f.e.	clustering	survey	obs.
(1)	males	.77	.20	−.59	−.66	−.95	−1.24	.96	.48	.81		−.54	.38			.48				−1.06		−.091	11.3					E2	514
		(.35)	(.75)	(.88)	(1.52)	(1.81)	(100.0)	(.26)	(.25)	(.31)		(.37)	(.75)			(.27)				(.43)		(.034)	(3.56)						
(2)	males	.26	−.19	.075	−.35	−.69	−.90	1.73	.39	.47					−.049	.52	1.53	1.42	1.09		1.61	−.066	6.59					ED	11606
		(.069)	(.29)	(.35)	(.47)	(.55)	(.69)	(.14)	(.098)	(.14)					(.087)	(.095)	(.19)	(.17)	(.31)		(.21)	(.014)	(1.49)						
(3)	males	.60	.59	−.26	−.67	−1.15	−.90	1.30	.38	.34	−.42	.055	−.023		−.16	.31	1.91	1.81	1.80	.72	1.77	−.11	10.6					G17	5605
		(.10)	(.26)	(.32)	(.47)	(.56)	(1.32)	(.13)	(.10)	(.11)	(.17)	(.15)	(.27)		(.069)	(.087)	(.30)	(.28)	(.44)	(.33)	(.30)	(.013)	(1.38)						
⟨1-3⟩	males	.37	.23	−.14	−.52	−.92	−.90	1.44	.39	.42	−.42	−.031	.022		−.12	.41	1.64	1.52	1.33	.063	1.66	−.091	8.91					⟨3⟩	17725
		(.057)	(.19)	(.23)	(.32)	(.38)	(.61)	(.090)	(.068)	(.083)	(.17)	(.14)	(.25)		(.054)	(.062)	(.16)	(.14)	(.25)	(.26)	(.17)	(.009)	(.98)						
(4)	males																							✓			CMA	E2	0
							(0)																						
(5)	males	.25	−.22	.15	−.29		−.11	1.82	.41	.44					.006	.54	1.50	1.42	1.11		1.64	−.069	6.84	✓			CMA	ED	11606
		(.063)	(.26)	(.29)	(.43)		(100.0)	(.15)	(.072)	(.094)					(.092)	(.096)	(.14)	(.11)	(.28)		(.17)	(.013)	(1.54)						
(6)	males	.60	.57	−.27	−.54		.37	1.28	.40	.32	−.41	.061	−.035		−.15	.34	1.92	1.84	1.81	.72	1.78	−.12	11.0	✓			CMA	G17	5598
		(.10)	(.22)	(.25)	(.60)		(.30)	(.18)	(.067)	(.081)	(.17)	(.10)	(.23)		(.073)	(.13)	(.31)	(.29)	(.42)	(.40)	(.30)	(.013)	(1.51)						
⟨4-6⟩	males	.35	.24	−.092	−.37		.37	1.60	.40	.37	−.41	.061	−.035		−.087	.47	1.57	1.47	1.33	.72	1.67	−.093	8.94	✓			CMA	⟨2⟩	17204
		(.053)	(.17)	(.19)	(.35)		(.30)	(.12)	(.049)	(.061)	(.17)	(.10)	(.23)		(.057)	(.077)	(.12)	(.100)	(.24)	(.40)	(.15)	(.009)	(1.08)						
(7)	males																								✓		CSD	E2	0
							(0)																						
(8)	males	.27	−.22	.14			.19	1.85	.41	.46					.052	.58	1.46	1.36	1.22		1.60	−.070	7.08		✓		CSD	ED	10974
		(.079)	(.35)	(.35)			(.13)	(.14)	(.079)	(.13)					(.090)	(.090)	(.18)	(.16)	(.28)		(.22)	(.015)	(1.73)						
(9)	males	.60	.51	−.26			.85	1.31	.30	.21	−.43	.049	−.032		−.086	.36	1.80	1.77	1.75	.66	1.72	−.11	10.5		✓		CSD	G17	4889
		(.12)	(.28)	(.26)			(.27)	(.13)	(.11)	(.13)	(.16)	(.13)	(.27)		(.077)	(.096)	(.34)	(.33)	(.47)	(.40)	(.36)	(.016)	(1.71)						
⟨7-9⟩	males	.37	.23	−.12			.31	1.55	.37	.33	−.43	.049	−.032		−.028	.48	1.54	1.43	1.36	.66	1.63	−.090	8.82		✓		CSD	⟨2⟩	15863
		(.066)	(.22)	(.21)			(.11)	(.096)	(.065)	(.091)	(.16)	(.13)	(.27)		(.059)	(.066)	(.16)	(.14)	(.24)	(.40)	(.19)	(.011)	(1.22)						
(10)	males																									✓	CT	E2	0
							(0)																						
(11)	males																									✓	CT	ED	0
							(0)																						
(12)	males																									✓	CT	G17	0
							(0)																						
⟨10-12⟩	males																									✓	CT	⟨0⟩	0

Continued on next page

		log(HH inc)	DA: log(HH inc)	CT: log(HH inc)	CSD: log(HH inc)	CMA: log(HH inc)	$\sum \beta_{\text{inc}}$	trust-N	married	asmarried	separated	divorced	widowed	male	noReligion	godImportance	student	employed	domestic	unemployed	retired	age	$(\text{age}/100)^2$	CMA f.e.	CSD f.e.	CT f.e.	clustering	survey	obs.
(13)	females	−.019	−.26	1.01	−1.83	−.79	−1.89	.22	.79	.062		−.17	.16			.16				−1.22		.0002	1.68					E2	627
		(.27)	(.72)	(.88)	(1.40)	(1.70)	(1.47)	(.24)	(.24)	(.29)		(.33)	(.36)			(.26)				(.39)		(.026)	(2.47)						
(14)	females	.20	.45	−.75	.53	−1.45	−1.03	1.73	.52	.46					.083	.60	1.00	.97	.83		1.18	−.051	5.30					ED	11983
		(.067)	(.27)	(.34)	(.45)	(.52)	(.51)	(.13)	(.079)	(.12)					(.093)	(.098)	(.22)	(.20)	(.21)		(.23)	(.012)	(1.29)						
(15)	females	.45	.45	−.68	−.40	−1.12	−1.31	1.27	.55	.34	−.43	−.23	−.14		−.29	.38	.90	.77	.85	.18	1.02	−.085	7.27					G17	6603
		(.072)	(.23)	(.29)	(.39)	(.49)	(.63)	(.11)	(.087)	(.10)	(.13)	(.11)	(.16)		(.066)	(.086)	(.24)	(.22)	(.23)	(.33)	(.24)	(.012)	(1.34)						
⟨13-15⟩	females	.30	.41	−.61	−.075	−1.26	−1.19	1.32	.55	.37	−.43	−.22	−.087		−.17	.46	.95	.88	.84	−.40	1.11	−.061	5.68					⟨3⟩	19213
		(.048)	(.17)	(.22)	(.29)	(.35)	(.38)	(.081)	(.057)	(.076)	(.13)	(.10)	(.15)		(.054)	(.063)	(.16)	(.15)	(.16)	(.25)	(.17)	(.008)	(.87)						
(16)	females																							✓			CMA	E2	0
							(0)																						
(17)	females	.19	.47	−.76	.67		.57	1.79	.54	.41					.18	.66	.99	.98	.84		1.15	−.053	5.57	✓			CMA	ED	11983
		(.055)	(.24)	(.34)	(.33)		(.36)	(.099)	(.058)	(.10)					(.084)	(.066)	(.29)	(.27)	(.28)		(.29)	(.009)	(.94)						
(18)	females	.45	.46	−.67	−.34		−.11	1.27	.57	.32	−.42	−.23	−.13		−.28	.40	.89	.77	.86	.17	1.00	−.087	7.41	✓			CMA	G17	6603
		(.075)	(.18)	(.22)	(.33)		(.12)	(.14)	(.063)	(.099)	(.13)	(.11)	(.17)		(.081)	(.092)	(.22)	(.19)	(.18)	(.30)	(.15)	(.007)	(.80)						
⟨16-18⟩	females	.28	.46	−.70	.17		−.042	1.62	.55	.36	−.42	−.23	−.13		−.057	.57	.93	.84	.85	.17	1.04	−.074	6.64	✓			CMA	⟨2⟩	18586
		(.044)	(.15)	(.18)	(.23)		(.11)	(.081)	(.043)	(.072)	(.13)	(.11)	(.17)		(.058)	(.054)	(.17)	(.15)	(.15)	(.30)	(.13)	(.006)	(.61)						
(19)	females																								✓		CSD	E2	0
							(0)																						
(20)	females	.21	.42	−.73			−.099	1.82	.56	.41					.18	.68	.99	.99	.84		1.12	−.063	6.70		✓		CSD	ED	11343
		(.060)	(.29)	(.34)			(100.0)	(.13)	(.078)	(.13)					(.10)	(.095)	(.27)	(.26)	(.28)		(.28)	(.012)	(1.27)						
(21)	females	.43	.48	−.66			.25	1.27	.54	.39	−.46	−.26	−.15		−.25	.40	.90	.79	.83	.19	.99	−.084	7.32		✓		CSD	G17	5844
		(.081)	(.22)	(.25)			(.12)	(.10)	(.085)	(.12)	(.14)	(.11)	(.19)		(.081)	(.081)	(.22)	(.21)	(.21)	(.28)	(.24)	(.012)	(1.35)						
⟨19-21⟩	females	.29	.46	−.69			.25	1.50	.55	.40	−.46	−.26	−.15		−.080	.52	.93	.87	.83	.19	1.05	−.073	6.99		✓		CSD	⟨2⟩	17187
		(.048)	(.17)	(.20)			(.12)	(.080)	(.057)	(.088)	(.14)	(.11)	(.19)		(.063)	(.062)	(.17)	(.16)	(.17)	(.28)	(.18)	(.008)	(.92)						
(22)	females																									✓	CT	E2	0
							(0)																						
(23)	females																									✓	CT	ED	0
							(0)																						
(24)	females																									✓	CT	G17	0
							(0)																						
⟨22-24⟩	females																									✓	CT	⟨0⟩	0
(25)	single	.27	.055	.18	.031	−2.16	−1.62	.70		.40		−.18	.12	.11		.056				−1.11		−.045	6.12					E2	613
		(.27)	(.71)	(.81)	(1.26)	(1.65)	(1.56)	(.23)		(.20)		(.25)	(.33)	(.16)		(.24)				(.32)		(.025)	(2.50)						
(26)	single	.19	.17	−.48	−.22	−.66	−1.00	1.64		.46				.0009	−.035	.47	1.33	1.24	.89		1.27	−.063	6.76					ED	11699

Continued on next page

		log(HH inc)	DA: log(HH inc)	CT: log(HH inc)	CSD: log(HH inc)	CMA: log(HH inc)	$\sum\beta_{\mathrm{inc}}$	trust-N	married	asmarried	separated	divorced	widowed	male	noReligion	godImportance	student	employed	domestic	unemployed	retired	age	$(\mathrm{age}/100)^2$	CMA f.e.	CSD f.e.	CT f.e.	clustering	survey	obs.
		(.058)	(.28)	(.35)	(.47)	(.54)	(.50)	(.13)		(.092)				(.058)	(.087)	(.10)	(.17)	(.15)	(.22)		(.21)	(.011)	(1.29)						
(27)	single	.41	.53	−.60	−.80	−.53	−.99	1.22		.47	−.37	−.061	−.10	−.077	−.27	.32	1.26	1.18	.81	.33	.94	−.12	12.1					G17	6700
		(.073)	(.23)	(.28)	(.42)	(.51)	(.77)	(.11)		(.078)	(.11)	(.093)	(.15)	(.055)	(.062)	(.083)	(.24)	(.23)	(.26)	(.29)	(.26)	(.011)	(1.26)						
⟨25-27⟩	single	.27	.37	−.50	−.51	−.66	−1.04	1.31		.46	−.37	−.075	−.065	−.031	−.19	.36	1.31	1.22	.85	−.31	1.14	−.088	9.10					⟨3⟩	19012
		(.045)	(.17)	(.21)	(.30)	(.36)	(.41)	(.081)		(.057)	(.11)	(.087)	(.14)	(.039)	(.050)	(.062)	(.14)	(.13)	(.17)	(.22)	(.16)	(.007)	(.85)						
(28)	single																							✓			CMA	E2	0
							(0)																						
(29)	single	.18	.15	−.47	.14		−.008	1.72		.43				−.008	.053	.55	1.32	1.25	.90		1.23	−.068	7.28	✓			CMA	ED	11699
		(.047)	(.16)	(.27)	(.36)		(100.0)	(.11)		(.063)				(.055)	(.088)	(.094)	(.21)	(.18)	(.34)		(.22)	(.009)	(.98)						
(30)	single	.40	.50	−.58	−.71		−.38	1.22		.45	−.35	−.050	−.10	−.071	−.22	.35	1.26	1.19	.81	.33	.97	−.12	12.1	✓			CMA	G17	6700
		(.075)	(.20)	(.29)	(.47)		(100.0)	(.16)		(.078)	(.10)	(.058)	(.13)	(.036)	(.051)	(.057)	(.32)	(.31)	(.29)	(.42)	(.31)	(.010)	(1.28)						
⟨28-30⟩	single	.24	.29	−.52	−.17		−.19	1.56		.44	−.35	−.050	−.10	−.052	−.15	.41	1.30	1.23	.85	.33	1.14	−.093	9.06	✓			CMA	⟨2⟩	18399
		(.040)	(.13)	(.20)	(.29)		(70.7)	(.093)		(.049)	(.10)	(.058)	(.13)	(.030)	(.044)	(.049)	(.17)	(.16)	(.22)	(.42)	(.18)	(.007)	(.78)						
(31)	single																								✓		CSD	E2	0
							(0)																						
(32)	single	.17	.082	−.38			−.13	1.78		.41				−.0003	.060	.55	1.29	1.23	.85		1.17	−.076	8.25		✓		CSD	ED	11000
		(.060)	(.27)	(.30)			(100.0)	(.13)		(.092)				(.055)	(.089)	(.093)	(.21)	(.21)	(.31)		(.24)	(.012)	(1.40)						
(33)	single	.38	.58	−.64			.31	1.25		.48	−.47	−.056	−.11	−.076	−.18	.37	1.20	1.16	.66	.29	.93	−.12	12.1		✓		CSD	G17	5987
		(.086)	(.24)	(.26)			(.13)	(.10)		(.089)	(.11)	(.081)	(.17)	(.050)	(.065)	(.077)	(.24)	(.23)	(.26)	(.33)	(.29)	(.012)	(1.46)						
⟨31-33⟩	single	.24	.36	−.52			.31	1.46		.45	−.47	−.056	−.11	−.042	−.095	.44	1.25	1.20	.74	.29	1.07	−.098	10.1		✓		CSD	⟨2⟩	16987
		(.049)	(.18)	(.20)			(.13)	(.080)		(.064)	(.11)	(.081)	(.17)	(.037)	(.052)	(.059)	(.16)	(.15)	(.20)	(.33)	(.19)	(.008)	(1.01)						
(34)	single																									✓	CT	E2	0
							(0)																						
(35)	single																									✓	CT	ED	0
							(0)																						
(36)	single																									✓	CT	G17	0
							(0)																						
⟨34-36⟩	single																									✓	CT	⟨0⟩	0
(37)	married	.34	−.10	.41	−2.90	1.04	−1.23	.38						−.15		.72				−1.26		−.001	2.10					E2	528
		(.32)	(.71)	(.98)	(1.52)	(1.78)	(100.0)	(.27)						(.16)		(.28)				(.92)		(.041)	(3.98)						
(38)	married	.32	.020	−.14	.49	−1.54	−.85	1.84						−.18	.071	.67	.82	1.17	1.10		1.51	−.044	4.41					ED	11890
		(.087)	(.28)	(.33)	(.44)	(.54)	(.65)	(.14)						(.058)	(.093)	(.092)	(.29)	(.23)	(.25)		(.25)	(.016)	(1.65)						
(39)	married	.79	.32	−.41	−.30	−1.45	−1.05	1.34						−.22	−.17	.39	1.23	1.42	1.65	.46	1.71	−.060	4.67					G17	5508
		(.13)	(.28)	(.33)	(.43)	(.54)	(100.0)	(.13)						(.059)	(.076)	(.091)	(.36)	(.31)	(.32)	(.39)	(.33)	(.016)	(1.70)						
⟨37-39⟩	married	.46	.15	−.24	−.032	−1.39	−.85	1.45						−.19	−.075	.54	.98	1.26	1.30	.20	1.59	−.048	4.34					⟨3⟩	17926

Continued on next page

	log(HH inc)	DA: log(HH inc)	CT: log(HH inc)	CSD: log(HH inc)	CMA: log(HH inc)	$\sum\beta_{\text{inc}}$	trust-N	married	asmarried	separated	divorced	widowed	male	noReligion	godImportance	student	employed	domestic	unemployed	retired	age	$(\text{age}/100)^2$	CMA f.e.	CSD f.e.	CT f.e.	clustering	survey	obs.
	(.071)	(.19)	(.23)	(.30)	(.37)	(.65)	(.090)						(.040)	(.059)	(.063)	(.22)	(.19)	(.20)	(.36)	(.20)	(.011)	(1.14)						
(40) married																							✓			CMA	E2	0
						(0)																						
(41) married	.32	−.021	−.061	.33		.57	1.92						−.18	.13	.68	.81	1.16	1.11		1.50	−.046	4.58	✓			CMA	ED	11890
	(.089)	(.30)	(.38)	(.28)		(.23)	(.13)						(.059)	(.082)	(.076)	(.22)	(.17)	(.20)		(.20)	(.016)	(1.60)						
(42) married	.79	.32	−.44	−.14		.54	1.32						−.22	−.17	.40	1.25	1.46	1.68	.48	1.73	−.065	5.09	✓			CMA	G17	5494
	(.16)	(.25)	(.23)	(.56)		(.34)	(.16)						(.052)	(.077)	(.12)	(.32)	(.24)	(.25)	(.33)	(.26)	(.012)	(1.28)						
⟨40-42⟩ married	.44	.18	−.34	.24		.56	1.67						−.20	−.026	.60	.95	1.26	1.33	.48	1.58	−.058	4.89	✓			CMA	⟨2⟩	17384
	(.077)	(.19)	(.20)	(.25)		(.19)	(.10)						(.039)	(.056)	(.065)	(.18)	(.14)	(.15)	(.33)	(.16)	(.010)	(1.00)						
(43) married																								✓		CSD	E2	0
						(0)																						
(44) married	.38	−.16	.15			.38	1.98						−.19	.17	.71	.84	1.11	1.13		1.56	−.051	5.10		✓		CSD	ED	11269
	(.090)	(.27)	(.30)			(.17)	(.12)						(.068)	(.096)	(.082)	(.28)	(.21)	(.24)		(.27)	(.017)	(1.76)						
(45) married	.75	.34	−.45			.64	1.35						−.22	−.12	.41	1.17	1.34	1.57	.35	1.59	−.059	4.58		✓		CSD	G17	4739
	(.15)	(.28)	(.30)			(.20)	(.14)						(.054)	(.071)	(.089)	(.40)	(.36)	(.36)	(.46)	(.39)	(.017)	(1.84)						
⟨43-45⟩ married	.48	.087	−.15			.48	1.70						−.21	−.021	.57	.95	1.17	1.27	.35	1.57	−.055	4.85		✓		CSD	⟨2⟩	16008
	(.077)	(.19)	(.21)			(.13)	(.093)						(.042)	(.057)	(.060)	(.23)	(.18)	(.20)	(.46)	(.22)	(.012)	(1.27)						
(46) married																									✓	CT	E2	0
						(0)																						
(47) married	.31	1.01				1.32	2.39						−.15	.14	.65	3.31	2.55	2.58		3.27	−.047	3.93			✓	CT	ED	967
	(.41)	(1.38)				(1.29)	(.52)						(.21)	(.36)	(.31)	(1.21)	(.87)	(.94)		(1.03)	(.065)	(7.07)						
(48) married																									✓	CT	G17	0
						(0)																						
⟨46-48⟩ married	.31	1.01				1.32	2.39						−.15	.14	.65	3.31	2.55	2.58		3.27	−.047	3.93			✓	CT	⟨1⟩	967
	(.41)	(1.38)				(1.29)	(.52)						(.21)	(.36)	(.31)	(1.21)	(.87)	(.94)		(1.03)	(.065)	(7.07)						
(49) employed	.27	.16	−.33	−.086	−1.21	−1.20	1.49	.47	.46				−.083	.036	.58						−.072	7.85					ED	14418
	(.076)	(.25)	(.30)	(.40)	(.48)	(.50)	(.12)	(.073)	(.11)				(.050)	(.079)	(.085)						(.016)	(1.87)						
(50) employed	.70	.26	−.45	−.68	−.95	−1.12	1.34	.55	.48	−.25	−.071	.14	−.14	−.27	.30						−.15	15.9					G17	7520
	(.13)	(.24)	(.28)	(.40)	(.48)	(1.07)	(.11)	(.084)	(.092)	(.13)	(.11)	(.29)	(.049)	(.061)	(.079)						(.016)	(1.89)						
⟨49-50⟩employed	.38	.21	−.39	−.38	−1.08	−1.18	1.41	.51	.47	−.25	−.071	.14	−.11	−.16	.43						−.11	11.9					⟨2⟩	21938
	(.066)	(.17)	(.20)	(.28)	(.34)	(.46)	(.082)	(.055)	(.070)	(.13)	(.11)	(.29)	(.035)	(.048)	(.058)						(.011)	(1.33)						
(51) employed	.25	.20	−.35	.065		.17	1.58	.50	.41				−.089	.12	.63						−.076	8.16	✓			CMA	ED	14418
	(.065)	(.19)	(.23)	(.38)		(.21)	(.12)	(.082)	(.089)				(.047)	(.11)	(.11)						(.017)	(2.02)						
(52) employed	.71	.24	−.46	−.57		−.072	1.33	.56	.46	−.25	−.072	.16	−.13	−.26	.31						−.16	16.3	✓			CMA	G17	7520
	(.10)	(.28)	(.31)	(.32)		(.025)	(.17)	(.059)	(.085)	(.10)	(.10)	(.28)	(.048)	(.052)	(.083)						(.015)	(1.80)						
⟨51-52⟩employed	.39	.22	−.39	−.31		−.069	1.49	.54	.44	−.25	−.072	.16	−.11	−.19	.42						−.12	12.7	✓			CMA	⟨2⟩	21938

Continued on next page

	log(HH inc)	DA: log(HH inc)	CT: log(HH inc)	CSD: log(HH inc)	CMA: log(HH inc)	$\sum \beta_{inc}$	trust-N	married	asmarried	separated	divorced	widowed	male	noReligion	godImportance	student	employed	domestic	unemployed	retired	age	$(age/100)^2$	CMA f.e.	CSD f.e.	CT f.e.	clustering	survey	obs.
	(.054)	(.16)	(.19)	(.24)		(.025)	(.100)	(.048)	(.061)	(.10)	(.10)	(.28)	(.034)	(.047)	(.066)						(.011)	(1.34)						
(53) employed	.27	.15	−.24			.19	1.62	.51	.40				−.096	.15	.64						−.082	8.89		✓		CSD	ED	13782
	(.083)	(.23)	(.28)			(.10)	(.12)	(.075)	(.11)				(.050)	(.081)	(.075)						(.017)	(1.96)						
(54) employed	.72	.19	−.48			.43	1.42	.52	.45	−.35	−.13	.14	−.17	−.24	.33						−.15	16.1		✓		CSD	G17	6729
	(.13)	(.27)	(.23)			(.14)	(.11)	(.087)	(.12)	(.13)	(.13)	(.29)	(.049)	(.057)	(.074)						(.018)	(2.05)						
⟨53-54⟩employed	.40	.17	−.38			.27	1.52	.52	.42	−.35	−.13	.14	−.13	−.11	.48						−.12	12.4		✓		CSD	⟨2⟩	20511
	(.070)	(.18)	(.18)			(.084)	(.082)	(.057)	(.080)	(.13)	(.13)	(.29)	(.035)	(.047)	(.053)						(.012)	(1.42)						
(55) employed	.12	.39				.52	2.13	.64	.59				−.14	.17	.70						−.056	6.13			✓	CT	ED	1551
	(.22)	(.93)				(.78)	(.34)	(.27)	(.52)				(.18)	(.24)	(.29)						(.056)	(6.61)						
(56) employed																									✓	CT	G17	0
						(0)																						
⟨55-56⟩employed	.12	.39				.52	2.13	.64	.59				−.14	.17	.70						−.056	6.13			✓	CT	⟨1⟩	1551
	(.22)	(.93)				(.78)	(.34)	(.27)	(.52)				(.18)	(.24)	(.29)						(.056)	(6.61)						

	log(HH inc)	DA: log(HH inc)	CT: log(HH inc)	CSD: log(HH inc)	CMA: log(HH inc)	health	CT: health	CSD: health	CMA: health	married	asmarried	separated	divorced	widowed	male	noReligion	godImportance	student	employed	domestic	unemployed	retired	age	$(age/100)^2$	f.e./clustering	survey	obs.	pseudo-R^2	$N_{clusters}$
(9)	.22	.16	.096	−1.14	−.69	1.14	.69	−1.55	1.36	.53	.26		−.34	.25	−.020		.45				−1.23		.020	.025		E2	1151	.035	
	(.20)	(.50)	(.61)	(1.03)	(1.24)	(.32)	(.47)	(1.49)	(1.97)	(.17)	(.21)		(.23)	(.31)	(.11)		(.17)				(.31)		(.004)	(.005)					
(10)	.36	.52	−.43	−.34	−1.29	2.85	.19	.008	1.15	.45	.35	−.45	−.11	−.042	−.16	−.18	.38	.68	.56	.65	−.17	.87	−.084	8.69		G17	12544	.058	
	(.060)	(.17)	(.21)	(.31)	(.38)	(.093)	(.16)	(.52)	(.74)	(.066)	(.075)	(.10)	(.086)	(.13)	(.039)	(.047)	(.061)	(.17)	(.16)	(.18)	(.21)	(.18)	(.008)	(.92)					
⟨9-10⟩	.35	.49	−.37	−.41	−1.24	2.72	.24	−.16	1.18	.46	.34	−.45	−.14	.001	−.15	−.18	.39	.68	.56	.65	−.50	.87	−.0008	.025		⟨2⟩	13695		
	(.057)	(.16)	(.20)	(.30)	(.36)	(.089)	(.15)	(.49)	(.69)	(.061)	(.071)	(.10)	(.081)	(.12)	(.037)	(.047)	(.058)	(.17)	(.16)	(.18)	(.17)	(.18)	(.004)	(.005)					
(11)	.25	.29	−.014	−.56		.95	.77	−1.35		.59	.16		−.25	.21	−.11		.51				−1.06		.017	.022	CMA	E2	1044	.038	23
	(.20)	(.42)	(.55)	(1.39)		(.31)	(.44)	(1.91)		(.17)	(.27)		(.41)	(.23)	(.077)		(.14)				(.41)		(.005)	(.007)					
(12)	.35	.53	−.45	−.28		2.86	.17	.004		.45	.34	−.45	−.10	−.055	−.16	−.16	.40	.67	.55	.66	−.18	.87	−.085	8.76	CMA	G17	12544	.060	46
	(.058)	(.16)	(.18)	(.25)		(.092)	(.16)	(.53)		(.042)	(.062)	(.084)	(.063)	(.12)	(.038)	(.033)	(.063)	(.17)	(.16)	(.15)	(.22)	(.12)	(.006)	(.58)					
⟨11-12⟩	.34	.50	−.41	−.29		2.71	.24	−.094		.46	.33	−.45	−.11	.005	−.15	−.16	.42	.67	.55	.66	−.38	.87	−.025	.023	CMA	⟨2⟩	13588		
	(.056)	(.15)	(.17)	(.25)		(.088)	(.15)	(.51)		(.041)	(.060)	(.084)	(.062)	(.11)	(.034)	(.033)	(.058)	(.17)	(.16)	(.15)	(.19)	(.12)	(.004)	(.007)					
(13)	.38	.83	−.91			.86	.69			.68	.33		−.26	.63	−.12		.51				−1.00		.018	.023	CSD	E2	814	.046	23
	(.21)	(.53)	(.52)			(.44)	(.61)			(.18)	(.38)		(.36)	(.24)	(.12)		(.18)				(.36)		(.005)	(.006)					
(14)	.34	.53	−.41			2.89	.097			.43	.30	−.46	−.12	−.099	−.17	−.13	.41	.58	.47	.57	−.23	.77	−.086	8.95	CSD	G17	11782	.064	188
	(.064)	(.18)	(.17)			(.10)	(.18)			(.061)	(.082)	(.10)	(.075)	(.14)	(.042)	(.044)	(.056)	(.18)	(.16)	(.16)	(.22)	(.17)	(.008)	(.89)					
⟨13-14⟩	.34	.56	−.46			2.78	.14			.46	.31	−.46	−.12	.083	−.16	−.13	.42	.58	.47	.57	−.44	.77	−.011	.023	CSD	⟨2⟩	12596		
	(.061)	(.17)	(.16)			(.10)	(.17)			(.058)	(.080)	(.10)	(.073)	(.12)	(.040)	(.044)	(.054)	(.18)	(.16)	(.16)	(.19)	(.17)	(.004)	(.006)					
(15)																									CT	E2	0		
(16)	.65	.16				3.39				.34	.59	−1.05	−.023	−.19	−.19	−.096	.69	.63	.33	.049	.18	.20	−.081	8.07	CT	G17	1474	.091	119
	(.20)	(.52)				(.35)				(.25)	(.26)	(.37)	(.35)	(.47)	(.15)	(.17)	(.18)	(.60)	(.52)	(.56)	(.59)	(.48)	(.032)	(3.44)					
⟨15-16⟩	.65	.16				3.39				.34	.59	−1.05	−.023	−.19	−.19	−.096	.69	.63	.33	.049	.18	.20	−.081	8.07	CT	⟨1⟩	1474		
	(.20)	(.52)				(.35)				(.25)	(.26)	(.37)	(.35)	(.47)	(.15)	(.17)	(.18)	(.60)	(.52)	(.56)	(.59)	(.48)	(.032)	(3.44)					

Significance: 1% 5% 10%

Table A.6: Detailed regressions for spillover effects of others' health. Standard controls are not shown explicitly. These results are summarised in Table 2.13 on page 31.

A.2 Survey descriptions, consistency, and summary statistics

This appendix provides a qualitative look at some of the key survey variables with the aim of assessing the consistency of results from different surveys.

A.2.1 Survey descriptions

We make use of three surveys conducted across Canada: the second wave of the Equality, Security, and Community survey (ESC2) from 2002-2003, described by *Soroka et al.* [2007] and online at `http://grad.econ.ubc.ca/cpbl/esc2`; the Ethnic Diversity Survey (EDS) from 2002; and the General Social Survey Cycle 17 (GSS) from 2003. The latter two surveys are described in detail on Statistics Canada's web site. See also *Helliwell and Huang* [2005] for some further description of, and differences between, these surveys. Survey data at the level of the individual respondent were accessed through Statistics Canada's Research Data Centre located at UBC.

The surveys comprise a total of $\sim$70,000 individuals and they have some key questions in common. Most importantly, respondents were asked to rate their overall life satisfaction on a 5 or 10 point scale. Figure A.1 on page 131 shows how the responses to these questions were distributed in each survey. Responses from the EDS' five point scale are shown rescaled to range between 1 and 10 as in the other surveys. Also included for comparison is the distribution from the 2005 General Social Survey, Cycle 19. The shape of the distribution is remarkably repeatable between the two surveys, GSS17 and GSS19, which have the most similar sequence of questions on subjective well-being, although they may indicate a significant decrease in average reported satisfaction. Life satisfaction reportde in the ESC2, at about the same time as the GSS17, appears to be similarly distributed except for the higher preponderance of fully satisfied respondents. Not surprisingly, the less detailed scale used in the EDS cannot resolve the features evident in the other surveys, yet it nevertheless indicates a similar mean scaled response as the others.

Numerous other questions relevant to social interactions and socioeconomic and cultural backgrounds were posed in these surveys. Table A.7 on page 130 shows the availability of some of these measures and compares the question wording used. In some cases, such as for the important measure of trust in neighbours, different questions were asked but, after being scaled, will be used as equivalent measures in our analysis.

Some differences between responses concerning trust and life satisfaction in GSS17 and GSS19 could be due to the order of modules in the questionnaires. In GS17, the well-being module is asked in the initial section, while in GSS19 a similar section of modules is asked in the middle of the survey, after details of the time use diary, unpaid work, and childcare had been covered.

There were also some notable sampling differences between the surveys.

> Data for Cycle 19 of the GSS were collected in 11 monthly samples from January to November 2005 with data collection for the November sample extending until mid-December. The sample was evenly distributed over the 11 months. Questions asked as part of the survey had a variety of reference periods, such as the past week, the past 12 months, and the past 5 years.

EDS was a post-censal survey. The target population was a subset (the majority) of those who were selected to answer the long form of the 2001 census questionnaire and represents 23 million Canadians. However, the sample selection was based on a stratification by ethnic origin, place of birth and place of birth of parents, rather than the geographic distribution behind the GSS sample selection. The EDS is therefore generally unsuitable for use in fine-scale geographic analysis.

A.2.2 Consistency of place-based characteristics

A key feature of all the surveys used is the availability of high resolution in the geographic location of respondents' places of residence. Because our work relies on the possibility that significant determinants of life satisfaction are rooted in geographic locations, this section assesses the repeatability of these features over time and between surveys.

A.2.2.1 Trust in neighbours

Figure A.2 on page 132 shows that questions about trust in neighbours elicit differences between provinces that are consistent from one survey to another. The agreement between GSS17 and EDS, in particular, is very close and it may be noted that these surveys pose the question in the same way (Table A.7 on page 130) as a five-point subjective assessment. The ESC2 survey asks a much more specific question which concerns the likelihood, on a three-point scale, of a neighbour returning a lost wallet. Nevertheless, responses from this measure are still very strongly correlated with those from the other surveys. Figures A.3 and A.4 show the same correlations at the CMA and CSD level and show a similar consistency. Taken together, these comparisons for different geographic scales suggest transforming the trust in neighbours means $\langle T\rangle_{\text{ESC2}}$ from ESC2 in order to be more comparable to those from the GSS17 and EDS:

$$\langle T\rangle_{\text{rescaled}} = -.51 \pm .07 + \langle T\rangle_{\text{ESC2}}\,(1.72 \pm 0.10)$$

A.2.2.2 Life satisfaction

Figures A.5 to A.7 present the analogous set of correlations for survey responses to the life satisfaction question, aggregated by province, CMA, and CSD. In comparison with the measure of trust there is less correlation between surveys for life satisfaction. It can be seen, especially in the case of CSDs, that the correlation is higher amongst regions with larger populations (represented by dark dots). This suggests that the relatively poor correlation may simply reflect the well recognised large individual variance in life satisfaction that results from aspects of personality and other factors which are non-geographic. In addition, however, there are as noted above significant differences in the mean reported life satisfaction between surveys, especially in the case of the EDS for which the question was given with a five-point scale.

These findings have several implications for statistical methods. First, the large variance of life satisfaction at the individual level combined with relatively precise measurements of geographical means of some of its correlates, such as trust in neighbours, suggests that there is a large advantage to using individual level data in regressions for life satisfaction when testing for effects of place-based determinants. This is indeed our primary method in this study. It also suggests that large sample sizes are likely to be necessary when estimating the effects of variables at small geographic scales.

Secondly, for any variables with significant dispersion of means from smaller regions, such as is evident for life satisfaction, aggregation of data from different surveys may prove useful in reducing standard errors of geographic means. In the section to follow, this is applied to improve the precision of estimates and rankings of Canadian regions by their mean life satisfaction.

Thirdly, estimates of mean life satisfaction should be adjusted for survey means before further aggregation. This is also demonstrated in Section A.2.3.

A.2.2.3 Other variables

Figures A.8 to A.19show similar survey comparisons for geographic means of other variables. In general, regions with higher populations and therefore sample sizes show more consistent results. Most variables are available for only a subset of the three surveys.

A.2.3 Variation across geographical regions

Figures A.20 to A.25 show the range of multi-survey means for life satisfaction and trust in neighbours calculated at the spatial scales of province and CMA.[1] In each figure, each dark horizontal bar shows the estimated mean and its standard error for one region, which is named to the left. Each mean V and standard error S are calculated by taking a weighted mean over the results from individual surveys,

$$\begin{aligned} V_r &= \frac{\sum_i \frac{v_{ir}}{\sigma_{ir}^2}}{\sum_i \frac{1}{\sigma_{ir}^2}} \qquad \text{(A.1)} \\ S_r &= \sqrt{\frac{1}{\sum_i \frac{1}{\sigma_{ir}^2}}} \end{aligned}$$

[1]Means calculated at the CSD level are available from the author.

where v_{ir} and σ_{ir} indicate the mean and standard error over region r from survey i. The surveys involved in each mean are listed to the right of the plots and the individual survey values are shown as light horizontal bars just above the dark bar corresponding to their mean. All bars are four standard deviations (standard error of the mean) wide and therefore indicate 95% confidence intervals.

In the case of reported life satisfaction, these charts are shown in two forms. In the second of each pair, the geographic averages from each survey have been adjusted to remove the differences in overall means from each of the three surveys, before being aggregated as in equation (A.1). As discussed above, this is especially useful considering the difference in the way life satisfaction was measured in the EDS survey. With this correction, there is in general a good consistency between different surveys, especially for large regions. This again indicates that weak correlation between surveys for life satisfaction is largely a result of the high degree of what might be measured as individual fixed effects in panel data; with adequate sample sizes, significant geographic differences in average life satisfaction are measureable and reproducible.

This geographic variation is even more evident for aggregated trust in neighbours, as shown in Figures A.24 to A.25.

A.2.4 Life satisfaction rankings based on ESC2 alone

The ESC2 survey generated somewhat more variation between geographical regions in reported life satisfaction than GSS17 or EDS. The remaining figures show the variation and standard errors based on this survey alone, which made its way into the public press in early 2008 as a ranking of Canadian cities by life satisfaction. Here it is evident that without inclusion of survey data from other surveys, the differences between CMA regions are barely signficant.

Table A.7: Summary of survey variable definitions.

Variable	GSS17	ESC1	ESC2	EDS	GSS19	
lsatis	√	√	√	√	√	**GSS17**: Using the same scale, how do you feel about your life as a whole right now? **ESC1**: ? **ESC2**: Now a question about life satisfaction. On a scale of 1-10 where ONE means dissatisfied and TEN means satisfied, all things considered how satisfied are you with your life as a whole these days? **EDS**: Using a scale of 1 to 5, where 1 means not satisfied at all and 5 means very satisfied. All things considered, how satisfied are you with your life as a whole these days? **GSS19**: Using the same scale, how do you feel about your life as a whole right now?
trustNeighbour	√	√	√	√	√	**GSS17**: [scaled]: Using a scale of 1 to 5 where 1 means 'Cannot be trusted at all' and 5 means 'Can be trusted a lot', how much do you trust each of the following groups of people: .people in your neighbourhood? **ESC1**: If you lost a wallet or a purse that contained two hundred dollars, how likely is it to be returned with the money in it if it was found by someone who lives close by; would you say very likely, somewhat likely or not at all likely? **ESC2**: If you lost a wallet or a purse that contained two hundred dollars, how likely is it to be returned with the money in it if it was found by someone who lives close by; would you say very likely, somewhat likely or not at all likely? **EDS**: [scaled]: Using a scale of 1 to 5 where 1 means cannot be trusted at all and 5 means can be trusted a lot, how much do you trust each of the following groups of people: People in your neighborhood? **GSS19**: [scaled]: Using a scale of 1 to 5 where 1 means 'Cannot be trusted at all' and 5 means 'Can be trusted a lot', how much do you trust each of the following groups of people:... people in your neighbourhood?
trustBool	√	√	√	√	√	**GSS17**: Generally speaking, would you say that most people can be trusted or that you cannot be too careful in dealing with people? **ESC1**: Now some questions about how much you trust other people. Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people? **ESC2**: Now some questions about how much you trust other people. Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people? **EDS**: Generally speaking, would you say that most people can be trusted or that you cannot be too careful in dealing with people? **GSS19**: [scaled]: Generally speaking, would you say that most people can be trusted or that you cannot be too careful in dealing with people?
godImportance	√	√	√	√	√	**GSS17**: [scaled]: How important are your religious or spiritual beliefs to the way that you live your life? Would you say it is: **ESC1**: [scaled]: How important is religion in your life? Would you say very important, somewhat important, not very important, or not important at all? **ESC2**: [scaled]: How important is religion in your life? Would you say very important, somewhat important, not very important, or not important at all? **EDS**: [scaled]: Using a scale of 1 to 5, where 1 is not important at all and 5 is very important, how important is your religion to you? **GSS19**: [scaled]: How important are your (religious or) spiritual beliefs to the way you live your life? Would you say they are:
godParticipateFrequency	√	√	√		√	**GSS17**: Other than on special occasions, (such as weddings, funerals or baptisms) how often did you attend religious services or meetings in the last 12 months? Was it: **ESC1**: How often do you attend religious services, NOT including weddings and funerals? **ESC2**: How often do you attend religious services, NOT including weddings and funerals? **GSS19**: Religious attendance of the respondent.
happy	√	√			√	**GSS17**: [scaled]: Presently, would you describe yourself as: **ESC1**: [scaled]: Now a question about life satisfaction. On a scale of 1-10 where ONE means dissatisfied and TEN means satisfied, all things considered how satisfied are you with your life as a whole these days? **GSS19**: [scaled]: Presently, would you describe yourself as:
health	√	√	√		√	**GSS17**: [scaled]: In general, would you say your health is: **ESC1**: [scaled]: How would you describe your health these days, would you say: poor, fair, good, very good, or excellent? **ESC2**: [scaled]: How would you describe your health these days, would you say: poor, fair, good, very good, or excellent? **GSS19**: [scaled]: In general, would you say your health is:

Table A.7: Summary of survey variable definitions.

Variable	GSS17	ESC1	ESC2	EDS	GSS19	
logTenureHouse	$\surd$		$\surd$		$\surd$	**GSS17**: How long have you lived in this dwelling? **ESC2**: How many years have you lived at your current address? **GSS19**: How long have you lived in this dwelling?
motherSchoolingYears	$\surd$	$\surd$	$\surd$		$\surd$	**GSS17**: Highest level of education obtained by the respondent's mother - 10 groups. **ESC1**: What is the highest level of education that your MOTHER completed? **ESC2**: What is the highest level of education that your MOTHER completed? **GSS19**: Highest level of education obtained by the respondent's mother - 10 groups.
fatherSchoolingYears	$\surd$	$\surd$	$\surd$		$\surd$	**GSS17**: Highest level of education obtained by the respondent's father - 10 groups. **ESC1**: What about your FATHER, what is the highest level of education he completed? **ESC2**: What about your FATHER, what is the highest level of education he completed? **GSS19**: Highest level of education obtained by the respondent's father - 10 groups.
belongCommunity	$\surd$				$\surd$	**GSS17**: [scaled]: How would you describe your sense of belonging to your local community? Would you say it is: **GSS19**: [scaled]: How would you describe your sense of belonging to your local community? Would you say it is:
belongCountry	$\surd$			$\surd$	$\surd$	**GSS17**: [scaled]: What about (your sense of belonging) to Canada? **EDS**: [scaled]: Some people have a stronger sense of belonging to some things than others. Using a scale of 1 to 5, where 1 is not strong at all and 5 is very strong, how strong is your sense of belonging to Canada? **GSS19**: [scaled]: What about (your sense of belonging) to Canada?
belongEthnicity				$\surd$		**EDS**: [scaled]: Some people have a stronger sense of belonging to some things than others. Using a scale of 1 to 5, where 1 is not strong at all and 5 is very strong, how strong is your sense of belonging to your ethnic or cultural group(s)?
belongFamily				$\surd$		**EDS**: [scaled]: Some people have a stronger sense of belonging to some things than others. Using a scale of 1 to 5, where 1 is not strong at all and 5 is very strong, how strong is your sense of belonging to your family?
belongProvince	$\surd$			$\surd$	$\surd$	**GSS17**: [scaled]: What about (your sense of belonging) to your province? **EDS**: [scaled]: Some people have a stronger sense of belonging to some things than others. Using a scale of 1 to 5, where 1 is not strong at all and 5 is very strong, how strong is your sense of belonging to your province? **GSS19**: [scaled]: What about (your sense of belonging) to your province?
belongTown				$\surd$		**EDS**: [scaled]: Some people have a stronger sense of belonging to some things than others. Using a scale of 1 to 5, where 1 is not strong at all and 5 is very strong, how strong is your sense of belonging to your town, city or municipality?
commutingWeekly	$\surd$					**GSS17**: (RAW CODEBOOK INFO MISSING)
confidenceBanks	$\surd$					**GSS17**: [scaled]: How much confidence do you have in: .banks?
confidenceBigCorps	$\surd$					**GSS17**: [scaled]: How much confidence do you have in: .major corporations?
confidenceHealthcare	$\surd$					**GSS17**: [scaled]: How much confidence do you have in: .the health care system?
confidenceJustice	$\surd$					**GSS17**: [scaled]: How much confidence do you have in: .the justice system and courts?
confidenceLocalCorps	$\surd$					**GSS17**: [scaled]: How much confidence do you have in: .local merchants and business people?
confidenceParliament	$\surd$					**GSS17**: [scaled]: How much confidence do you have in: .federal parliament?
confidencePolice	$\surd$	$\surd$	$\surd$			**GSS17**: [scaled]: How much confidence do you have in: .the police? **ESC1**: If you lost a wallet or a purse that contained two hundred dollars, how likely is it to be returned with the money in it if it was found by a police officer; would you say very likely, somewhat likely or not at all likely? **ESC2**: If you lost a wallet or a purse that contained two hundred dollars, how likely is it to be returned with the money in it if it was found by a police officer; would you say very likely, somewhat likely or not at all likely?
confidenceSchools	$\surd$					**GSS17**: [scaled]: How much confidence do you have in: .the school system?
confidenceWelfare	$\surd$					**GSS17**: [scaled]: How much confidence do you have in: .the welfare system?
ethnicHeterophile	$\surd$					**GSS17**: [scaled]: Using a scale of 1 to 5, where 1 is not important at all and 5 is very important, how important is it for you to establish and maintain ties: .with people who have different ethnic or cultural origins than you?

Table A.7: Summary of survey variable definitions.

Variable	GSS17	ESC1	ESC2	EDS	GSS19	
ethnicHomophile	$\surd$					**GSS17**: [scaled]: Using a scale of 1 to 5, where 1 is not important at all and 5 is very important, how important is it for you to establish and maintain ties: . with other people who have similar ethnic or cultural origin as you?
ethnicImportance				$\surd$		**EDS**: Maximum of reported importances of ethnicity
foreignBorn				$\surd$		**EDS**: Derived - Place of birth - inside or outside Canada
gaySpouse	$\surd$				$\surd$	**GSS17**: Type of partner the respondent has within the household. **GSS19**: Type of partner the respondent has within the household.
healthBadSleep	$\surd$				$\surd$	**GSS17**: [bool,.=0]: Do you regularly have trouble going to sleep or staying asleep? **GSS19**: [bool,.=0]: Do you regularly have trouble going to sleep or staying asleep?
healthStress	$\surd$				$\surd$	**GSS17**: [scaled]: Thinking about the amount of stress in your life, would you say that most days are: **GSS19**: [scaled]: Thinking about the amount of stress in your life, would you say that most days are:
helpfulNeighbours	$\surd$				$\surd$	**GSS17**: [scaled]: Would you say this neighbourhood is a place where neighbours help each other? **GSS19**: [scaled]: Would you say this neighbourhood is a place where neighbours help each other?
honestNeighbour	$\surd$					**GSS17**: [scaled]: If you lost a wallet or purse that contained two hundred dollars, how likely is it to be returned with the money in it if it was found: .by someone who lives close by? Would it be:
honestStranger	$\surd$					**GSS17**: [scaled]: If you lost a wallet or purse that contained two hundred dollars, how likely is it to be returned with the money in it if it was found: .by a complete stranger?
honesty	$\surd$					**GSS17**: (RAW CODEBOOK INFO MISSING)
knowNeighbours	$\surd$				$\surd$	**GSS17**: [scaled]: Now I would like to ask you a few questions about your more immediate neighbourhood. Would you say that you know: **GSS19**: [scaled]: Now I would like to ask you a few questions about your more immediate neighbourhood. Would you say that you know:
livingWithFriends	$\surd$				$\surd$	**GSS17**: Number of respondent's close friends living in household. **GSS19**: Number of respondent's close friend(s) living in household.
logTenureCity					$\surd$	**GSS19**: Length of time respondent has lived in current city or local community.
logTenureNeighbourhood					$\surd$	**GSS19**: Length of time respondent has lived in current neighbourhood.
mastery	$\surd$					**GSS17**: [scaled]: Mastery scale.
noReligion	$\surd$			$\surd$	$\surd$	**GSS17**: Religion of respondent. In fifteen categories. **EDS**: Derived - Religion - Christian or non-Christian **GSS19**: Religion of respondent. In fifteen categories.
safeAtHome	$\surd$					**GSS17**: [scaled]: When alone in your home in the evening or at night, do you feel:
safeAtNight	$\surd$					**GSS17**: [scaled]: How safe do you feel from crime walking alone in your area after dark? Do you feel:
satisFinances	$\surd$				$\surd$	**GSS17**: [scaled]: Please rate your feelings about certain areas of your life, using a scale of 1 to 10 where 1 means "Very dissatisfied" and 10 means "Very satisfied". What about: .your finances? **GSS19**: [scaled]: Please rate your feelings about them, using a scale of 1 to 10 where 1 means "Very dissatisfied" and 10 means "Very satisfied". What about: your finances?
satisHealth	$\surd$				$\surd$	**GSS17**: [scaled]: Please rate your feelings about certain areas of your life, using a scale of 1 to 10 where 1 means "Very dissatisfied" and 10 means "Very satisfied". What about: .your health? **GSS19**: [scaled]: Please rate your feelings about them, using a scale of 1 to 10 where 1 means "Very dissatisfied" and 10 means "Very satisfied". What about: your health?
satisJob	$\surd$				$\surd$	**GSS17**: [scaled]: Please rate your feelings about certain areas of your life, using a scale of 1 to 10 where 1 means "Very dissatisfied" and 10 means "Very satisfied". What about: .your job or main activity? **GSS19**: [scaled]: Please rate your feelings about them, using a scale of 1 to 10 where 1 means "Very dissatisfied" and 10 means "Very satisfied". What about: your job or main activity?

Table A.7: Summary of survey variable definitions.

Variable	GSS17	ESC1	ESC2	EDS	GSS19	
satisTime	√				√	**GSS17**: [scaled]: Please rate your feelings about certain areas of your life, using a scale of 1 to 10 where 1 means "Very dissatisfied" and 10 means "Very satisfied". What about: .the way you spend your other time? **GSS19**: [scaled]: Please rate your feelings about them, using a scale of 1 to 10 where 1 means "Very dissatisfied" and 10 means "Very satisfied". What about: the way you spend your other time?
trustColleagues	√			√	√	**GSS17**: [scaled]: Using a scale of 1 to 5 where 1 means 'Cannot be trusted at all' and 5 means 'Can be trusted a lot', how much do you trust each of the following groups of people: .people you work with or go to school with? **EDS**: [scaled]: Using a scale of 1 to 5 where 1 means cannot be trusted at all and 5 means can be trusted a lot, how much do you trust each of the following groups of people: People that you work with or go to school with? **GSS19**: [scaled]: Using a scale of 1 to 5 where 1 means 'Cannot be trusted at all' and 5 means 'Can be trusted a lot', how much do you trust each of the following groups of people:... people you work with or go to school with?
trustFamily	√			√	√	**GSS17**: [scaled]: Using a scale of 1 to 5 where 1 means 'Cannot be trusted at all' and 5 means 'Can be trusted a lot', how much do you trust each of the following groups of people: .people in your family? **EDS**: [scaled]: Using a scale of 1 to 5 where 1 means cannot be trusted at all and 5 means can be trusted a lot, how much do you trust each of the following groups of people: People in your family? **GSS19**: [scaled]: Using a scale of 1 to 5 where 1 means 'Cannot be trusted at all' and 5 means 'Can be trusted a lot', how much do you trust each of the following groups of people:... people in your family?
trustNeighbourFraction	√					**GSS17**: [scaled]: Would you say that you trust:
trustStrangers	√				√	**GSS17**: [scaled]: Using a scale of 1 to 5 where 1 means 'Cannot be trusted at all' and 5 means 'Can be trusted a lot', how much do you trust each of the following groups of people: .strangers? **GSS19**: [scaled]: Using a scale of 1 to 5 where 1 means 'Cannot be trusted at all' and 5 means 'Can be trusted a lot', how much do you trust each of the following groups of people:... strangers?
valueSocial	√					**GSS17**: [scaled]: Using a scale of 1 to 5, where 1 is not important at all and 5 is very important, how important is it for you to establish and maintain ties: .with other people?

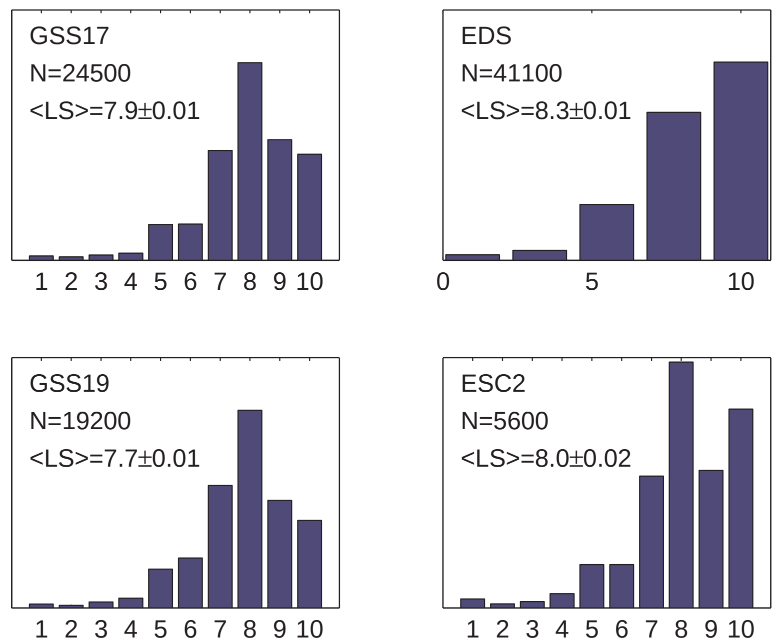


Figure A.1: Histograms of reported *life satisfaction* in several Canada-wide surveys. Estimates and standard errors of the mean are shown, as are sample sizes.

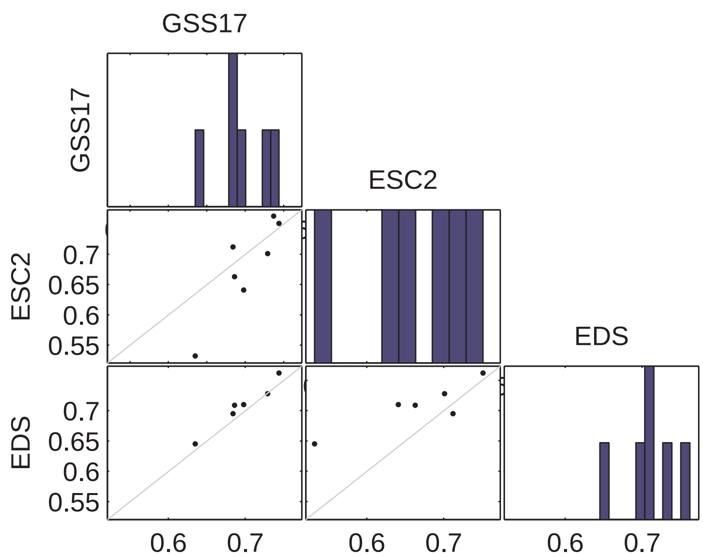


Figure A.2: Correlation between surveys of mean trust in neighbours by province. Trust in neighbours is scaled to lie between 0 and 1. The light gray line represents perfect correspondence. Along the diagonal are shown histograms of the provincial averages.

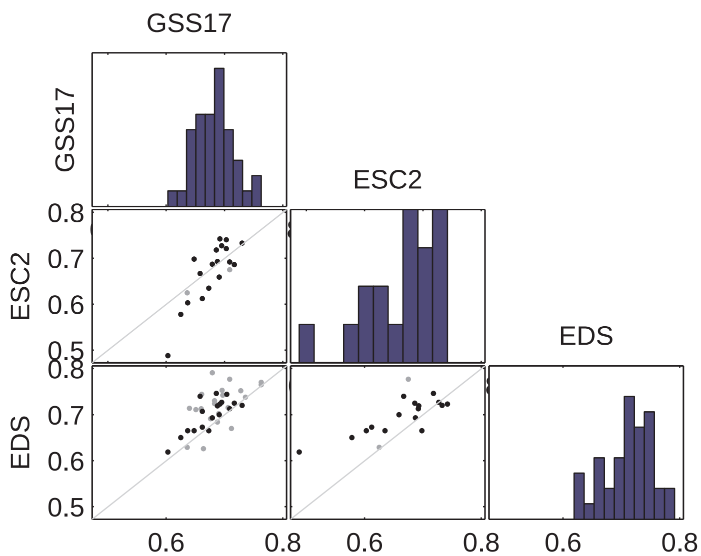


Figure A.3: Correlation between surveys of mean trust in neighbours by CMA. In this and the subsequent several figures, the heavier dots represent regions with higher populations than the lighter dots.

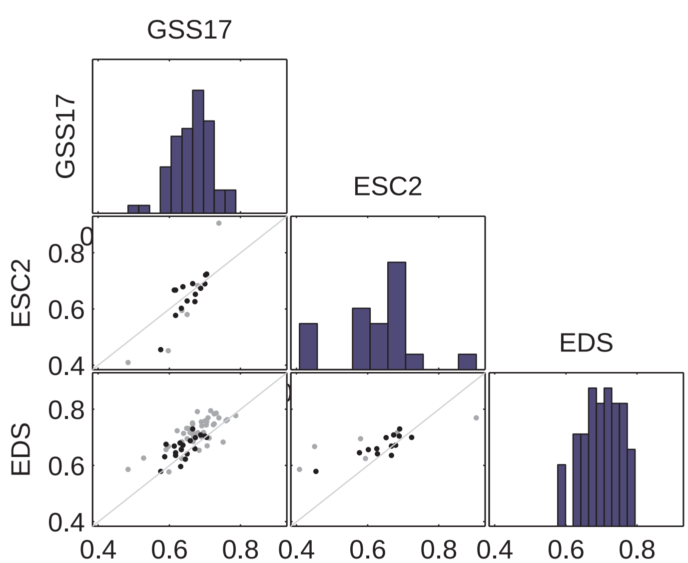


Figure A.4: Correlation between surveys of mean trust in neighbours by CSD.

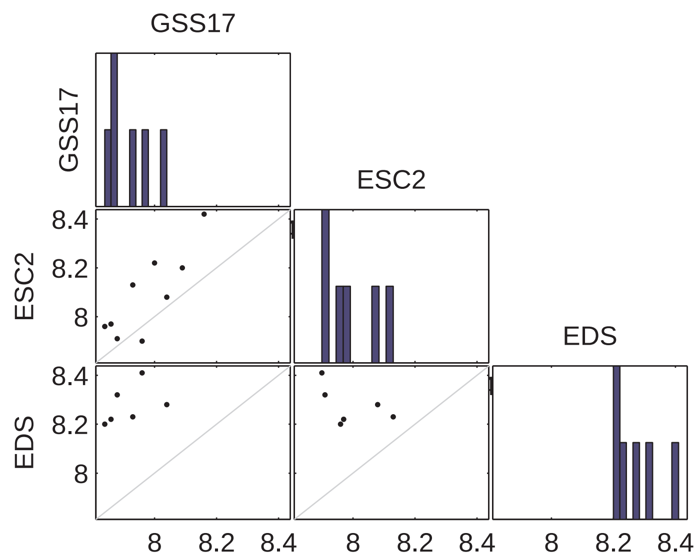


Figure A.5: Comparison of provincial mean life satisfaction from different surveys.

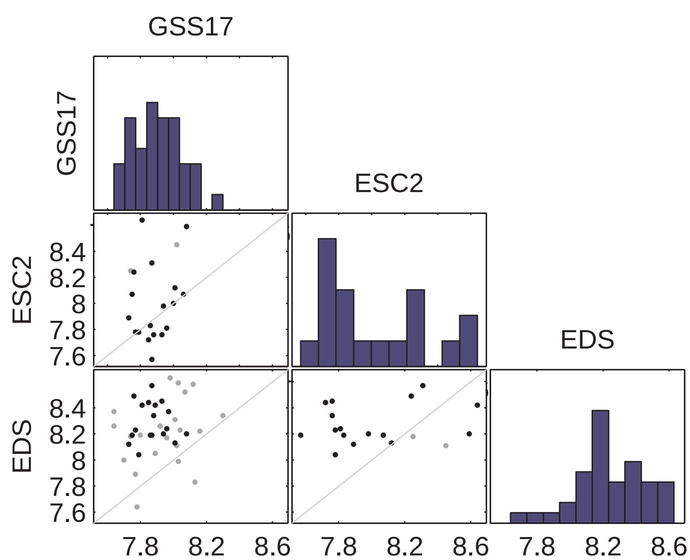


Figure A.6: Comparison of CMA mean life satisfaction from different surveys.

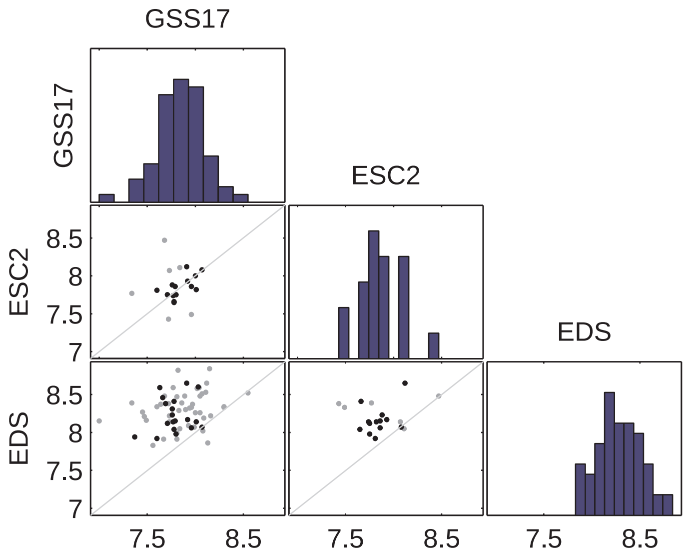


Figure A.7: Comparison of CSD mean life satisfaction from different surveys.

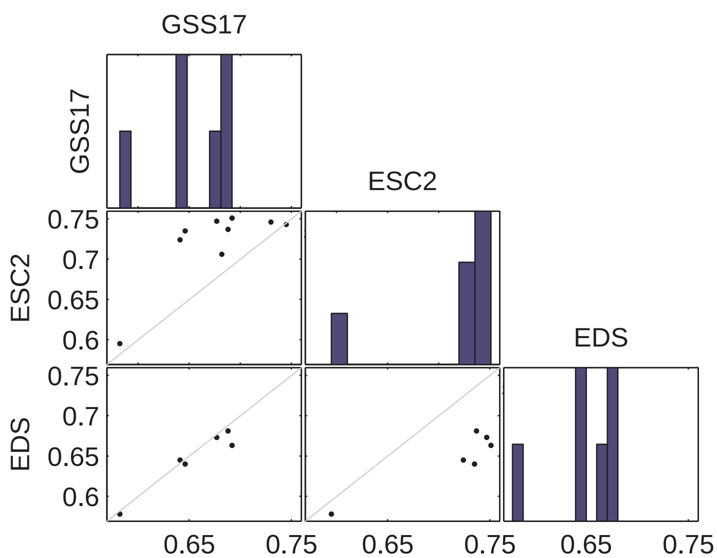


Figure A.8: Comparison of provincial mean *importance of religion* from different surveys.

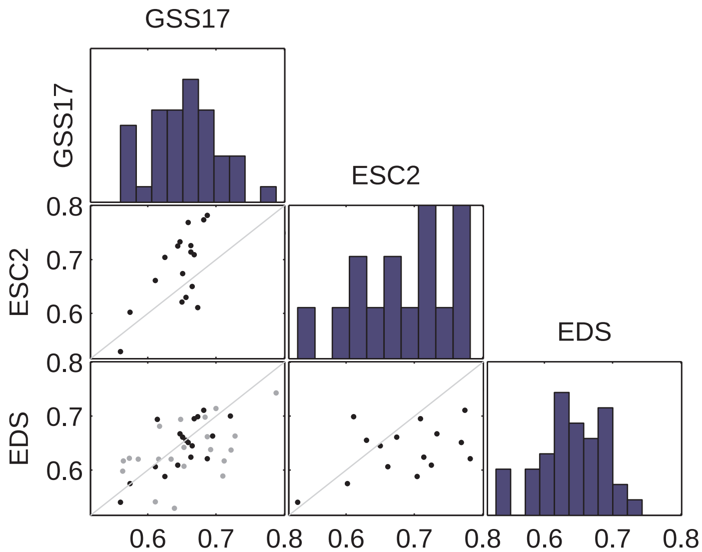


Figure A.9: Comparison of CMA mean *importance of religion* from different surveys.

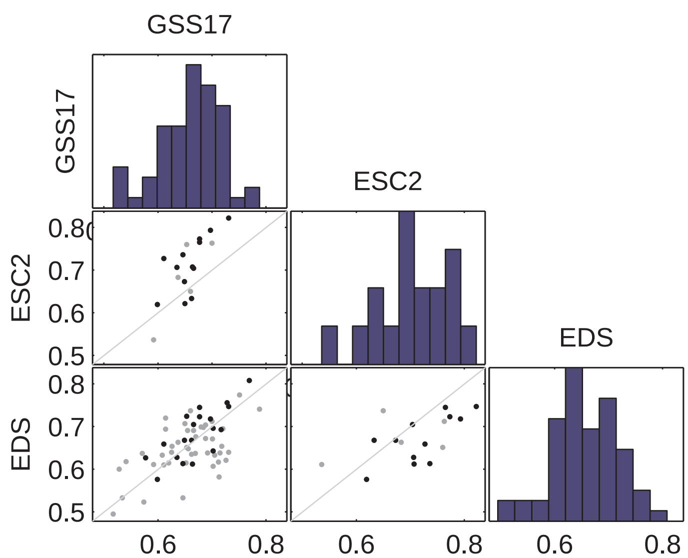


Figure A.10: Comparison of CSD mean *importance of religion* from different surveys.

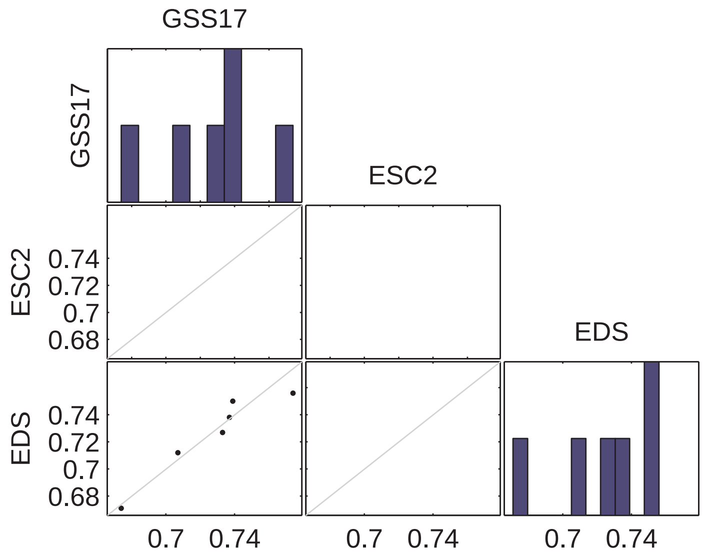


Figure A.11: Comparison of provincial mean trust in colleagues from different surveys.

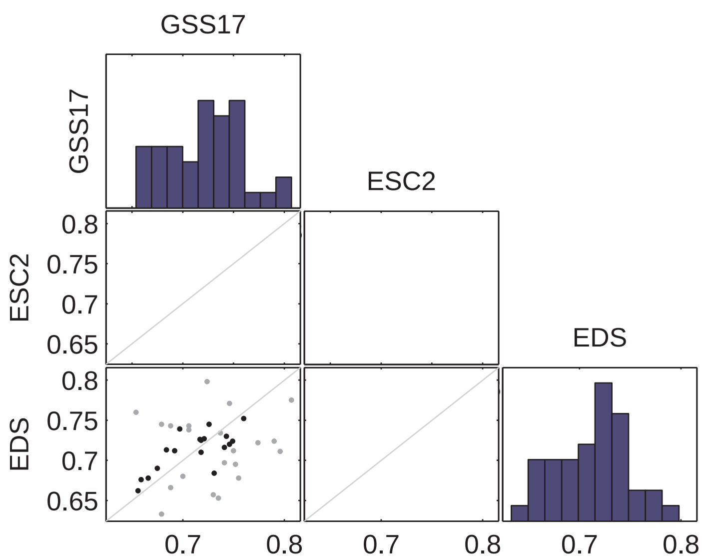


Figure A.12: Comparison of CMA mean trust in colleagues from different surveys.

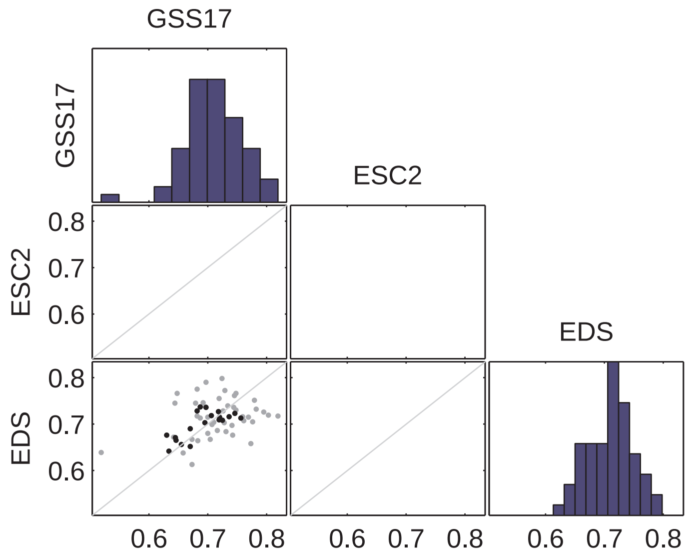


Figure A.13: Comparison of CSD mean trust in colleagues from different surveys.

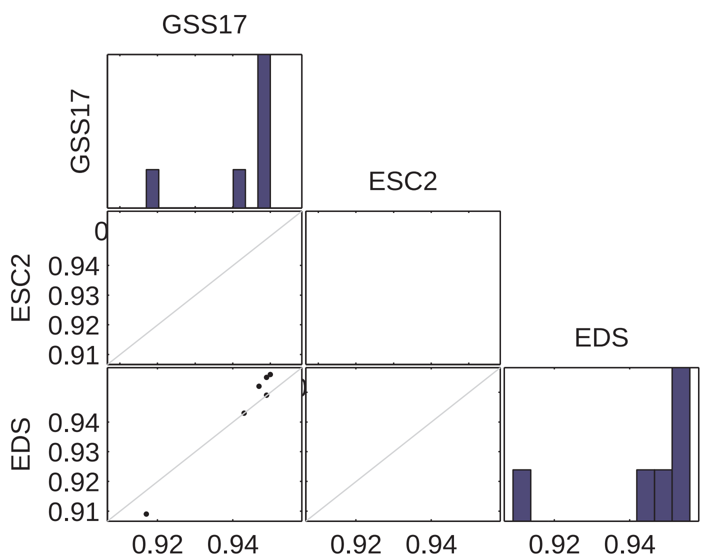


Figure A.14: Comparison of provincial mean trust in family from different surveys.

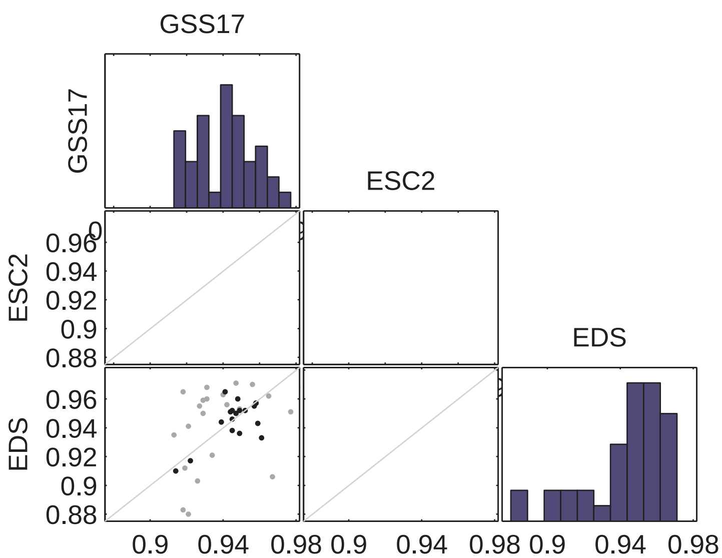


Figure A.15: Comparison of CMA mean trust in family from different surveys.

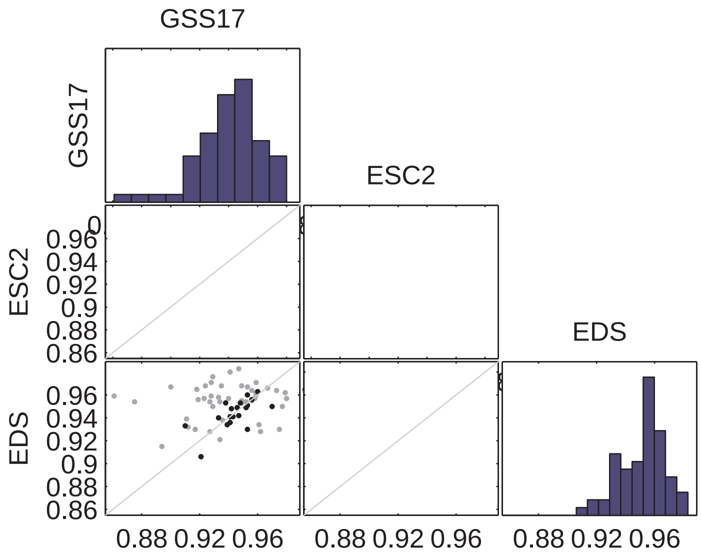


Figure A.16: Comparison of CSD mean trust in family from different surveys.

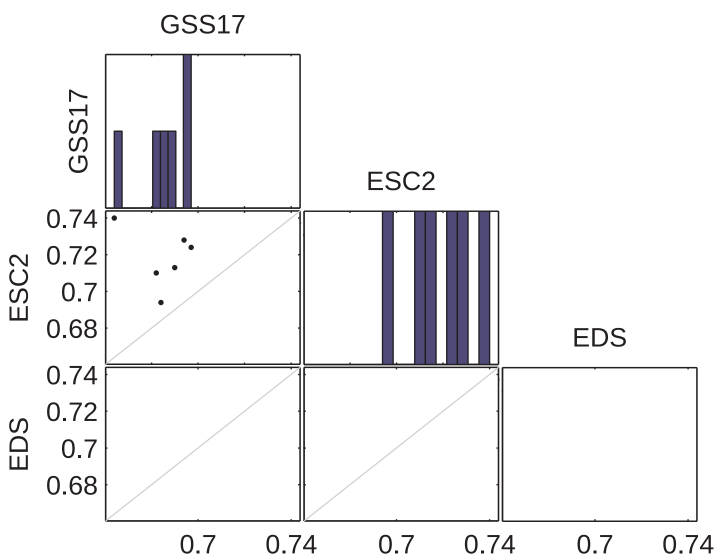


Figure A.17: Comparison of provincial mean subjective health from different surveys.

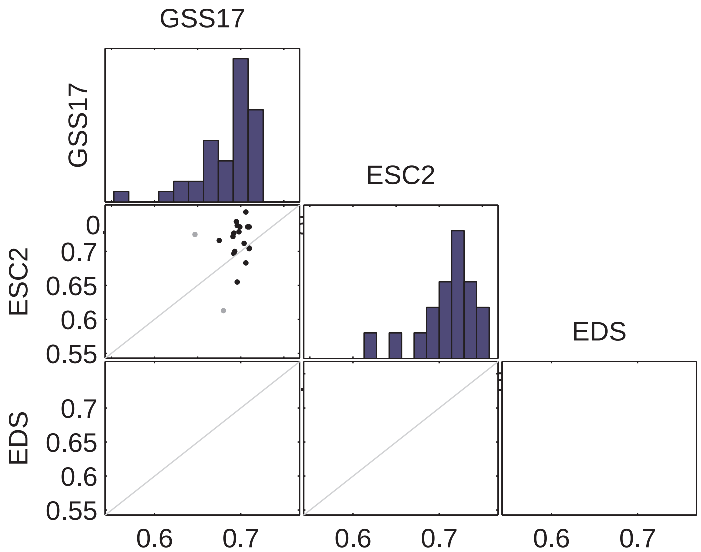


Figure A.18: Comparison of CMA mean subjective health from different surveys.

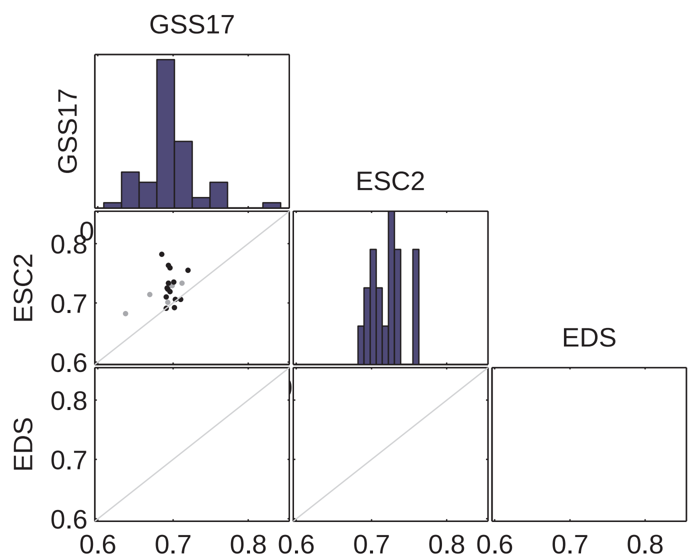


Figure A.19: Comparison of CSD mean subjective health from different surveys.

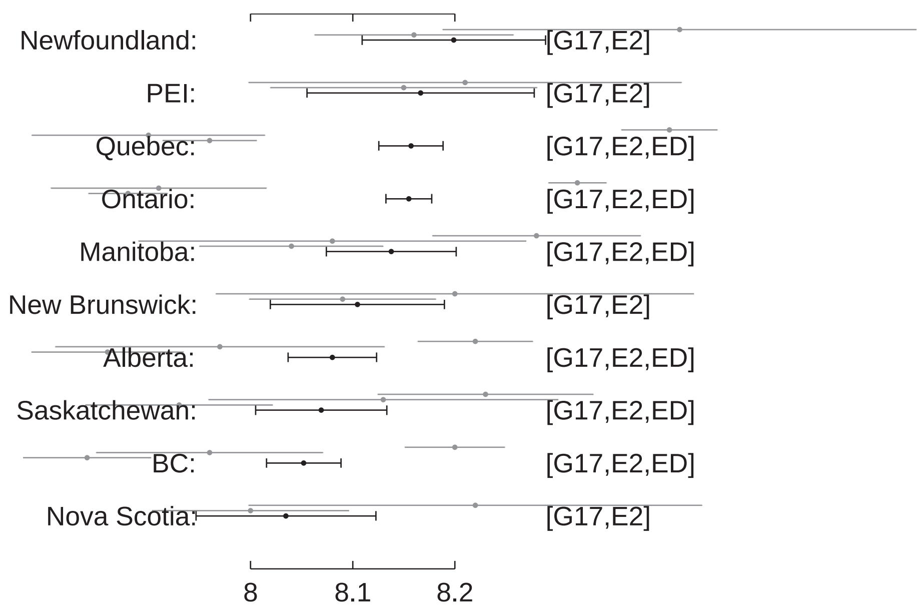


Figure A.20: Life satisfaction means by province. In this and the subsequent several figures, the right column indicates which surveys provide sufficient samples to include in the means. The light error bars show the individual means from each of these surveys, while the darker bars show the appropriately weighted mean using all available surveys.

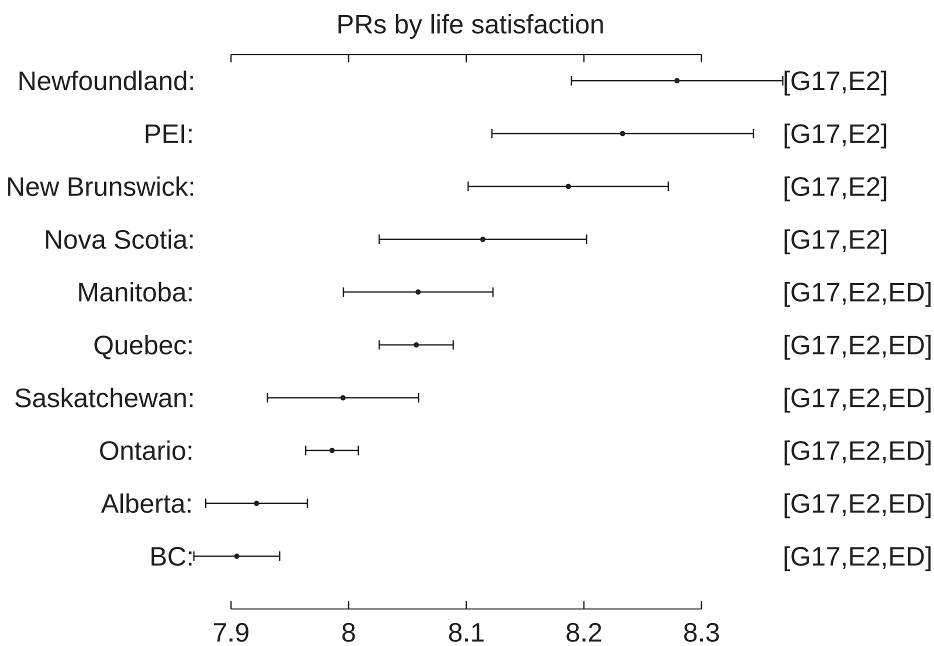


Figure A.21: Life satisfaction means by province, corrected for survey averages.

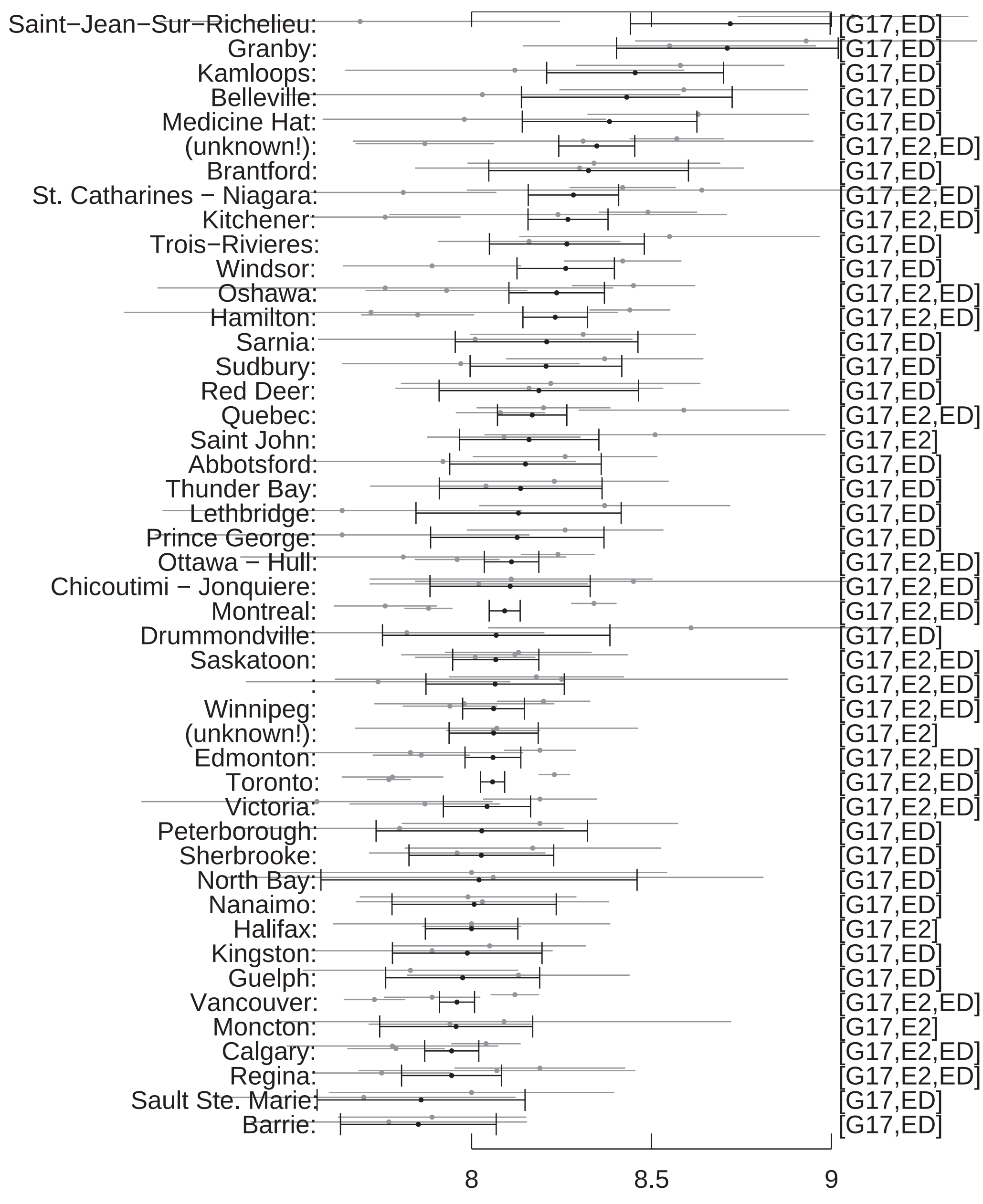


Figure A.22: Life satisfaction means by CMA.

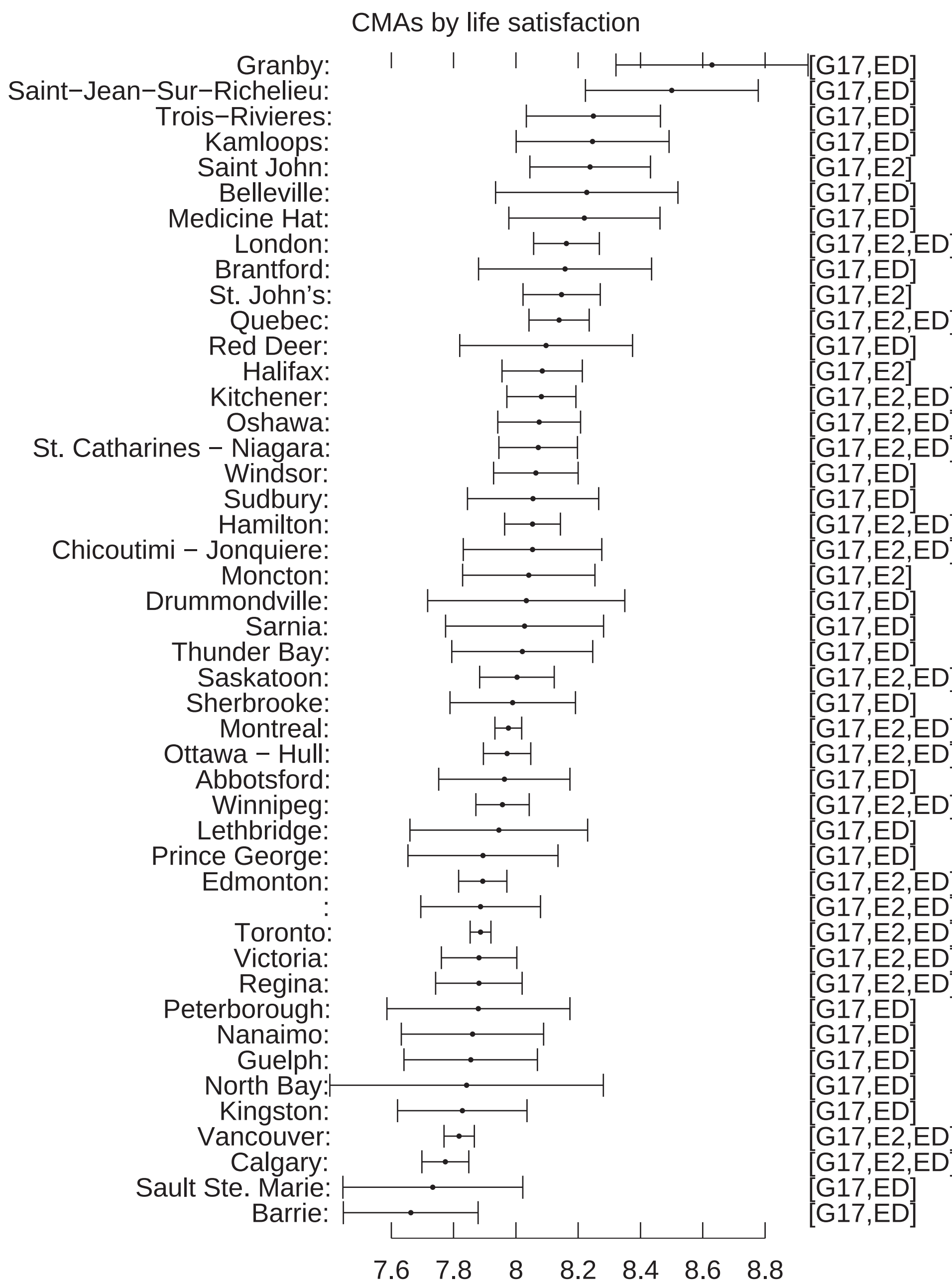


Figure A.23: Life satisfaction means by CMA, corrected for survey averages.

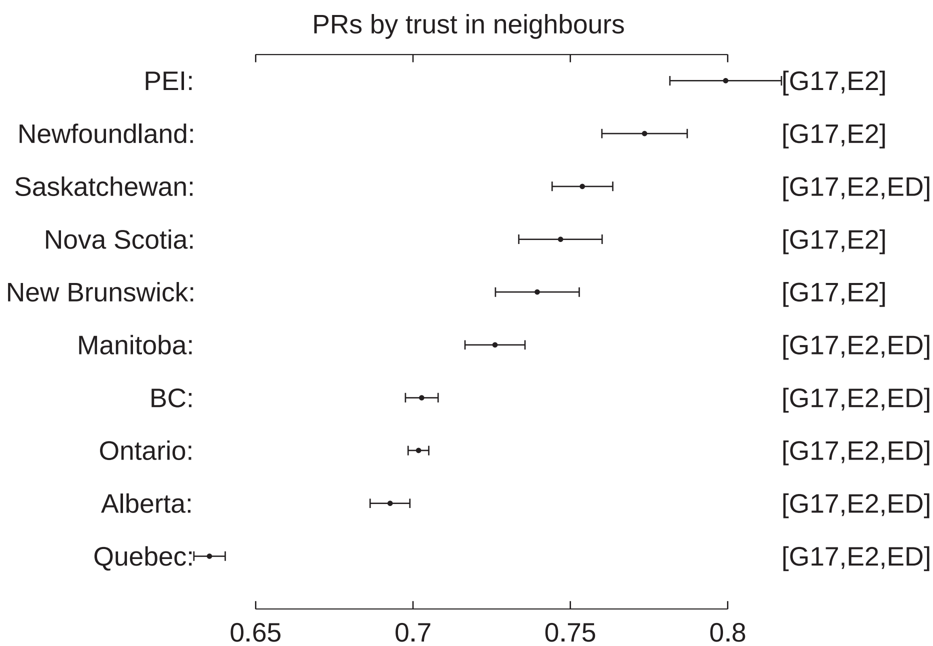


Figure A.24: Trust in neighbours by province.

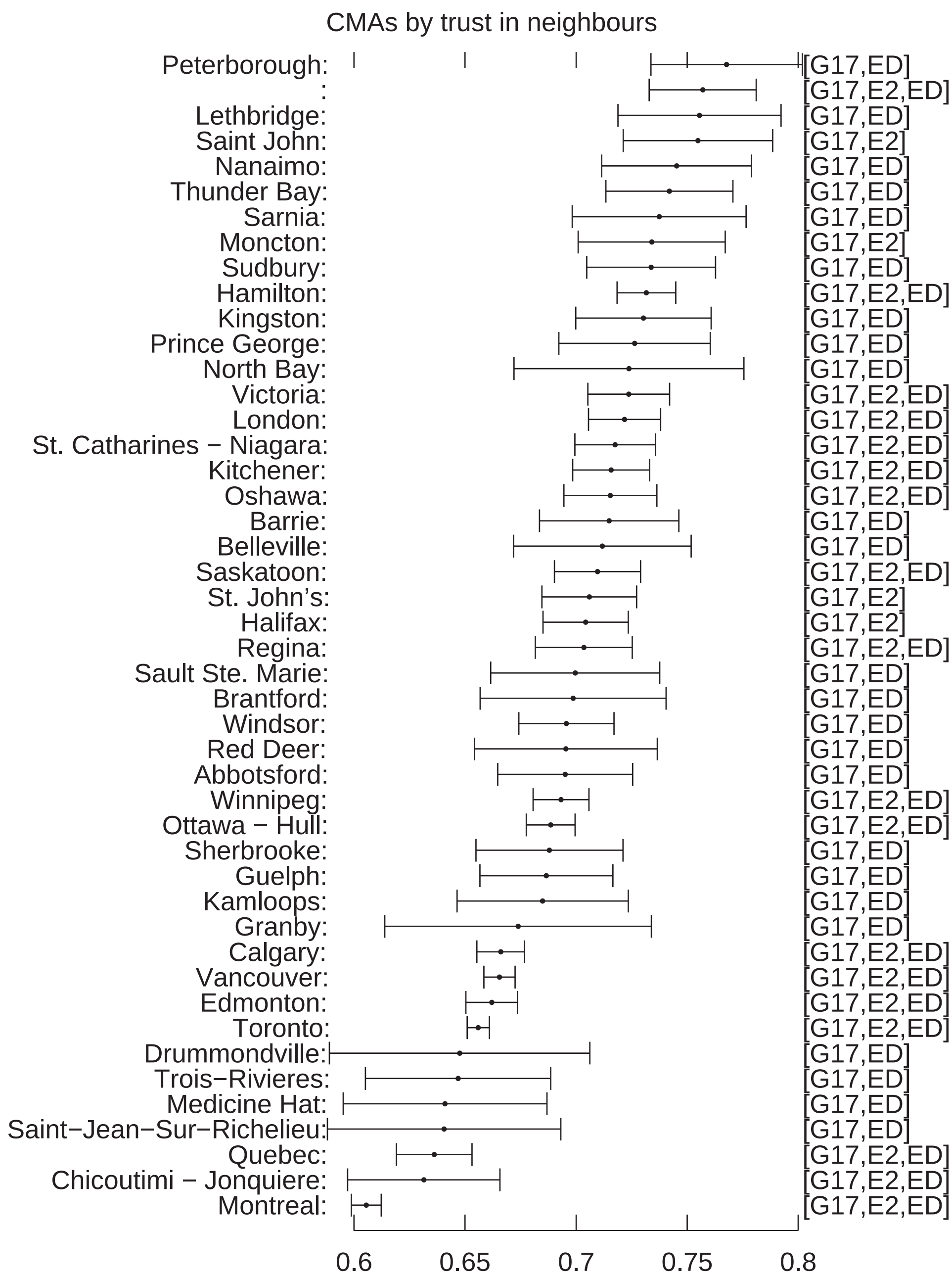


Figure A.25: Trust in neighbours by CMA.

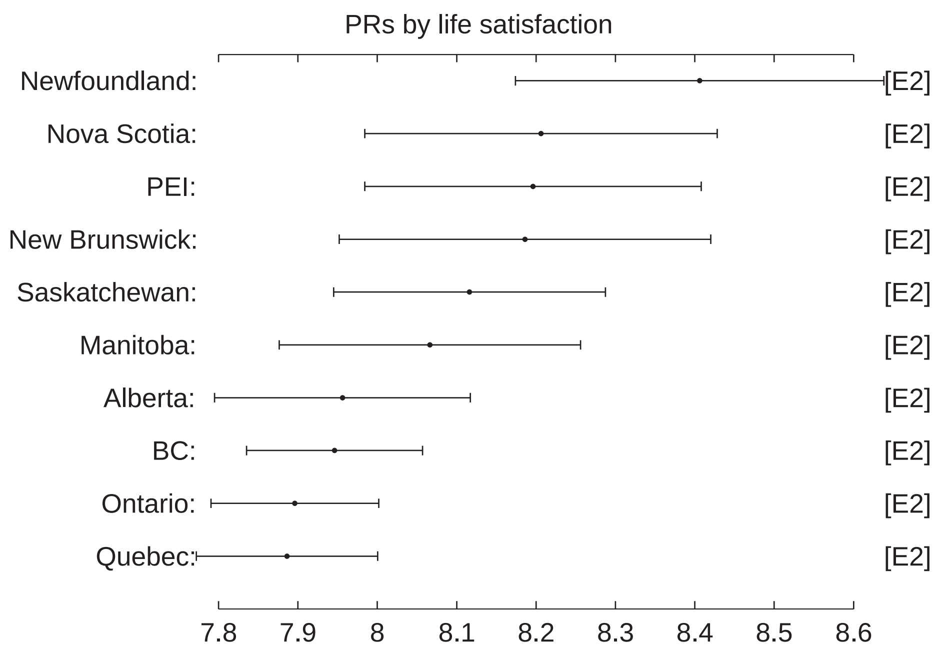


Figure A.26: Life satisfaction from ESC2 by province.

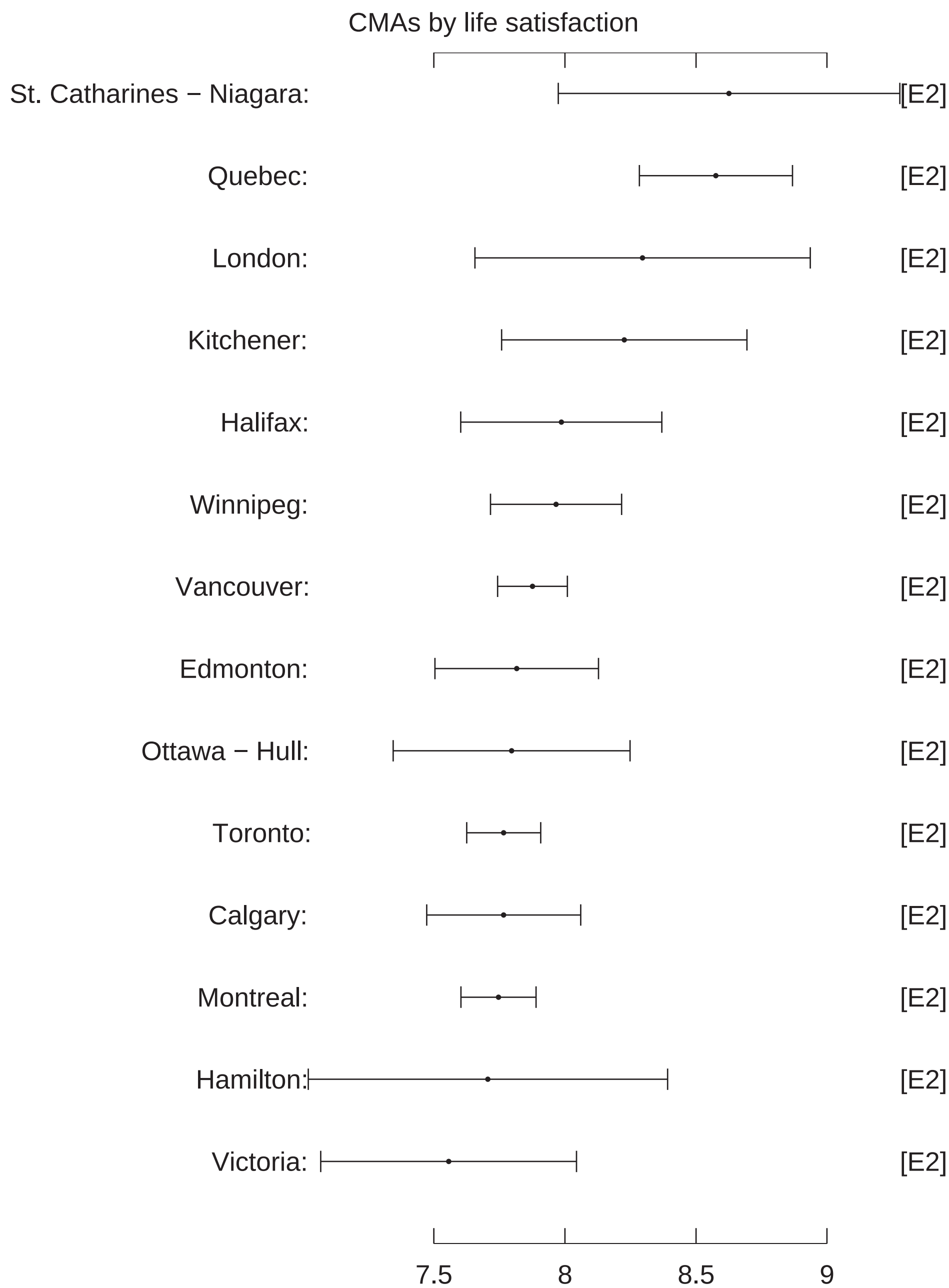


Figure A.27: Life satisfaction from ESC2 by CMA.

Bibliography for Appendix to Chapter 2

Helliwell, J., and H. Huang, How's the Job? Well-Being and Social Capital in the Workplace, *Forthcoming in Industrial and Labor Relations Review*, 2005.

Soroka, S., R. Johnston, and J. F. Helliwell, *Diversity, Social Capital and the Welfare State*, chap. Measuring and Modelling Interpersonal Trust, UBC Press, 2007.

Appendix B

Appendix to Chapter 3

This appendix provides detail and more in-depth discussion of several issues raised in or tangential to the main body of the paper. Section B.1 explains the relationship of this work to the literature on club economies. Section B.2 discusses some very simple models of heterogeneity and Veblen preferences which give intuition for the more general case. Section B.3 discusses the choice of functional forms used in the paper. Section B.5 outlines proofs of propositions stated earlier, and Section B.6 describes how to construct the separating equilibrium used in numerical examples.

B.1 Endogenous reference groups are not club goods

There is a large existing literature on local public goods and agglomeration into clubs. The problem I address here is distinct in a couple of ways. Neighbourhoods are unlike clubs in that their membership does not explicitly choose an entry price, nor do they coordinate (typically through a voting scheme, or just through coordination in the core equilibrium) on the nature of the public good they provide. That is, I assume that neighbourhoods do not set standards for how lawns and gardens must be kept or how ornate new houses must be. Rather, in the models here the homogeneous behaviour within neighbourhoods is a result only of relativities in preferences and possibly of the individual ability to pay in a competitive market for land.

I also ignore for the moment congestion and the endogenous sizing of communities or of lots of land, although these are clearly relevant to spatial development patterns. A relevant observation is that in Canada the lowest density settlements are populated by the richest and poorest. High density areas are populated by more median incomes, presumably in urban high-rises.

Standard efficiency considerations for models of local public goods and "capitalisation" are not appropriate when the public goods are Veblen goods. Existing models tend to focus on tax and price systems which afford efficient allocations within jurisdictions and efficient location choice for individuals. Because I do not assume that differentiation of types occurs on the same geographic scale as tax taking institutions or those offering public services, I just ignore those policy instruments and look for equilibria without them.

Although the models I consider below do not include actions of coordinated neighbourhoods, insights from the work could inform local providers of public goods such as neighbourhood associations, clubs, or local governments. Along with public good problems, these groups face migration and changing distribution of local wealth, pressures on land price, and economic growth, which are the key concepts to follow.

B.2 An introduction to neighbourhood segregation in the presence of Veblen goods

To introduce the ideas to come, consider a straightforward application of heterogeneity to the *Pure Veblen 1* formulation of *Eaton and Eswaran* [2006], as described below.

B.2.1 Exogenous segregation and Veblen consumption

Let there be two types of household, differentiated only by their endowed labour productivities, $w_H > w_L$. Preferences are generated by utility

$$U = F(x) + H(h - \bar{h})$$

where $0 < x < 1$ is chosen leisure and the numeraire good, h, is purchased according to the budget, $h \leq w[1-x]$. This good is a Veblen good in that its benefit is derived only through consumption relative to a reference consumption level, $\bar{h}$. We may imagine that h measures a form of intrinsically useless conspicuous consumption such as living in a grandiose house.[1] In this model economy and several of those to follow, building such houses is the only industry. The function $H(\cdot)$ then represents the status value of living with consumption level h amongst neighbours with average consumption level $\bar{h}$.

One can begin by considering the welfare implications of inequality. Is society better off with distinct types, w_H and w_L, segregated or integrated? If neighbourhoods characterised by their averge consumption $\bar{h}$ can be completely separated into homogeneous groups, then the utilities of the two types will be

$$\begin{aligned} U_L^s &= F(x_L^s) + H(0) \\ U_H^s &= F(x_H^s) + H(0) \end{aligned}$$

whereas in a homogeneously mixed community, outcomes are, for the case of equal populations in the two types,

[1]Such "pure Veblen" goods represent the case when any intrinsic value to increased consumption of the good suffers strongly from diminishing returns. Similar outcomes may be seen in the more general case when both absolute and relative benefits accrue from consumption. When others' consumption is roughly on par with one's own, the relative effects, or Veblen terms, remain while the absolute benefits saturate [*Eaton and Eswaran*, 2006]. In the current context, all houses are large enough to satisfy needs and provide most benefits that dwellings can provide their owners directly.

$$
\begin{aligned}
U_L^m = & \quad F(x_L^m) + H(-\Delta) & < U_L^s \\
U_H^m = & \quad F(x_H^m) + H(\Delta) & > U_H^s
\end{aligned}
$$

where $\Delta \equiv \frac{1}{2}h_H^m - \frac{1}{2}h_L^m$. The given inequalities follow if $H(\cdot)$ is strictly concave. They indicate that the high types are better off in the integrated community while the low types are worse off. Furthermore, concavity of both $H(\cdot)$ and $F(\cdot)$ implies that $x_H^m > x_H^s$ and $x_L^m < x_L^s$. That is, the high productivity individuals will work less in the integrated community than in the segregated one, and conversely for the less productive type. In addition, $H(\Delta) + H(-\Delta) < 2H(0)$; that is, the summed benefits derived from housing alone are higher in the segregated case. However, without specifying functional forms, no definitive statement can be made about the relative efficiencies of the two cases on the basis of summed utilities which include both leisure and housing.[2]

Similarly, the efficiency implications of growth are ambiguous. In the segregated case, growth in either w_H or w_L beyond some minimum level is unequivocally bad for welfare, as in *Eaton and Eswaran* [2006]. On the other hand and in contrast to their homogeneous case, the implications of growth for the mixed community is indeterminate without more assumptions.

Several of these ambiguities recur below when segregation is an endogenous outcome. Rather than pursue welfare analysis for specific functional forms at this stage, I consider next the implications of disaggregated choice in this simple economy in order to motivate necessary subsequent extensions to the model.

B.2.2 Endogenous segregation without neighbourhood benefits

To continue the introduction of heterogeneity to the competitive Veblen economy described in *Eaton and Eswaran* [2006], I now incorporate disaggregated decision making in the two-neighbourhood world of Section B.2.1. Consider the case where multiple neighbourhood locations exist and each household chooses where to live as well as how much to spend on its own consumption of housing. Assume that households are aware of the possibility of relocating, even though their consumption comparisons extend only to their own, chosen neighbours. This represents an endogenous choice of reference group for the Veblen good.

For any form of utility $U = F\left(x, v\left(h, \bar{h}\right)\right)$ with $\partial U / \partial \bar{h} < 0$, however, such neighbourhood differentiation is not possible. The uniform undesirability of high-consumption neighbours means that all decision makers will prefer the lowest available average neighbourhood con-

[2] Using the concept of "transferable utility" to justify the common practice of adding utilities and ordering outcomes on the basis of social (i.e., aggregate) welfare may seem a dubious method when utility has its normal modern interpretation as an abstract determinant of decision making. The widespread availability now of measured subjective well-being, however, gives the current exercise the slightly more empirical interpretation of comparing regional average life satisfaction levels under the two scenarios.

sumption $\bar{h}$. That is, all types want to live with the poorest neighbours. Even if there is a free market for land in different neighbourhoods, the most able to buy are the most wealthy and therefore those with the least desirable externality to their neighbours. As a result, no differentiated neighbourhoods may exist in equilibrium.

B.2.3 Neighbourhood benefits

A segregated equilibrium where migration is a choice can thus only exist if there is coordination of some kind between members of a neighbourhood[3] or if there is another, countervailing externality acting in addition to the negative consumption externality considered above.

That is, in order to explain why wealthy types and poor types might each prefer to live amongst their own, there must be a neighbourhood benefit simultaneously with the local consumption comparison. There are several obvious reasons for higher productivity neighbourhoods to be desirable:

1. Productivity w could be an exogenous feature of location rather than of the individual; in cross-sectional or short-term studies this amounts to historical factors which determine opportunity or availability of resources.

2. Neighbourhoods could be characterised by the intrinsic quality of their residents in a way which confers either pecuniary benefits to neighbours (through higher income opportunities due to networking or signaling) or social (non-wage) benefits through higher quality social interactions or more efficient child-rearing and home production.

3. Conspicuous consumption by neighbours (for instance and in particular, fancy houses) could determine a common status value enjoyed by all residents of a neighbourhood. In this case, the consumption comparison group is other neighbourhoods in the region.

In Section 3.2, I begin by considering benefits of type 2 and later focus on a self-consistent economy incorporating benefits of type 3. Either of these positive spillovers within a neighbourhood may exist simultaneously with the negative spillovers due to local, neighbour-to-neighbour consumption externalities.

One can imagine a model economy in which decision makers weigh these two effects against each other in order to choose a place to live. A household's choice of neighbourhood will be optimal if the benefit from that neighbourhood less the consumption externality suffered from living there is better than any other available option. This is the decision problem if entry into each neighbourhood is free to anyone who chooses it. I will show below that this scenario does not always lead to an economy with more than one neighbourhood.

[3]For instance, to assess taxes or membership fees. I do not consider this possibility, which as mentioned above is well covered by the literature on "club economies".

Alternatively, entry into a neighbourhood might have a further cost due to the price of land. Whatever are the reasons behind the benefit from living in a particular location, that benefit may be captured in land prices that arise endogenously in the economy; this is known as *capitalisation* in the literature on club economies.[4] When no differentiated neighbourhoods are possible with free land, there may still be separated equilibria when a land market exists. I address both cases of priced and unpriced land in the discussion that follows.

Because the form of preferences under discussion, which incorporates relativities, is unusual in economics, it will prove useful to explore some qualitative features using simple functional forms which, based on experience with non-Veblen utility functions, one might expect to be tractable. It turns out that a prominent feature of even simple forms of preferences involving endogenous reference groups is that the utility function is not globally concave. As a result, in many cases a desirable equilibrium does not exist or exists only for special sets of parameters.

The next section introduces some functional forms used in subsequent analysis, before building simple equilibria in which endogenous choice of reference groups leads to differentiated neighbourhoods.

B.3 Functional forms for Veblen preferences

Economists tend to use a narrow class of functional forms in parameterising utility. These functions are selected for their convenient macroeconomic properties and they typically have a domain restricted to positive values, which are appropriate for the study of preferences over absolute consumption levels. When describing preferences over relative consumption levels, these may be insufficient and new classes of functions which tend to be unfamiliar to consumption theory may be useful.

Clark and Oswald [1998] point out that utility which is concave in relative consumption leads to emulation, while comparison-convex utility leads to deviant behaviour. *Tversky and Kahneman* [1991]'s empirical findings on loss aversion might be rationale for expressing comparison utilities using a form of sigmoid curve, easily expressed using a hypertrigonometric form. The bounded extremes and slopes of such a function are also conducive to efficient numerical simulation. However, not being concave, sigmoid utility makes marginal analysis difficult.

Eaton and Eswaran [2006] consider two classes of comparison-concave utility. These are general concave increasing functions of either a difference or a ratio of own and average consumption levels. In the present work, I employ explicit forms for each of these two classes.

$$H(h,\bar{h}) = \Lambda \log\left(1 + h/\bar{h}\right)$$

and

[4] See *Scotchmer* [2002] for a review.

$$H(h,\bar{h}) = -\Lambda \exp\left(-\lambda\left[h-\bar{h}\right]\right)$$

Both forms are increasing, comparison-concave, and continuous for any nonnegative h and positive $\bar{h}$. Both are relatively simple and likely to have analytic tractability.

B.4 Nonexistence of separating equilibrium for discrete types model

In Section B.4.1, I introduce a slightly simplified utility form in which *benefits* from neighbours' consumption enters directly into the utility function, without comparison to other neighbourhoods. Section B.4.2 analyses one functional form in this class of utility functions, showing that there can be no equilibrium in which types separate. Sections B.4.3-B.4.5 consider several variations on the story which play a role in the treatment of a continuum of agent types in Section 3.3.

B.4.1 Direct neighbourhood benefits

This section formalises the endogenous reference group choice problem outlined above, in which households derive benefit from their relative consumption of housing but their absolute consumption of neighbourhood quality. Preferences are defined over leisure $x \geq 0$, the conspicuous extravagance $h \geq 0$ of one's house, and the average value $\bar{h}$ of houses in one's choice of a neighbourhood.

Utility is, as before, additively separable into a leisure term $F(\cdot)$, a pure Veblen term $H(\cdot)$ comparing own consumption with that of one's chosen peers, and a further absolute benefit $N(\cdot)$ derived from the consumption level of one's chosen peers:

$$U = F(x) + H(h,\bar{h}) + N(\bar{h}) \tag{B.1}$$

The benefits represented by $N(\cdot)$ are derived through one of the first two channels described in section Section B.2.3.

In maximising this utility function, an agent of type w is constrained by the budget

$$w[1-x] \geq h$$

Here $F(\cdot)$, $H(\cdot)$, and $N(\cdot)$ each obey standard convenient assumptions made concrete below. Given the optimality condition

$$x = 1 - h/w \tag{B.2}$$

the household's decision problem may be reduced to a nested choice of an optimal housing purchase $h^{\star}(\bar{h})$ for each possible neighbourhood $\bar{h}$, followed by a choice of optimal neighbourhood

$\bar{h}^{\star}$. Holding $\bar{h}$ fixed, $U(h)$ is concave and its global optimum is consistent with the first order condition

$$F'(1-\frac{h}{w}) = wH_h(h,\bar{h}) \quad \text{or} \quad h=0 \tag{B.3}$$

The indirect utility $U(w,\bar{h})$ is then derived by substituting into the utility function (B.1) the housing choice $h^{\star}\left(\bar{h}\right)$ which would be selected in a given neighbourhood with average consumption $\bar{h}$:

$$U(w,\bar{h}) = U\left(w, h^{\star}\left(\bar{h}\right), \bar{h}\right) \tag{B.4}$$

If there is a discrete set of available neighbourhoods, each household must choose the one offering the highest utility in (B.4). However, in order to gain insight into the discrete choice optima, consider the case (treated further in Section 3.3) in which a continuum of neighbourhoods is available. Then (B.4) presents a continuous choice maximisation problem with no constraints on $\bar{h}$. Notice, however, that there is no guarantee that this optimisation over $\bar{h}$ is also characterised by a concave objective function. The slope $dU(w,\bar{h})/d\bar{h}$ may have a nonmonotonic dependence on $\bar{h}$, meaning that the global optimum may be difficult to find analytically. Moreover, a global maximum does not necessarily even exist, since $U(\cdot)$ may be unbounded even subject to the budget constraint equation (B.2).

One may understand this by noting that in the scenario described above there is no direct cost to choosing one neighbourhood over another. Without a price for entry to a neighbourhood, for instance in the form of a market for land that is independent from the cost of constructing a house, it is possible for the benefit from having wealthier neighbours to outweigh the penalty from having a relatively less desirable house compared with the one next door.

With this caveat about existence in mind, I now define an equilibrium of interest in which endogenously chosen reference groups are consistent with households being sorted by type. To be more precise, consider a world with, as before, two types of household differentiated only by their endowed labour productivities, $w_H > w_L$, and two neighbourhoods into which individuals may move and build a house.

Definition Then a *discrete separating Nash equilibrium* is a set of allocations $\{h_L \equiv h(w_L)$, $\bar{h}_L \equiv \bar{h}(w_L)$, $h_H \equiv h(w_H)$, $\bar{h}_H \equiv \bar{h}(w_H)\}$ satisfying the necessary optimality conditions for each type w

$$\begin{aligned} \bar{h}(w) &= \bar{h}^{\star}(w) \\ h(w) &= h^{\star}(w, \bar{h}^{\star}(w)) \end{aligned}$$

and the consistency condition

$$\bar{h}(w) = h(w) \tag{B.5}$$

This last condition states that a neighbourhood's average consumption level $\bar{h}$ is equal to the consumption choice h of its residents.

It turns out that this equilibrium, in which types sort themselves into distinct neighbourhoods, is not possible for some preferences such as the one described next.

B.4.2 "Log-log-log" preferences with two types

Consider the following particular case of utility given in equation (B.1):

$$\begin{aligned} F(x) &= \Phi \log(x) \\ H(h,\bar{h}) &= \Lambda \log\left(1 + \frac{h}{\bar{h}}\right) \\ N(\bar{h}) &= N \log\left(\bar{h}\right) \end{aligned}$$

That is, let

$$U(x,h,\bar{h}) = \Phi \log(x) + \Lambda \log\left(1 + \frac{h}{\bar{h}}\right) + N \log\left(\bar{h}\right) \tag{B.6}$$

The optimal choice of housing within a given neighbourhood takes the simple form

$$h^{\star}(w,\bar{h}) = \max\left\{0, \frac{\Lambda w - \Phi \bar{h}}{\Phi + \Lambda}\right\} \tag{B.7}$$

with the corresponding leisure choices[5] given by equation (B.2):

$$x^{\star}(w,\bar{h}) = \Phi \frac{1 + \bar{h}/w}{\Phi + \Lambda} \tag{B.8}$$

Equation (B.7) states that households will choose to consume less (and enjoy more leisure) when their neighbours consume more. This substitution effect between neighbours' consumption and own consumption is a counterintuitive effect for a Veblen good.

Substituting these values into equation (B.6) generates the indirect utility $U(w,\bar{h})$ for each household type w. For interior solutions,

[5]Note that despite the superficial appearance of (B.6), the preferences do not conform to a Cobb-Douglas type, and the optimal allocation to leisure is not independent of others' allocations.

$$
\begin{aligned}
U(w,\bar{h}) &= \log\left(\left[\frac{\Phi}{\Phi+\Lambda}\right]^{\Phi}\left[\frac{\Lambda}{\Phi+\Lambda}\right]^{\Lambda}\right) \\
&\quad +\Phi\log\left(1+\frac{\bar{h}}{w}\right)+\Lambda\log\left(1+\frac{w}{\bar{h}}\right)+N\log\left(\bar{h}\right)
\end{aligned}
\tag{B.9}
$$

The household's problem involves finding the best choice amongst two alternative neighbourhoods $\bar{h}$ available in equilibrium. This goal, or finding a global optimum value $\bar{h}^{\star}(w)$ for this continuous equation, are both nontrivial tasks because $U(w,\bar{h})$ is not concave. Moreover, I next show that a separating equilibrium cannot exist.

Proposition B.4.1. *When group entry (land) is costless and preferences are given by equation* (B.6)*, there is no discrete separating group Nash equilibrium with two types.*

A proof is given on page 164 in an Appendix. The only endogenous choice equilibrium is an unsorted one in which all households end up pooling in the same reference group, characterised by the average value of housing consumption. For instance, if productivities are high enough to avoid corner choices, there is a pooling equilibrium where $\bar{h}$ is given by equation (B.15) with w replaced by its population average.

In order to understand this result more intuitively, it is useful to consider the continuous properties of equation (B.9) in further detail. A significant feature of the indirect utility $U(w,\bar{h})$ is that it is in general neither monotonic nor concave in the choice of neighbourhood $\bar{h}$. The marginal utility of a shift in neighbourhood consumption is derived from equation (B.9):

$$
\frac{dU}{d\bar{h}} = \frac{1}{w+\bar{h}}\left[\Phi+N+[N-\Lambda]\,\frac{w}{\bar{h}}\right] \tag{B.10}
$$

When neighbourhood benefits are valued highly enough in comparison with local relative consumption, $N > \Lambda$ and $U(w,\bar{h})$ is strictly increasing in $\bar{h}$. In that case, the lower type will always prefer to move up to the higher type's neighbourhood when the two are separated.

If instead $\Lambda > N$, utility is initially decreasing but eventually increasing with $\bar{h}$. According to equation (B.10), utility is in this case unbounded as $\bar{h} \to 0$ and as $\bar{h} \to \infty$ and has a minimum value U_{min} at

$$
\bar{h}_{\min U} = \frac{\Lambda-N}{\Phi+N}w \tag{B.11}
$$

Because both $\bar{h}_{\text{eq}}$ and $\bar{h}_{\min U}$ scale directly with w, households occupying their separating equilibrium neighbourhoods will always either both prefer any higher neighbourhood to their own or both prefer any lower neighbourhood to their own. Thus it is impossible for both types to fulfill the equilibrium requirements.

Figure B.1 on page 161 shows the possible cases for preferences conforming to equation (B.6). The left panels show the dependence of the indirect utility on the neighbourhood location $\bar{h}$ for the high type (red) and low type (green). The dependence is characterised by a minimum value which is proportional to the endowments w, in accordance with equation (B.11). Marked on each plot as h_L and h_H are the values $\bar{h}_{\text{eq}}$ for which a type's housing choice is consistent with that of its neighbours, *i.e.,* where $h^\star(\bar{h}) = \bar{h}$. The right panels show indifference contours for $U(h,\bar{h})$ for two values of w. The dashed lines indicate the optimum housing choice $h^\star(\bar{h})$ within each neighbourhood $\bar{h}$. The dotted line is the solution to $h = \bar{h}$, the blue squares show the values of $\bar{h}_{\text{eq}}$, and the red and green squares show each type's optimal choice of h in the *other* type's neighbourhood.

In (a), the left panel shows that $\bar{h}_{\text{eq}}$ is to the left of $\bar{h}_{\min U}$ for both types. Hence both types prefer to move to a less affluent neighbourhood and, in accordance with equation (B.7) and equation (B.8), to build a slightly smaller house and to consume less leisure. Because such a move is available to the high type, the $\bar{h}_{\text{eq}}$ values do not constitute an equilibrium. The right hand panel shows that the optimal housing consumption $h^\star(\bar{h})$ passes near a saddle point in the utility function $U(h,\bar{h})$.

Figure B.1(b) shows the opposite case, when $\bar{h}_{\text{eq}}$ is greater than $\bar{h}_{\min U}$ and thus the low-type household prefers to move locations. Panels (c) are the same as (b) with the values of Λ and N reversed such that $N > \Lambda$. In this case, $U(w,\bar{h})$ is increasing in $\bar{h}$ and a move to a higher expenditure neighbourhood is always beneficial.

B.4.3 Mixed strategies

For simplicity, equation (B.5) describes a pure strategy equilibrium. A less restrictive definition of equilibrium in which mixed strategies are allowed would require only that for each neighbourhood j,

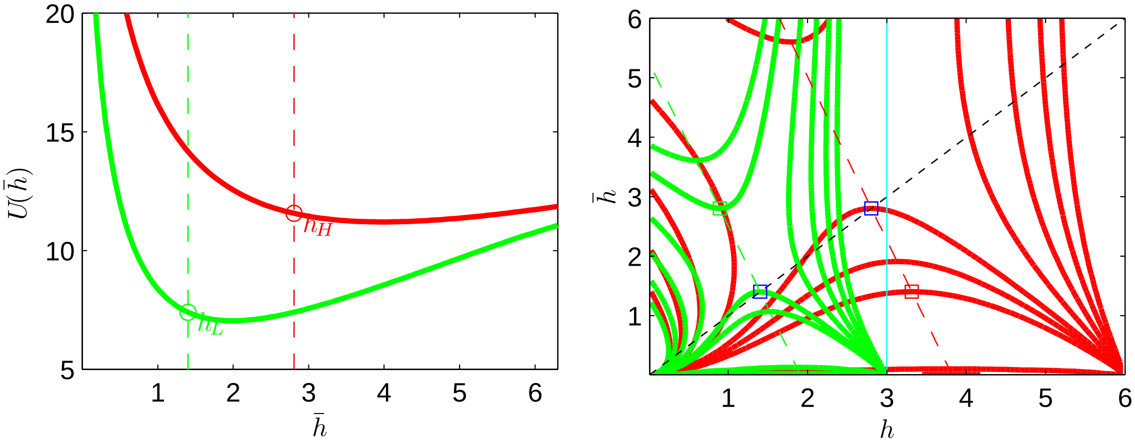
$$\bar{h}_n = \langle h \rangle_{\text{residents}(n)} \tag{B.12}$$

where the average $\langle \cdot \rangle$ is taken over all residents in the neighbourhood. This weaker condition will still not admit any separating outcome in which different types tend to live in different neighbourhoods. This is because for either type to be indifferent between two neighbourhoods, the neighbourhoods must have identical $\bar{h}$ and hence identical mixtures of the two types of household, resulting in a pooling equilibrium.

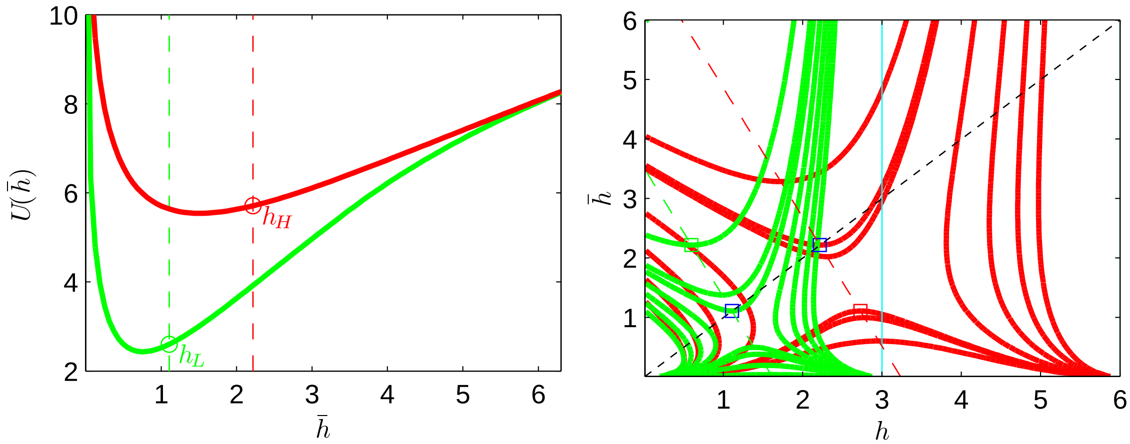
B.4.4 Neighbourhood benefits compared with other neighbourhoods

In equation (B.6), the functional form of $N(\cdot)$ provides unbounded benefits from consumption of the public good $\bar{h}$ while $H(\cdot)$ represents a bounded cost of Veblen comparison as $\bar{h}$ becomes large. As a result, households will always prefer moving to a sufficiently high-consumption neighbourhood rather than remain in their own.

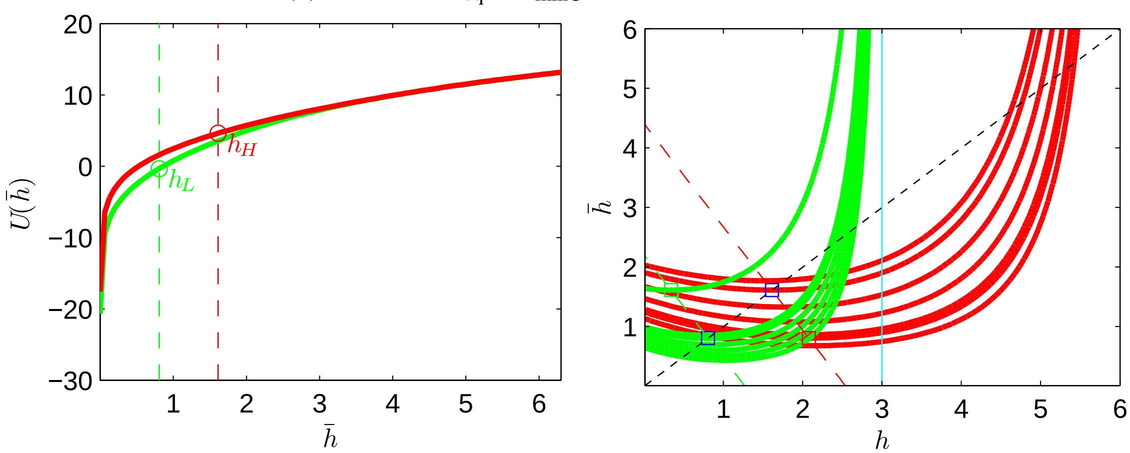
An alternate specification of preferences pertains to neighbourhood status benefits of type 3 on page 154 and is also more consistent with the empirical results outlined in Section 3.1.



(a) Case with $\bar{h}_{\mathrm{eq}} < \bar{h}_{\mathrm{minU}}$: $\Phi = 9, \Lambda = 16$, $N = 6$



(b) Case with $\bar{h}_{\mathrm{eq}} > \bar{h}_{\mathrm{minU}}$: $\Phi = 6, \Lambda = 7$, $N = 4.5$



(c) Case with $N > \Lambda$: $\Phi = 6, \Lambda = 4.5$, $N = 7$

Figure B.1: Non-existence of separating equilibrium. No separating equilibrium exists for "log-log-log" preferences given by equation (B.6). In all cases shown, $w_L = 3$ and $w_H = 6$.

In this functional form, the neighbourhood consumption $\bar{h}$ confers utility only through comparison to a yet broader average consumption, $\bar{\bar{h}}$, which may be taken to be the average over all neighbourhoods. A new consistency condition states this additional relationship,

$$\bar{\bar{h}} = \langle \bar{h} \rangle$$

and the comparison between neighbourhoods is captured in the final term of the utility function, $N(\bar{h}, \bar{\bar{h}})$. For instance, a form similar to that analysed in Section B.4.2 is

$$U(x,h,\bar{h}) = \Phi \log(x) + \Lambda \log\left(1 + \frac{h}{\bar{h}}\right) + N \log\left(1 + \frac{\bar{h}}{\bar{\bar{h}}}\right) \tag{B.13}$$

This utility function provides a more natural limit to the benefit obtained in equilibrium from neighbourhood consumption when the number of neighbourhoods is finite. Nevertheless, it is shown on page 166 in Appendix B.5 that there is still no separating equilibrium for households with these preferences. The proof is similar to the case of absolute benefits, above.

B.4.5 "Log-log-exp" preferences with two types

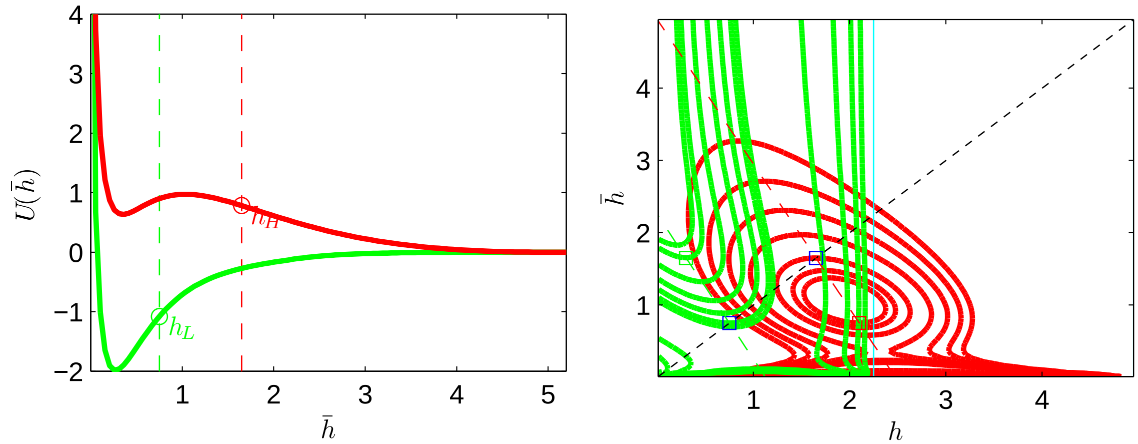
In this section and the next, other convenient functional forms described in Section B.3 are used to vary the qualitative assumptions on utility. Using the inverse exponential form for $N(\cdot)$ imposes a bound on the benefits from living in an affluent neighbourhood, which may be a more defensible assumption and circumvents one apparent problem with the specification give in equation (B.6). For simplicity, consider again the case in which $N(\cdot)$ depends only on absolute consumption of one's neighbours, but now with the following form:

$$U(x,h,\bar{h}) = \Phi \log(x) + \Lambda \log\left(1 + \frac{h}{\bar{h}}\right) - N \exp\left(-\nu \bar{h}\right) \tag{B.14}$$

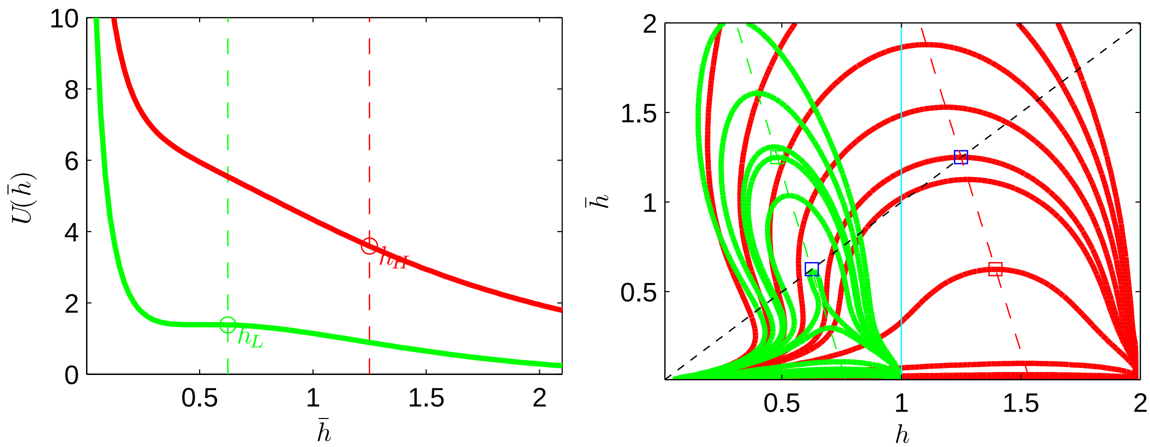
Topologically this form is richer than equation (B.6), with multiple inflection points in the indirect utility $U(w,\bar{h})$. Numerical analysis indicates that it also is incompatible with a separating equilibrium. Figure B.2 on page 163 shows some parameter sets for which in (a) both types prefer to switch neighbourhoods if assigned to their $\bar{h}_{\mathrm{eq}}$, and in (b) the high type prefers to switch and the low type prefers to stay.

B.5 Proofs

Proposition B.5.1. (Decreasing leisure in economy with a continuum of types) *If $F(x)$ is concave and either (a) $H(h,\bar{h})$ takes the form $H = f\left(h - \bar{h}\right)$ or (b) $H(h,\bar{h})$ takes the form $H = f\left(\frac{h}{\bar{h}}\right)$ and $p = 0$, then leisure x is decreasing in w amongst interior equilibria.*



(a) Case with $\Phi = 4$, $\Lambda = 4$, $N = 10$, $\nu = 2$, $w_L \approx 2.5$, and $w_H \approx 5$



(b) Case with $\Phi = 3$, $\Lambda = 10$, $N = 17$, $\nu = 3$, $w_L = 1$, and $w_H = 2$

Figure B.2: Non-existence of separating equilibrium. No separating equilibrium exists for "log-log-exp" preferences given by equation (B.14).

Proof. For interior equilibria, $F'(x) = wH_h(h,\bar{h})$ and $h = \bar{h}\forall w$. When $H = f\left(h - \bar{h}\right)$, taking a derivative gives

$$\frac{dx}{dw} = \frac{f_0'}{F''} < 0$$

For $H = f\left(\frac{h}{\bar{h}}\right)$,

$$\frac{dx}{dw} = \frac{f_0'}{\bar{h}F''}\left[1 - \frac{d\bar{h}}{dw}\right]$$

and if $p(\bar{h}) = 0$,

$$\frac{d\bar{h}}{dw} = 1 - x - w\frac{dx}{dw}$$

Combining these expressions gives

$$\frac{dx}{dw} = \frac{1-x}{F'' - \frac{f_0'}{\bar{h}}w^2} < 0$$

□

Proposition B.4.1 on page 159.

Proof. According to equation (B.7) and equation (B.5), the interior equilibrium choices of type w can be written:

$$h_{\text{eq}} = \bar{h}_{\text{eq}} = \frac{\Lambda}{2\Phi + \Lambda}w \tag{B.15}$$

$$x_{\text{eq}} = \frac{2\Phi}{2\Phi + \Lambda} \tag{B.16}$$

This says that whenever a separating equilibrium exists, households of type w will always be seen to populate the same kind of neighbourhood, regardless of which other types also exist.

First, note that the equilibrium cannot include corner allocations. According to equation (B.7), the choice of h is interior whenever $\bar{h} \leq \frac{\Lambda}{\Phi}w$. Since $\bar{h}_{\text{eq}}$ given in equation (B.15) always satisfies $\bar{h}_{\text{eq}} \leq \frac{\Lambda}{\Phi}w$ and since, again according to equation (B.7), $h^\star > 0$ if $\bar{h} = 0$, allocations in the separating equilibrium must be interior.

Now, a sufficient condition for the existence of a separating equilibrium is that each type prefers to remain in its own neighbourhood. Formally, the net benefit ΔU from moving to the other available neighbourhood and choosing a new level of housing there must be negative for each type:

$$\begin{aligned}
\Delta U_L \equiv U\left(w_L, \bar{h}_H\right) - U\left(w_L, \bar{h}_L\right) &\leq 0 \quad \text{and} &\text{(B.17)}\\
\Delta U_H \equiv U\left(w_H, \bar{h}_L\right) - U\left(w_H, \bar{h}_H\right) &\leq 0 &\text{(B.18)}
\end{aligned}$$

where $\bar{h}_L$ and $\bar{h}_H$ are the equilibrium neighbourhood housing choices in equation (B.15). These conditions can be evaluated using equation (B.9) with the convenient notation $\Theta \equiv \frac{w_H}{w_L} > 1$ and $B \equiv h\text{eq}/w = \frac{\Lambda}{2\Phi+\Lambda} < 1$:

$$\begin{aligned}
\Delta U_L &= \Phi \log\left(1+\frac{h_H}{w_L}\right) + \Lambda \log\left(1+\frac{w_L}{h_H}\right) + N \log\left(h_H\right) \\
&\quad -\Phi \log\left(1+\frac{h_L}{w_L}\right) - \Lambda \log\left(1+\frac{w_L}{h_L}\right) - N \log\left(h_L\right) \\
&= \Phi \log\left(\frac{1+B\Theta}{1+B}\right) + \Lambda \log\left(\frac{1+1/B\Theta}{1+1/B}\right) + N \log\left(\Theta\right) \\
&= \Phi \log\left(\frac{1+B\Theta}{1+B}\right) + \Lambda \log\left(\frac{1}{\Theta}\frac{1+B\Theta}{1+B}\right) + N \log\left(\Theta\right) \\
&= \left[\Phi+\Lambda\right] \log\left(\frac{1+B\Theta}{1+B}\right) + \left[N-\Lambda\right] \log\left(\Theta\right) \qquad \text{(B.19)}
\end{aligned}$$

Similarly,

$$\begin{aligned}
\Delta U_H &= \Phi \log\left(1+\frac{h_L}{w_H}\right) + \Lambda \log\left(1+\frac{w_H}{h_L}\right) + N \log\left(h_L\right) \\
&\quad -\Phi \log\left(1+\frac{h_H}{w_H}\right) - \Lambda \log\left(1+\frac{w_H}{h_H}\right) - N \log\left(h_H\right) \\
&= \Phi \log\left(\frac{1+B/\Theta}{1+B}\right) + \Lambda \log\left(\frac{1+\Theta/B}{1+1/B}\right) - N \log\left(\Theta\right) \\
&= \Phi \log\left(\frac{1}{\Theta}\frac{B+\Theta}{1+B}\right) + \Lambda \log\left(\frac{B+\Theta}{1+B}\right) - N \log\left(\Theta\right) \\
&= \left[\Phi+\Lambda\right] \log\left(\frac{B+\Theta}{1+B}\right) - \left[N+\Phi\right] \log\left(\Theta\right)
\end{aligned}$$

Note that the first term in equation (B.19) must be positive, since $\Theta > 1$. Whenever $N > \Lambda$ the second term is also positive and $\Delta U_L > 0$, which means that the low type will always prefer to move up to the high types's neighbourhood. This makes the separating equilibrium impossible when $N < \Lambda$, *i.e.* for agents who value neighbourhood-level benefits sufficiently more than they value their status within a neighbourhood.

The two terms of ΔU_H can also be unambiguously signed; the first is always positive and the second always negative. I now show that when the low types are content in their neighbourhood, the high types cannot be content in theirs.

A necessary equilibrium condition follows from combining the two inequalities in equation (B.17) and equation (B.18) into the weaker requirement that

$$\Delta U_H + \Delta U_L \leq 0$$

which becomes

$$\begin{aligned}
\Delta U_L + \Delta U_H &= [\Phi + \Lambda] \log\left(\frac{1+B\Theta}{1+B}\right) + [N - \Lambda] \log(\Theta) \\
&\quad + [\Phi + \Lambda] \log\left(\frac{B+\Theta}{1+B}\right) - [N + \Phi] \log(\Theta) \\
&= [\Phi + \Lambda] \log\left(\frac{1+B\Theta}{1+B}\,\frac{B+\Theta}{1+B}\,\frac{1}{\Theta}\right) \\
&= [\Phi + \Lambda] \log\left(\frac{\Theta\left[1+B^2\right] + \left[\Theta + \frac{1}{\Theta}\right] B\Theta}{\Theta\left[1+B^2\right] + 2B\Theta}\right) \qquad \text{(B.20)}
\end{aligned}$$

Because $\Theta + \frac{1}{\Theta} > 2$ for all $\Theta > 1$, the argument of log in equation (B.20) is always greater than 1; thus $\Delta U_L + \Delta U_H > 0$. Therefore, there is no separating equilibrium. □

Proposition B.5.2. *When group entry (land) is costless and preferences are given by equation* (B.13)*, there is no discrete separating group Nash equilibrium with two types.*

Proof. The proof closely resembles that of Proposition B.4.1; differences are noted here.

Let the equilibrium neighbourhoods be h_L and h_H according to equation (B.15). The global reference level can then be expressed

$$\bar{\bar{h}} = \Lambda \frac{w_L + w_H}{2\Lambda + 4\phi}$$

The conditions for an equilibrium then become:

$\Delta U_L = [\Phi + \Lambda] \log\left(\frac{\Phi + \Lambda \frac{1}{2}[1+\Theta]}{\Phi + \Lambda}\right) - \Lambda \log(\Theta) + N \log\left(\frac{1+3\Theta}{3+\Theta}\right) \leq 0$

$\Delta U_H = -[\Phi + \Lambda] \log\left(\frac{\Phi + \Lambda}{\Phi + \Lambda \frac{1}{2}\left[1 + \frac{1}{\Theta}\right]}\right) + \Lambda \log(\Theta) - N \log\left(\frac{1+3\Theta}{3+\Theta}\right) \leq 0$

Again, a weaker necessary condition that follows from combining these two inequalities, $\Delta U_H \leq 0 \leq -\Delta U_L$, is that

$$\Delta U_H + \Delta U_L \leq 0$$

$$\begin{aligned}
\Delta U_H + \Delta U_L &= [\Phi+\Lambda]\log\left(\frac{\Phi+\Lambda\frac{1}{2}\left[1+\frac{1}{\Theta}\right]}{\Phi+\Lambda}\frac{\Phi+\Lambda\frac{1}{2}[1+\Theta]}{\Phi+\Lambda}\right)\\
&= [\Phi+\Lambda]\log\left(\frac{\Phi^2+\Phi\Lambda\frac{1}{2}\left[2+\Theta+\frac{1}{\Theta}\right]+\Lambda^2\frac{1}{4}\left[2+\frac{1}{\Theta}+\Theta\right]}{\Phi^2+2\Phi\Lambda+\Lambda^2}\right)\leq 0
\end{aligned}$$

Because $\Theta+\frac{1}{\Theta}>2$ for all $\Theta>1$, the argument of log is always greater than 1 and therefore the above inequality is impossible. There is no separating equilibrium. □

Lemma B.5.3. (A useful exponential form) *Let* $\Psi(a,b)\equiv\frac{1}{b}-\frac{a}{e^{ab}-1}$. *Then for a and b positive,* $\lim_{b\to 0}\Psi(a,b)=\frac{1}{2}a$ *and* $\Psi(a,b)$ *is always positive. For* $\Psi(-a,b)=\frac{1}{b}-\frac{a}{1-e^{-ab}}$ *and a and b positive,* $\lim_{b\to 0}\Psi(-a,b)=-\frac{1}{2}a$ *and* $\Psi(-a,b)$ *is always negative. Furthermore,* $\frac{d\Psi(a,b)}{da}>0$, $\frac{d\Psi(a,b)}{db}<0$, $\frac{d\Psi(-a,b)}{da}<0$, $\frac{d\Psi(-a,b)}{db}<0$.

Proof. For $a>0$, $b>0$, the function $b\cdot\Psi(a,b)$ is nonnegative iff $ab\leq e^{ab}-1$. For $ab=0$, this is an equality. For $ab>0$, the slope of the right hand side strictly dominates the slope of the left hand side. Therefore the inequality holds for all $ab\geq 0$. By similar reasoning, the inequality $ab\geq 1-e^{-ab}$ holds for all positive ab.

Using a Taylor expansion to find the limits,

$$\begin{aligned}
\Psi(a,b) &= \frac{1}{b}-\frac{a}{ab+\frac{1}{2}a^2b^2+\frac{1}{3!}a^3b^3+\ldots}\\
&= \frac{b+\frac{1}{2}ab^2+\frac{1}{3!}a^2b^3+\ldots-b}{b^2+\frac{1}{2}ab^3+\frac{1}{3!}a^2b^4+\ldots}\\
&= \frac{\frac{1}{2}a+\frac{1}{3!}a^2b+\ldots}{1+\frac{1}{2}ab+\frac{1}{3!}a^2b^2+\ldots}\\
&\to \begin{cases} 0 & \text{as } a\to 0\\ \frac{1}{2}a & \text{as } b\to 0\end{cases}
\end{aligned}$$

A similar transformation finds the limits for $\Psi(-a,b)$. The above results can be used to sign the first derivatives as follows:

$$\begin{aligned}\frac{d}{da}\Psi(a,b) &= -\frac{1}{e^{ab}-1}+\frac{abe^{ab}}{[e^{ab}-1]^2}\\ &= -\frac{1}{e^{ab}-1}\left[1-\frac{ab}{1-e^{-ab}}\right]\\ &= -\frac{1}{e^{ab}-1}\Psi(-ab,1)>0\end{aligned}$$

and

$$\begin{aligned}\frac{d}{db}\Psi(a,b) &= -\frac{1}{b^2}+\frac{a^2e^{ab}}{[e^{ab}-1]^2}\\ &= -\frac{1}{b^2}+\frac{a^2}{[e^{ab}-1][1-e^{-ab}]}\\ &= -\frac{1}{b^2}+\frac{a^2}{e^{ab}+e^{-ab}-2}\\ &= -\frac{1}{b^2}+\frac{a^2}{\frac{2a^2b^2}{2!}+\frac{2a^4b^4}{4!}+\frac{2a^6b^6}{6!}\dots}\frac{a^2}{e^{ab}+e^{-ab}-2}\\ &= -\frac{1}{b^2}\left[1-\frac{1}{1+\frac{2a^2b^2}{4!}+\frac{2a^4b^4}{6!}+\dots}\right]<0\end{aligned}$$

For the modified function $\Psi(-a,b)$, both derivatives are negative:

$$\begin{aligned}\frac{d}{da}\Psi(-a,b) &= -\frac{1}{1-e^{-ab}}-\frac{abe^{-ab}}{[1-e^{-ab}]^2}\\ &= -\frac{1}{1-e^{-ab}}\left[1+\frac{ab}{e^{ab}-1}\right]<0\end{aligned}$$

and

$$\frac{d}{db}\Psi(-a,b) = -\frac{1}{b^2}\left[1-\frac{1}{1+\frac{2a^2b^2}{4!}+\frac{2a^4b^4}{6!}+\dots}\right]<0$$

It follows from these monotonic properties that $\Psi(a,b)\in(0,\frac{1}{2}a)$ and $\Psi(-a,b)\in(-\frac{1}{2}a,0)$.

$\square$

B.6 Construction of equilibrium

This section outlines the steps taken to compute a separating equilibrium in the numerical examples to follow. Given exogenous parameters including the range of types w_L, w_H, values for $x(w)$, $w_m = \max\{w_L, w_0\}$, $\bar{\bar{h}}$, and $r - p_0$ can be directly computed from equations (3.18), (3.19), (3.24), and (3.25). From these, the remaining allocations $\bar{h}(w)$ follow using (3.23) . What remains is to calculate a price schedule $p(\bar{h})$.

In order to select a value of p_0 sufficiently high to clear the market both for the most affluent desired neighbourhoods, *i.e.,* $p\left(\bar{h}(w_H)\right) > 0$, and for the least affluent neighbourhood, *i.e.,* $p(\bar{h}_{\min}) \geq 0$, limiting values of p_0 for both conditions must be calculated and the higher of the two adopted. First, to ensure that there is a nonnegative price for the highest type, I impose $\bar{h}_{\max} = \bar{h}(w_H)$ and $p(\bar{h}_{\max}) = 0$ in (3.20), giving

$$\bar{h}_{\max} = r + w_H - w_0$$

Then (3.22) can be evaluated at this upper limit in order to solve for p_0:

$$\begin{aligned}
0 = p(\bar{h}_{\max}) &= p_0 - \bar{h}_{\max} + \frac{N}{\Lambda\lambda}\log\left(1 + \frac{\bar{h}_{\max}}{\bar{\bar{h}}}\right) \\
\frac{N}{\Lambda\lambda}\log\left(1 + \frac{r + w_H - w_0}{\bar{\bar{h}}}\right) &= [r - p_0 + w_H - w_0] \\
1 + \frac{r + w_H - w_0}{\bar{\bar{h}}} &= \exp\left(\frac{\Lambda\lambda}{N}[r - p_0 + w_H - w_0]\right) \\
r &= \bar{\bar{h}}\left[\exp\left(\frac{\Lambda\lambda}{N}[r - p_0 + w_H - w_0]\right) - 1\right] + w_0 - w_H
\end{aligned}$$

Hence,

$$\begin{aligned}
p_0 &= r - [r - p_0] \\
&= \bar{\bar{h}}\left[\exp\left(\frac{\Lambda\lambda}{N}[r - p_0 + w_H - w_0]\right) - 1\right] + w_0 - w_H - [r - p_0] \qquad \text{(B.21)}
\end{aligned}$$

Since the value of $[r - p_0]$ is already calculated in terms of exogenous parameters (B.21) provides a lower bound on the constant p_0. To calculate a second lower bound satisfying the condition $p(\bar{h}_{\min}) = 0$, note that the consumption level implied by this condition is $\bar{h}_{\min} = r$. Thus, the minimum p_0 can again be calculated in terms of $[r - p_0]$:

$$
\begin{aligned}
0 &= p(r) \\
&= p_0 - r + \frac{N}{\Lambda\lambda}\log\left(1+\frac{r}{\bar{\bar{h}}}\right) \\
\frac{N}{\Lambda\lambda}\log\left(1+\frac{r}{\bar{\bar{h}}}\right) &= r - p_0 \\
1+\frac{p_0+[r-p_0]}{\bar{\bar{h}}} &= \exp\left([r-p_0]\frac{\Lambda\lambda}{N}\right) \\
&= p_0+[r-p_0]-\frac{N}{\Lambda\lambda}\log\left(1+\frac{p_0+[r-p_0]}{\bar{\bar{h}}}\right) \\
\rightarrow 1+\frac{p_0+[r-p_0]}{\bar{\bar{h}}} &= \exp\left([r-p_0]\frac{\Lambda\lambda}{N}\right) \\
p_0 &= \bar{\bar{h}}\left[\exp\left([r-p_0]\frac{\Lambda\lambda}{N}\right)-1\right]-[r-p_0] \qquad \text{(B.22)}
\end{aligned}
$$

For $r - p_0$ sufficiently high for a real solution $\bar{h}(w)$, above, a market-clearing price schedule can now be found by setting p_0 to the larger of the two values in (B.21) and (B.22). The price schedule follows from (3.22).

Bibliography for Appendix to Chapter 3

Clark, A., and A. Oswald, Comparison-concave utility and following behaviour in social and economic settings, *Journal of Public Economics*, *70*, 133–155, 1998.

Eaton, B. C., and M. Eswaran, Well-Being and Affluence in the Presence of a Veblen Good, 2006.

Scotchmer, S., Local Public Goods and Clubs, *Handbook of Public Economics*, *4*, 1997–2042, 2002.

Tversky, A., and D. Kahneman, Loss Aversion in Riskless Choice: A Reference-Dependent Model, *The Quarterly Journal of Economics*, *106*, 1039–1061, 1991.

Appendix C

Appendix to Chapter 4

C.1 Detailed Tables

Below are more detailed versions of estimation results presented in the main body of the paper. For space reasons, tables exclude coefficients of the set of demographic, individual, and household controls used for all models. The complete table is available from the author.

Table C.1: Complete regression results for weather effects on survey-reported SWL, happiness, health, trust, and income. Trust-G is the general social trust question, while trust-N is the stated trust in neighbours. The dependent variable is indicated at the left end of each row. All coefficients are raw ordered logit coefficients, except for the regressions for income, in which case OLS coefficients with robust standard errors are shown.
Significance: 1% 5% 10%

		clouds	clouds (7 days)	T_{high} (°C)	T_{low} (°C)	rain (mm)	snow (cm)	log(HH inc)	health	trust-N	constant	controls	mnth f.e.	stn f.e.	mnthStn f.e.	clustering	survey	obs.	pseudo-R^2	$N_{clusters}$	R^2(adj)
(1)	SWL		−.77									✓					G19	6359	.014		
			(.22)																		
(2)	SWL		−.43									✓					E2	1632	.031		
			(.36)																		
⟨1-2⟩	SWL		−.68									✓					⟨2⟩	7991			
			(.19)																		
(3)	SWL		−.94					.64				✓					G19	5167	.018		
			(.24)					(.11)													
(4)	SWL		−.52					.47				✓					E2	1496	.033		
			(.38)					(.16)													
⟨3-4⟩	SWL		−.81					.59				✓					⟨2⟩	6663			
			(.20)					(.091)													
(5)	SWL		−.78					.36	2.81	.51		✓					G19	5161	.056		
			(.24)					(.11)	(.15)	(.17)											
(6)	SWL		−.49					.34	1.66	.42		✓					E2	1495	.043		
			(.38)					(.15)	(.28)	(.14)											
⟨5-6⟩	SWL		−.70					.35	2.55	.46		✓					⟨2⟩	6656			
			(.20)					(.091)	(.13)	(.11)											
(7)	SWL	−.19	−.81	.002	−.0006	.001	−.008	.42	2.85			✓					G19	4956	.055		
		(.15)	(.26)	(.009)	(.009)	(.004)	(.016)	(.12)	(.16)												
(8)	SWL	−.12	−.58	−6e-05	.007	−.007	−.003	.40	1.70			✓					E2	1495	.042		
		(.22)	(.39)	(.011)	(.012)	(.010)	(.024)	(.15)	(.28)												
⟨7-8⟩	SWL	−.17	−.74	.001	.002	−.0001	−.006	.41	2.58			✓					⟨2⟩	6451			
		(.12)	(.21)	(.007)	(.007)	(.004)	(.013)	(.094)	(.14)												
(19)	SWL		−.83									✓	✓			mnth	G19	6359	.015	12	
			(.31)																		
(20)	SWL		−.57									✓	✓			mnth	E2	1632	.033	8	
			(.47)																		
⟨19-20⟩	SWL		−.75									✓	✓			mnth	⟨2⟩	7991			
			(.26)																		
(21)	SWL		−.99					.64				✓	✓			mnth	G19	5167	.020	12	
			(.29)					(.15)													
(22)	SWL		−.69					.47				✓	✓			mnth	E2	1496	.035	8	
			(.40)					(.12)													
⟨21-22⟩	SWL		−.89					.54				✓	✓			mnth	⟨2⟩	6663			
			(.23)					(.094)													
(23)	SWL		−.91					.36	2.81	.51		✓	✓			mnth	G19	5161	.057	12	
			(.26)					(.14)	(.14)	(.19)											
(24)	SWL		−.68					.33	1.66	.44		✓	✓			mnth	E2	1495	.045	8	
			(.39)					(.13)	(.26)	(.12)											
⟨23-24⟩	SWL		−.84					.34	2.56	.46		✓	✓			mnth	⟨2⟩	6656			
			(.22)					(.094)	(.12)	(.098)											

Continued on next page

		clouds	clouds (7 days)	T_{high} (°C)	T_{low} (°C)	rain (mm)	snow (cm)	log(HH inc)	health	trust-N	constant	controls	mnth f.e.	stn f.e.	mnthStn f.e.	clustering	survey	obs.	pseudo-R^2	$N_{clusters}$	R^2(adj)
(25)	SWL	−.23	−.87	−.004	−.009	.0002	−.010	.42	2.84			✓	✓			mnth	G19	4956	.057	12	
		(.21)	(.27)	(.007)	(.007)	(.003)	(.012)	(.13)	(.14)												
(26)	SWL	−.17	−.69	.0006	.001	−.008	−.003	.38	1.70			✓	✓			mnth	E2	1495	.044	8	
		(.22)	(.44)	(.012)	(.009)	(.008)	(.030)	(.14)	(.25)												
⟨25-26⟩	SWL	−.20	−.82	−.003	−.005	−.0008	−.009	.40	2.58			✓	✓			mnth	⟨2⟩	6451			
		(.15)	(.23)	(.006)	(.006)	(.003)	(.011)	(.096)	(.12)												
(37)	SWL		−.71									✓		✓		stn	G19	6334	.020	50	
			(.26)																		
(38)	SWL		−.23									✓		✓		stn	E2	1594	.036	22	
			(.30)																		
⟨37-38⟩	SWL		−.50									✓		✓		stn	⟨2⟩	7928			
			(.20)																		
(39)	SWL		−.84					.67				✓		✓		stn	G19	5147	.025	50	
			(.26)					(.13)													
(40)	SWL		−.18					.51				✓		✓		stn	E2	1461	.039	22	
			(.32)					(.17)													
⟨39-40⟩	SWL		−.58					.61				✓		✓		stn	⟨2⟩	6608			
			(.20)					(.10)													
(41)	SWL		−.65					.39	2.84	.50		✓		✓		stn	G19	5141	.062	50	
			(.31)					(.12)	(.12)	(.16)											
(42)	SWL		−.20					.38	1.74	.38		✓		✓		stn	E2	1460	.049	22	
			(.31)					(.15)	(.26)	(.17)											
⟨41-42⟩	SWL		−.42					.38	2.64	.44		✓		✓		stn	⟨2⟩	6601			
			(.22)					(.092)	(.11)	(.12)											
(43)	SWL	−.14	−.68	−.003	.007	.0006	−.012	.45	2.89			✓		✓		stn	G19	4928	.063	49	
		(.10)	(.28)	(.009)	(.008)	(.004)	(.020)	(.13)	(.12)												
(44)	SWL	−.013	−.25	−.007	.015	−.010	−.003	.41	1.76			✓		✓		stn	E2	1460	.048	22	
		(.22)	(.32)	(.013)	(.013)	(.011)	(.024)	(.17)	(.26)												
⟨43-44⟩	SWL	−.12	−.49	−.004	.009	−.0006	−.008	.44	2.70			✓		✓		stn	⟨2⟩	6388			
		(.094)	(.21)	(.007)	(.007)	(.004)	(.015)	(.10)	(.10)												
(55)	SWL		−.47									✓			✓	mnthStn	G19	5144	.027	169	
			(.34)																		
(56)	SWL		−.65									✓			✓	mnthStn	E2	1245	.033	44	
			(.54)																		
⟨55-56⟩	SWL		−.52									✓			✓	mnthStn	⟨2⟩	6389			
			(.29)																		
(57)	SWL		−.71					.67				✓			✓	mnthStn	G19	4040	.033	152	
			(.35)					(.13)													
(58)	SWL		−.56					.72				✓			✓	mnthStn	E2	1122	.036	42	
			(.52)					(.20)													
⟨57-58⟩	SWL		−.67					.68				✓			✓	mnthStn	⟨2⟩	5162			
			(.29)					(.11)													
(59)	SWL		−.67					.35	2.95	.62		✓			✓	mnthStn	G19	4017	.073	150	
			(.37)					(.12)	(.17)	(.20)											
(60)	SWL		−.58					.56	1.53	.48		✓			✓	mnthStn	E2	1122	.045	42	
			(.55)					(.20)	(.23)	(.23)											
⟨59-60⟩	SWL		−.64					.41	2.44	.56		✓			✓	mnthStn	⟨2⟩	5139			
			(.31)					(.10)	(.13)	(.15)											
(61)	SWL	−.23	−.67	−.004	−.011	.004	−.010	.42	2.99			✓			✓	mnthStn	G19	3833	.074	143	
		(.19)	(.38)	(.012)	(.013)	(.006)	(.035)	(.13)	(.17)												
(62)	SWL	−.35	−.58	−.006	.009	−.011	−.037	.61	1.58			✓			✓	mnthStn	E2	1122	.044	42	

Continued on next page

		clouds	clouds (7 days)	T_{high} (°C)	T_{low} (°C)	rain (mm)	snow (cm)	log(HH inc)	health	trust-N	constant	controls	mnth f.e.	stn f.e.	mnthStn f.e.	clustering	survey	obs.	pseudo-R^2	$N_{clusters}$	R^2(adj)
		(.22)	(.53)	(.014)	(.013)	(.010)	(.041)	(.21)	(.23)												
⟨61-62⟩	SWL	−.29	−.64	−.005	−.001	3e-05	−.021	.47	2.47			✓			✓	mnthStn	⟨2⟩	4955			
		(.14)	(.31)	(.009)	(.009)	(.005)	(.027)	(.11)	(.14)												
(100)	happy		−.28									✓			✓	mnthStn	G19	5169	.048	169	
			(.38)																		
(101)	happy		−.31					.63				✓			✓	mnthStn	G19	4052	.057	152	
			(.43)					(.14)													
(102)	happy		−.27					.35	2.63	.37		✓			✓	mnthStn	G19	4029	.107	150	
			(.43)					(.14)	(.19)	(.23)											
(103)	happy	−.11	−.15	.016	−.040	.005	−.041	.31	2.67			✓			✓	mnthStn	G19	3846	.108	143	
		(.17)	(.48)	(.015)	(.020)	(.007)	(.029)	(.16)	(.19)												
(163)	health		−.33									✓			✓	mnthStn	G19	5200	.030	169	
			(.30)																		
(164)	health		.29									✓			✓	mnthStn	E2	1247	.024	44	
			(.50)																		
⟨163-164⟩	health		−.16									✓			✓	mnthStn	⟨2⟩	6447			
			(.26)																		
(165)	health		−.30					.90				✓			✓	mnthStn	G19	4071	.037	152	
			(.33)					(.14)													
(166)	health		.50					.60				✓			✓	mnthStn	E2	1124	.032	42	
			(.56)					(.17)													
⟨165-166⟩	health		−.095					.78				✓			✓	mnthStn	⟨2⟩	5195			
			(.28)					(.10)													
(167)	health		−.30					.86		.87		✓			✓	mnthStn	G19	4071	.041	152	
			(.33)					(.13)		(.18)											
(168)	health		.52					.56		.24		✓			✓	mnthStn	E2	1124	.033	42	
			(.56)					(.17)		(.14)											
⟨167-168⟩	health		−.088					.75		.49		✓			✓	mnthStn	⟨2⟩	5195			
			(.28)					(.10)		(.11)											
(169)	health	.006	−.26	−.004	−.001	.006	−.017	.95				✓			✓	mnthStn	G19	3885	.039	145	
		(.19)	(.37)	(.012)	(.015)	(.004)	(.026)	(.14)													
(170)	health	.022	.53	−.006	−.010	−.002	.080	.62				✓			✓	mnthStn	E2	1124	.037	42	
		(.23)	(.56)	(.015)	(.016)	(.019)	(.016)	(.17)													
⟨169-170⟩	health	.013	−.015	−.005	−.005	.006	.053	.82				✓			✓	mnthStn	⟨2⟩	5009			
		(.15)	(.31)	(.009)	(.011)	(.004)	(.014)	(.11)													
(235)	trust-N		−.86									✓			✓	mnthStn	G19	2140	.035	100	
			(.53)																		
(236)	trust-N		−.64									✓			✓	mnthStn	E2	1250	.072	44	
			(.44)																		
⟨235-236⟩	trust-N		−.73									✓			✓	mnthStn	⟨2⟩	3390			
			(.34)																		
(237)	trust-N		−.19					.59				✓			✓	mnthStn	G19	1558	.038	82	
			(.61)					(.25)													
(238)	trust-N		−.46					.86				✓			✓	mnthStn	E2	1125	.086	42	
			(.48)					(.18)													
⟨237-238⟩	trust-N		−.35					.76				✓			✓	mnthStn	⟨2⟩	2683			
			(.38)					(.15)													
(239)	trust-N		−.31					.49	1.06			✓			✓	mnthStn	G19	1558	.044	82	
			(.62)					(.25)	(.24)												
(240)	trust-N		−.48					.83	.45			✓			✓	mnthStn	E2	1124	.087	42	
			(.49)					(.18)	(.26)												

Continued on next page

		clouds	clouds (7 days)	T_{high} (°C)	T_{low} (°C)	rain (mm)	snow (cm)	log(HH inc)	health	trust-N	constant	controls	mnth f.e.	stn f.e.	mnthStn f.e.	clustering	survey	obs.	pseudo-R^2	$N_{clusters}$	R^2(adj)
⟨239-240⟩	trust-N		−.42					.71	.78			✓			✓	mnthStn	⟨2⟩	2682			
			(.39)					(.15)	(.18)												
(241)	trust-N	.13	−.43	.006	.017	−.016	.013	.46	1.12			✓			✓	mnthStn	G19	1479	.045	77	
		(.26)	(.71)	(.018)	(.024)	(.007)	(.042)	(.26)	(.26)												
(242)	trust-N	.18	−.77	.018	−.019	.039	.077	.88	.42			✓			✓	mnthStn	E2	1124	.092	42	
		(.24)	(.51)	(.015)	(.015)	(.014)	(.042)	(.17)	(.26)												
⟨241-242⟩	trust-N	.15	−.66	.013	−.009	−.005	.045	.75	.77			✓			✓	mnthStn	⟨2⟩	2603			
		(.18)	(.41)	(.012)	(.013)	(.006)	(.030)	(.14)	(.18)												
(307)	trust-G		.67								−1.65	✓			✓	mnthStn	G19	2534	.093	166	
			(.72)								(.52)										
(308)	trust-G		−.11								.24	✓			✓	mnthStn	E2	1219	.079	44	
			(.65)								(.26)										
⟨307-308⟩	trust-G		.24								−.14	✓			✓	mnthStn	⟨2⟩	3753			
			(.48)								(.23)										
(309)	trust-G		1.16					.61			−4.18	✓			✓	mnthStn	G19	1964	.112	148	
			(.77)					(.23)			(1.14)										
(310)	trust-G		.22					1.31			−6.29	✓			✓	mnthStn	E2	1103	.109	42	
			(.67)					(.19)			(1.00)										
⟨309-310⟩	trust-G		.62					1.02			−5.38	✓			✓	mnthStn	⟨2⟩	3067			
			(.50)					(.15)			(.75)										
(311)	trust-G		1.50					.33	.75	2.85	−4.92	✓			✓	mnthStn	G19	1957	.203	146	
			(.84)					(.23)	(.29)	(.31)	(1.21)										
(312)	trust-G		.26					1.11	1.17	1.23	−7.24	✓			✓	mnthStn	E2	1102	.142	42	
			(.67)					(.21)	(.39)	(.20)	(1.01)										
⟨311-312⟩	trust-G		.74					.75	.90	1.70	−6.29	✓			✓	mnthStn	⟨2⟩	3059			
			(.52)					(.16)	(.23)	(.17)	(.78)										
(313)	trust-G	−.23	1.51	.012	−.040	−.022	−.011	.53	1.17		−4.31	✓			✓	mnthStn	G19	1865	.128	139	
		(.34)	(.92)	(.026)	(.030)	(.009)	(.057)	(.24)	(.30)		(1.22)										
(314)	trust-G	.45	.26	−.041	.008	.025	−.020	1.28	1.25		−6.85	✓			✓	mnthStn	E2	1102	.126	42	
		(.38)	(.70)	(.021)	(.022)	(.020)	(.042)	(.20)	(.42)		(.96)										
⟨313-314⟩	trust-G	.068	.72	−.020	−.009	−.014	−.017	.98	1.20		−5.87	✓			✓	mnthStn	⟨2⟩	2967			
		(.25)	(.56)	(.016)	(.018)	(.008)	(.034)	(.16)	(.25)		(.76)										
(381)	log(HH inc)		−.048								4.36	✓			✓	mnthStn	G19	4209		169	.237
			(.065)								(.11)										
(382)	log(HH inc)		−.12								4.77	✓			✓	mnthStn	E2	1141		44	.222
			(.069)								(.029)										
⟨381-382⟩	log(HH inc)		−.082								4.74	✓			✓	mnthStn	⟨2⟩	5350			
			(.047)								(.028)										
(383)	log(HH inc)		−.039						.19	.12	3.98	✓			✓	mnthStn	G19	4195		168	.258
			(.063)						(.029)	(.034)	(.10)										
(384)	log(HH inc)		−.12						.14	.11	4.59	✓			✓	mnthStn	E2	1140		44	.240
			(.069)						(.037)	(.023)	(.045)										
⟨383-384⟩	log(HH inc)		−.074						.17	.11	4.49	✓			✓	mnthStn	⟨2⟩	5335			
			(.047)						(.023)	(.019)	(.041)										
(385)	log(HH inc)	−.041	−.054	.002	5e-06	−.0006	−.008		.20		4.07	✓			✓	mnthStn	G19	4001		160	.256
		(.033)	(.072)	(.002)	(.002)	(.001)	(.005)		(.030)		(.11)										
(386)	log(HH inc)	−.066	−.12	.002	−.003	−.003	−.012		.16		4.65	✓			✓	mnthStn	E2	1140		44	.232
		(.049)	(.068)	(.002)	(.002)	(.003)	(.004)		(.037)		(.068)										
⟨385-386⟩	log(HH inc)	−.049	−.087	.002	−.001	−.0008	−.010		.19		4.49	✓			✓	mnthStn	⟨2⟩	5141			
		(.027)	(.049)	(.001)	(.001)	(.0009)	(.003)		(.023)		(.058)										

Table C.2: Climate and satisfaction with life. Covariates include local climatic expectations in the form of probabilities and means for each station's overall climate (YEAR) and for its averages for the month (MONTH) and day (DAY) of the interview. Standard errors are calculated with clustering at the level of the fixed effects (f.e.) indicated. Results in this table are summarised in Table 4.8 on page 93. Significance: 1% 5% 10%

	YEAR: $\langle T_{max} \rangle$ (°C)	YEAR: $\langle T_{min} \rangle$ (°C)	YEAR: days sun	MONTH: days sun	MONTH: sun fraction	MONTH: $\langle T \rangle$ (°C)	MONTH: rain>5mm	MONTH: snow>5cm	DAY: precipitation	DAY: $\langle T_{max} \rangle$ (°C)	DAY: $\langle T_{min} \rangle$ (°C)	clouds (7 days)	log(HH inc)	controls	f.e./clustering	survey	obs.
(1)	.013	−.010	−.003										.59	✓	mnth	G19	1930
	(.042)	(.013)	(.002)										(.17)				
(2)	.23	−.013	−.008										.52	✓	mnth	E2	355
	(.071)	(.026)	(.005)										(.26)				
⟨1-2⟩	.068	−.011	−.004										.57	✓	mnth	⟨2⟩	2285
	(.036)	(.012)	(.002)										(.14)				
(3)	.013	−.014	.0002	.060	−.024	.001	.025	.003					.59	✓	mnth	G19	1930
	(.048)	(.017)	(.003)	(.13)	(.030)	(.035)	(.055)	(.15)					(.17)				
(4)	.24	−.059	−.019	.038	.037	.088	.078	−.12					.58	✓	mnth	E2	355
	(.10)	(.046)	(.008)	(.18)	(.056)	(.051)	(.13)	(.14)					(.24)				
⟨3-4⟩	.052	−.019	−.002	.053	−.010	.029	.033	−.059					.59	✓	mnth	⟨2⟩	2285
	(.044)	(.016)	(.003)	(.11)	(.026)	(.029)	(.051)	(.10)					(.14)				
(5)	.037	−.002	.002						.015	−.012	−.008		.60	✓	mnth	G19	1930
	(.048)	(.020)	(.005)						(.005)	(.082)	(.075)		(.17)				
(6)	.21	−.050	−.017						−.011	.094	−.002		.51	✓	mnth	E2	355
	(.078)	(.043)	(.010)						(.018)	(.056)	(.062)		(.24)				
⟨5-6⟩	.085	−.011	−.002						.013	.060	−.004		.57	✓	mnth	⟨2⟩	2285
	(.041)	(.018)	(.004)						(.005)	(.046)	(.048)		(.14)				
(7)	.034	−.013	−.003											✓	mnth	G19	2388
	(.035)	(.012)	(.003)														
(8)	.20	−.004	−.006											✓	mnth	E2	386
	(.075)	(.021)	(.006)														
⟨7-8⟩	.063	−.011	−.004											✓	mnth	⟨2⟩	2774
	(.032)	(.010)	(.003)														
(9)	.034	−.013	−.002	.090	−.022	−.001	.032	.059						✓	mnth	G19	2388
	(.046)	(.018)	(.003)	(.11)	(.025)	(.033)	(.049)	(.12)									
(10)	.19	−.054	−.016	.020	.015	.10	.006	−.064						✓	mnth	E2	386
	(.085)	(.037)	(.009)	(.20)	(.066)	(.049)	(.15)	(.14)									

Continued on next page

	YEAR: $\langle T_{max} \rangle$ (°C)	YEAR: $\langle T_{min} \rangle$ (°C)	YEAR: days sun	MONTH: days sun	MONTH: sun fraction	MONTH: $\langle T \rangle$ (°C)	MONTH: rain>5mm	MONTH: snow>5cm	DAY: precipitation	DAY: $\langle T_{max} \rangle$ (°C)	DAY: $\langle T_{min} \rangle$ (°C)	clouds (7 days)	log(HH inc)	controls	f.e./clustering	survey	obs.
⟨9-10⟩	.070	−.021	−.003	.075	−.017	.031	.030	.010						✓	mnth	⟨2⟩	2774
	(.040)	(.016)	(.003)	(.094)	(.023)	(.027)	(.047)	(.091)									
(11)	.039	.002	−.003						.013	.044	−.064			✓	mnth	G19	2388
	(.039)	(.020)	(.004)						(.004)	(.075)	(.074)						
(12)	.20	−.044	−.012						−.006	.048	.035			✓	mnth	E2	386
	(.080)	(.039)	(.009)						(.017)	(.062)	(.073)						
⟨11-12⟩	.069	−.008	−.004						.012	.046	−.014			✓	mnth	⟨2⟩	2774
	(.035)	(.018)	(.004)						(.004)	(.048)	(.052)						
(15)				−.016	.004	−.010	.026	−.11					.60	✓	stn	G19	3631
				(.044)	(.011)	(.015)	(.036)	(.083)					(.12)				
(16)				−.024	−.004	.010	.066	.020					.42	✓	stn	E2	907
				(.048)	(.024)	(.018)	(.087)	(.074)					(.18)				
⟨15-16⟩				−.020	.003	−.002	.032	−.038					.54	✓	stn	⟨2⟩	4538
				(.032)	(.010)	(.012)	(.033)	(.055)					(.100)				
(17)									−.004	−.070	.075		.62	✓	stn	G19	9654
									(.004)	(.024)	(.027)		(.092)				
(18)									.015	−.028	.047		.52	✓	stn	E2	2562
									(.005)	(.045)	(.050)		(.12)				
⟨17-18⟩									.003	−.061	.069		.59	✓	stn	⟨2⟩	12216
									(.003)	(.021)	(.024)		(.074)				
(21)				−.006	.008	−.016	.025	−.12						✓	stn	G19	4457
				(.033)	(.009)	(.014)	(.031)	(.072)									
(22)				−.010	−.003	.005	.068	.053						✓	stn	E2	996
				(.044)	(.024)	(.021)	(.093)	(.060)									
⟨21-22⟩				−.007	.007	−.010	.029	−.018						✓	stn	⟨2⟩	5453
				(.026)	(.008)	(.012)	(.029)	(.046)									
(23)									−.004	−.051	.056			✓	stn	G19	11924
									(.003)	(.023)	(.026)						
(24)									.012	−.027	.045			✓	stn	E2	2829
									(.005)	(.039)	(.043)						
⟨23-24⟩									.0002	−.045	.053			✓	stn	⟨2⟩	14753
									(.003)	(.020)	(.022)						
(29)									−.013	−.30	.32		.75	✓	mnthStn	G19	6309
									(.013)	(.14)	(.16)		(.10)				
(30)									.002	−.22	.29		.57	✓	mnthStn	E2	1781

Continued on next page

	YEAR: $\langle T_{max} \rangle$ (°C)	YEAR: $\langle T_{min} \rangle$ (°C)	YEAR: days sun	MONTH: days sun	MONTH: sun fraction	MONTH: $\langle T \rangle$ (°C)	MONTH: rain>5mm	MONTH: snow>5cm	DAY: precipitation	DAY: $\langle T_{max} \rangle$ (°C)	DAY: $\langle T_{min} \rangle$ (°C)	clouds (7 days)	log(HH inc)	controls	f.e./clustering	survey	obs.
									(.013)	(.21)	(.24)		(.16)				
⟨29-30⟩									−.005	−.27	.31		.70	✓	mnthStn	⟨2⟩	8090
									(.009)	(.12)	(.13)		(.086)				
(39)									−.002	−.28	.31			✓	mnthStn	G19	8207
									(.008)	(.11)	(.13)						
(40)									.038	−.23	.30			✓	mnthStn	E2	2045
									(.015)	(.18)	(.21)						
⟨39-40⟩									.007	−.26	.31			✓	mnthStn	⟨2⟩	10252
									(.007)	(.095)	(.11)						
(41)									−.027	−.46	.52	−.53	.68	✓	mnthStn	G19	4040
									(.016)	(.17)	(.18)	(.34)	(.13)				
(42)									−.040	−.31	.38	−.62	.71	✓	mnthStn	E2	1122
									(.032)	(.21)	(.24)	(.57)	(.20)				
⟨41-42⟩									−.030	−.41	.47	−.56	.69	✓	mnthStn	⟨2⟩	5162
									(.014)	(.13)	(.15)	(.29)	(.11)				