



Universiteit
Leiden
The Netherlands

Essays on wealth, health, and data collection

Kools, L.

Citation

Kools, L. (2018, November 21). *Essays on wealth, health, and data collection*. Meijers-reeks.
Retrieved from <https://hdl.handle.net/1887/67120>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/67120>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/67120> holds various files of this Leiden University dissertation.

Author: Kools, L.

Title: Essays on wealth, health, and data collection

Issue Date: 2018-11-21

Essays on wealth, health, and
data collection

Essays on wealth, health, and data collection

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 21 november 2018
klokke 15.00 uur

door

LIEKE KOOLS
geboren te Boarnsterhim
in 1990

PROMOTORES: prof.dr. K.P. Goudswaard
prof.dr. C.L.J. Caminada
CO-PROMOTOR: prof.dr. M.G. Knoef

PROMOTIECOMMISSIE: prof.dr. R.J.M. Alessie (Rijksuniversiteit Groningen)
dr. V. Angelini (Rijksuniversiteit Groningen)
prof.dr. P.W.C. Koning (VU Amsterdam en Universiteit Leiden)
prof.dr. J.I. van der Rest

Preface

Even though I have often complained that a PhD is written in solitary confinement, in reality the existence of this thesis should be accredited to many others than myself.

First of all, I would like to thank my promotors, Kees Goudswaard and Koen Caminada, for taking the gamble of appointing me a PhD position and providing me with ample freedom to explore outside the boundaries of both the country and the profession. I am not sure if this thesis would have existed had I not been granted such possibilities, but I am absolutely sure this thesis would not have existed in absence of my daily supervisor Marike Knoef. Marike, no matter how pessimistic I entered a meeting with you, your enthousiasm and constant stream of new ideas always made me leave fueled up and ready to tackle the challenges ahead. I would also like to thank the members of the committee, Rob Alessie, Viola Angelini, Pierre Koning, and Jean-Pierre van der Rest for taking the time to read the manuscript of this thesis and providing helpful comments.

To my colleagues, thank you for the daily lunch breaks, trips to the Hoogvliet, Chinese language exchange, and interesting discussions. To Jim in particular, thank you for sharing your \LaTeX template. A special thanks goes out to my co-authors Jochem de Bresser, Pierre Koning and Dana Thomson. Thank you for sharing your wisdom with me. Dana, thank you for bringing me inspiration in Southampton.

I leave this PhD not only with new knowledge, but also with many new friends. I am grateful to both the members of the PhD committee and the members of SMO promovendi for providing me the necessary work-related distraction. You made the infamous valley of shit smell a lot

nicer. Daphne, I have learned a lot from your can-do mentality. Gintare and Muied, I am so happy to have met you at Wooldridge's course and to have shared all the fun times and struggles of PhD life with you. I wish I could have chosen you both as a paranimph, but well, Gintare saw me first.

I could not have thought of a better complement to the quiet PhD life, than the chaos at 'de Schans'. All the delicious food, dancing, dating, and long conversations kept me energized and ready to focus. Hanna, who but you, who has known me for most of my life and knows my *hersenspingsels* better than I do, could be my paranimph?

Thank you to all my friends for supporting me throughout. And thank you Tomas for teaching me not to take myself too serious.

Contents

Preface	vii
1 Introduction	1
1.1 Spreading wealth shocks over the life course	2
1.2 The relationship between health and consumption	3
1.3 Counteracting moral hazard problems	5
1.4 Innovations in data collection methods	6
2 Cutting One's Coat According to One's Cloth - How did the Great Recession affect retirement resources and expenditure goals?	9
2.1 Introduction	10
2.2 Pension reforms and the crisis	14
2.2.1 Public pension	15
2.2.2 Occupational pensions	15
2.2.3 Voluntary private savings	17
2.3 Theory	17
2.4 Data	20
2.4.1 Survey data	20
2.4.2 Administrative data	24
2.5 Empirical strategy	27
2.5.1 Size and nature of co-movements between wealth and retirement expenditure goals	28
2.5.2 Co-movements and the development of retirement savings adequacy	30
2.6 Results	32
2.6.1 Results on size and nature of co-movements	32
2.6.2 Simulation results on retirement savings adequacy	39
2.7 Conclusion	40
2.A Funding ratios	43

2.B	Development consumer confidence	44
2.C	Thinking about retirement and difficulty of the questions	45
2.D	Descriptives socio-economic variables	48
2.E	Descriptive statistics assets and debts	50
2.F	Assumptions underlying the annuities	52
2.G	Quantile models of changes in expenditure goals	57
2.H	Models of changes in expenditure goals that only control for family composition	59
2.I	Estimation results SUR model	62
2.J	Distribution of differences between goals and annuities	71
3	Health and Consumption Preferences - Estimating the Health State Dependence of Utility using Equivalence Scales	73
3.1	Introduction	74
3.2	Method	76
3.2.1	Theoretical framework	77
3.2.2	Empirical model	81
3.2.3	Underlying assumptions and threats to identification	83
3.3	Data	87
3.3.1	Sample selection	87
3.3.2	Health	88
3.3.3	Income and assets	90
3.3.4	Making ends meet	91
3.4	Results	92
3.4.1	Baseline	92
3.4.2	Medical costs	94
3.4.3	Alternative health measures	97
3.4.4	Robustness checks	99
3.5	Conclusion	104
3.A	Additional descriptive statistics	107
3.A.1	Health	107
3.A.2	Income and assets	111
3.A.3	Positive and negative feelings	113
4	Graded Return-to-Work as a Stepping Stone to Full Work Resumption	117
4.1	Introduction	118
4.2	Institutional setting	122
4.2.1	Gatekeeper Protocol	122
4.2.2	Private insurance of continued wage payments and case management	124
4.3	Data	126
4.3.1	Characteristics of sick-listed employees	126

4.3.2	Characteristics of case managers	132
4.3.3	Setup of graded return-to-work trajectories	134
4.4	Estimation strategy	136
4.4.1	Specification of the effect of graded work	139
4.4.2	Specification of the effect of timing and initial degree of graded work	141
4.5	Results	142
4.5.1	Overall effects of graded return-to-work	142
4.5.2	Effects of the timing and initial level of graded work	147
4.5.3	Effects for different types of medical conditions	149
4.5.4	Sensitivity tests	151
4.6	Conclusion	161
4.A	Additional data descriptives	163
4.B	Additional results	165
5	One-stage versus two-stage cluster sampling, a simulation study	177
5.1	Introduction	178
5.2	Method	182
5.2.1	Generating a realistic synthetic population	182
5.2.2	Calculating the minimal number of clusters	189
5.3	Data description	195
5.4	Results	201
5.4.1	Baseline analysis	201
5.4.2	Increasing sample sizes per cluster	204
5.5	Discussion	206
5.6	Conclusion	207
5.A	Maps of Oshikoto	209
5.B	Data description for synthetic populations 2-5	211
5.C	Results for each of the 5 synthetic populations	220
5.C.1	Baseline results	220
5.C.2	Additional results	225
5.D	Additional statistics	230
	Bibliography	231
	Nederlandse samenvatting	245
	Curriculum Vitae	253

List of Tables

2.1	Descriptive statistics retirement expenditure goals	23
2.2	Descriptive statistics of assets, debt and annuities	26
2.3	Shocks to annuities and changes in expenditure goals	33
2.4	Shocks to annuities and changes in expenditure goals - heterogeneity by age group	35
2.5	Shocks to annuities and changes in expenditure goals – heterogeneity by income	36
2.6	Aggregate simulation results: differences between annuities and expenditure goals	38
2.7	Descriptive statistics of minimum expenditures during retirement	45
2.8	Descriptives of self-reported question difficulty	46
2.9	Descriptives of retirement expenditure goals by level of question difficulty	47
2.10	Descriptives of socio-economic variables: individual-level variables	48
2.11	Descriptives of socio-economic variables: household-level variables	49
2.12	Descriptive statistics of assets and debts	51
2.13	Quantiles of shocks to annuities and changes in expenditure goals – heterogeneity by age	57
2.14	Quantiles of shocks to annuities and changes in expenditure goals – heterogeneity by income	58
2.15	Shocks to annuities and changes in expenditure goals	59
2.16	Shocks to annuities and changes in expenditure goals – heterogeneity by age	60
2.17	Shocks to annuities and changes in expenditure goals– heterogeneity by income	61
2.18	Joint models of annuities and retirement expenditures – expenditure equations – men	63

2.19	Joint models of annuities and retirement expenditures – expenditure equations – women	64
2.20	Joint models of annuities and retirement expenditure goals – annuity equations – pensions.	67
2.21	Joint models of annuities and retirement expenditure goals – annuity equations – pensions + wealth + housing.	68
2.22	Error correlations	70
3.1	Summary statistics	88
3.2	Baseline results	93
3.3	Test for coverage of health care costs	96
3.4	Results for alternative health measures	98
3.5	Specification checks	100
3.6	Robustness check: method	103
3.7	Summary statistics for health variables	107
3.8	Income split up by component and country	112
3.9	Percentage of respondents experiencing positive and negative feelings.	114
3.10	Results including measures of positive and negative feelings	115
4.1	Descriptive statistics sick-listed employees.	128
4.2	Descriptive statistics of the 68 case managers	133
4.3	Variation in grading practices across case managers.	135
4.4	Overall effects of graded return-to-work on full work resumption	144
4.5	Effect of starting graded return-to-work one week later or at a higher starting level: IV estimates.	148
4.6	IV estimation results on work resumption for different medical conditions.	150
4.7	Sensitivity tests for specialization effects – Return-to-work within one year – Overall effect	153
4.8	Sensitivity tests for specialization effects – Return-to-work within one year – Timing and intensity	154
4.9	Sensitivity tests for the importance of case manager quality	157
4.10	First stage results for detailed subcategories	159
4.11	First stage results for rough subcategories	160
4.12	Data selection steps	163
4.13	Additional case manager characteristics	164
4.14	Reasons for sick-listing	164
4.15	Effect of graded return-to-work when started in week 1-52, including coefficients on control variables.	166
4.16	Effect of graded return-to-work when started in week 1-26, including coefficients on control variables.	167

4.17	Effect of starting moment of graded return-to-work, including coefficients on control variables	168
4.18	Effect of initial degree of graded return-to-work, including coefficients on control variables.	169
4.19	IV estimation results for different medical conditions – weeks worked.	171
4.20	Overall results (1-52 weeks) using different cut-offs for the minimum number of clients per case manager	172
4.21	Overall results (1-26 weeks) using different cut-offs for the minimum number of clients per case manager	173
4.22	Weeks waited results using different cut-offs for the minimum number of clients per case manager	174
4.23	Degree grading results using different cut-offs for the minimum number of clients per case manager	175
5.1	Overview of data sources used for simulations	183
5.2	Household types	185
5.3	Within EA clustering scenarios	194
5.4	Summary statistics - Synthetic population 1	197
5.5	Required number of clusters - Baseline	202
5.6	Required number of clusters - increasing cluster size	205
5.7	summary statistics for synthetic populations 2-5	211
5.8	Required number of clusters - Baseline - Population 1	220
5.9	Required number of clusters - Baseline - Population 2	221
5.10	Required number of clusters - Baseline - Population 3	222
5.11	Required number of clusters - Baseline - Population 4	223
5.12	Required number of clusters - Baseline - Population 5	224
5.13	Required number of clusters - Increasing cluster size - Population 1	225
5.14	Required number of clusters - Increasing cluster size - Population 2	226
5.15	Required number of clusters - Increasing cluster size - Population 3	227
5.16	Required number of clusters - Increasing cluster size - Population 4	228
5.17	Required number of clusters - Increasing cluster size - Population 5	229
5.18	Observed ICCs in DHS Namibia 2013	230

List of Figures

2.1	Kernel regressions of retirement expenditure goals on household income (shaded areas are 95% confidence bands). . . .	23
2.2	Simulated preparedness for retirement: fraction that cannot afford retirement expenditure goals (spikes are 90% CIs) . .	38
2.3	Relationship between regulatory solvency and relative decline in funding ratios during 2008, source: DNB (2009) . . .	43
2.4	Development of people's confidence in their financial situation in the next 12 months, by education level	44
2.5	Indexation scenario's and realisations, after Knoef et al. (2016b)	54
2.6	Simulated differences between annuities and expenditure goals	71
3.1	Benchmark levels making ends meet	81
3.2	Prevalence of limitations in activities of daily living across age	89
3.3	'Making ends meet' across health and income	92
3.4	Limitations in instrumental activities of daily living across age	108
3.5	Limitations in physical mobility across age	108
3.6	Different type of limitations across age	110
3.7	Chronic diseases across age	110
4.1	Time line of the gatekeeper protocol.	123
4.2	Histogram of application moments.	129
4.3	Recovery patterns by type of diagnosis	129
4.4	Survival and hazard rates for individuals with and without graded return-to-work in first year of absence	131
4.5	Percentage of individuals participating in graded return-to-work per week.	135
4.6	Cumulative effects of graded work per sick weeks	146
4.7	Duration coefficients	170

4.8	Propensities to treat before scaling.	170
5.1	An example of 2-stage sampling of ordered household points and selection.	190
5.2	An example of 1-stage sampling segments.	191
5.3	An example of the spatial clustering of wealth within one EA for scenario 1 (top) and scenario 3 (bottom).	193
5.4	Prevalence of characteristics per EA - Synthetic population 1	196
5.5	ICC (◆) and boxplots of prevalences per EA/segment - Synthetic population 1	199
5.6	Population density in Oshikoto expressed in people per pixel (roughly 100m ²).	209
5.7	Distance to major roads (km).	210
5.8	Prevalence of characteristics per EA - Synthetic population 2	212
5.9	Prevalence of characteristics per EA - Synthetic population 3	213
5.10	Prevalence of characteristics per EA - Synthetic population 4	214
5.11	Prevalence of characteristics per EA - Synthetic population 5	215
5.12	ICC (◆) and boxplots of prevalences per EA/segment - Synthetic population 2	216
5.13	ICC (◆) and boxplots of prevalences per EA/segment - Synthetic population 3	217
5.14	ICC (◆) and boxplots of prevalences per EA/segment - Synthetic population 4	218
5.15	ICC (◆) and boxplots of prevalences per EA/segment - Synthetic population 5	219