CALY-SWE GODA ÅR

Capability for b r o a d e r cost-effectiveness in public health and social welfare

Developing, valuing, and applying capability-adjusted life years Sweden

Kaspar Walter Meili

UMEÅ UNIVERSITY

UMEÅ UNIVERSITY

CAPABILITY FOR BROADER COST-EFFECTIVENESS IN PUBLIC HEALTH AND SOCIAL WELFARE

Developing, valuing, and applying capability-adjusted life years Sweden (CALY-SWE)

Kaspar Walter Meili

Department of Epidemiology and Global Health
Umeå 2024

This work is protected by the Swedish Copyright Legislation (Act 1960:729)
Dissertation for PhD
ISBN: 978-91-8070-257-7 (print)
ISBN: 978-91-8070-258-4 (pdf)
ISSN: 0346-6612
New Series Number 2280
Cover design: G. Dekombis, K. Meili. Image back: L. Richter Sundberg, front: K. Meili
Electronic version available at: http://umu.diva-portal.org/
Printed by: CityPrint i Norr AB
Umeå, 2024

To Helen, Johan, and Théo – I love you

To family and friends – I cherish you

To those I never met – Alice, Medy, and Walter

To those passed away – Agi, Beatrice and Kurt

The personal, economic and cultural welfare
of the individual shall be fundamental aims of
public activity.
In particular, the public institutions shall
secure the right to employment, housing and
education, and shall promote social care and
social security, as well as favourable
conditions for good health.

The Swedish Instrument of Government (*1974:152*)

The success of a society is to be
evaluated primarily by the
substantive freedoms that the
members of that society enjoy.

Amartya Sen
Development as Freedom (*p. 18*)

...and in the knowledge that only those
who use their freedom remain free, and
that the strength of a people is
measured by the well-being of its
weakest members.

Federal Constitution of the Swiss Confederation (*SR 101*)

Front Matter

Front Matter **i**
English summary v
Svensk sammanfattning vii
Deutsche Zusammenfassung ix
Supervisors and key roles xi
Studies in the PhD thesis xi
Glossary and abbreviations xii

1 The moral case for broader cost-effectiveness in social welfare **1**
1.1 Cost-effectiveness in social welfare: shortcomings of the current logic 1
1.2 Cost-effectiveness in the healthcare context: cost per quality-adjusted life year (QALY) 2
1.3 Deriving the value of a QALY 3
1.4 Financial footprints in Swedish social welfare 7
1.5 Theoretical and philosophical bases for measuring welfare in economic evaluations 9
1.6 The capability approach 11
1.7 Point of departure: The CALY-SWE proposal 12
1.8 CALY-SWE as an alternative to describe societal welfare 14
1.9 Applying CALY-SWE in an economic cost-effectiveness evaluation 15
1.10 Existing capability measures for economic evaluations 16
1.11 Thesis aim, goals, and objectives 18
1.11.1 Outline of the thesis 19

2 Capability of what: selecting capability attributes and phrasing the statements (S1) **20**
2.1 Overview 20
2.2 Methods 20
2.2.1 Selection of capabilities using a Delphi panel 20
2.2.2 Delphi panel procedure 21

2.2.3 Empirical capability distribution to inform statement phrasing 22
2.3 Results . . . 24
2.3.1 Delphi panel . . . 24
2.3.2 Empirical capability distribution to inform statement phrasing 25
2.3.3 Phrasing the statements . . . 25
2.3.4 Meaning of the highest state (333333) . . . 29

3 For what it is worth: eliciting capability weights (S2) 30
3.1 Overview . . . 30
3.2 Methods . . . 30
3.2.1 Planning and timeline . . . 31
3.2.2 Determining the cornerstones . . . 31
3.2.3 Technical infrastructure for the survey . . . 32
3.2.4 Survey evolution, changes, and improvements in multiple stages . . . 33
3.2.5 TTO iteration procedure . . . 34
3.2.6 Experimental design and coefficient coding . . . 35
3.2.7 Sample size simulation . . . 36
3.2.8 TTO and DCE data quality assessment . . . 37
3.2.9 Model comparison and value set modeling . . . 38
3.3 Results . . . 39
3.3.1 TTO and DCE answers . . . 39
3.3.2 Data quality assessment . . . 40
3.3.3 Model comparison and value set modelling . . . 41

4 How do we fare? Applying CALY-SWE to measure the distribution and inequalities of capability (S3) 44
4.1 Overview . . . 44
4.2 Methods . . . 44
4.2.1 Questionnaire . . . 44
4.2.2 Equity and comparing groups . . . 45
4.2.3 Statistical analysis . . . 45
4.3 Results . . . 46
4.3.1 Attribute level answers . . . 46
4.3.2 Aggregated weights . . . 47
4.3.3 Weight components . . . 48
4.3.4 Counterfactual contributing covariables . . . 48
4.3.5 Population scale CALY consequences . . . 49

5 Is rolling back the payroll tax paying back? Applying CALY-SWE in a cost-effectiveness study of NEET prevention (S4) 54
5.1 Overview . . . 54
5.2 Methods . . . 54
5.2.1 Data sources . . . 54
5.2.2 Modelling NEET proportions . . . 55
5.2.3 Cost-effectiveness . . . 56
5.3 Results . . . 57
5.3.1 Modelling NEET proportions . . . 57
5.3.2 Cost-effectiveness . . . 57

6 Data collection and ethical approval 61
6.1 Methods . . . 61
6.1.1 Datasets . . . 61
6.1.2 Ethical approval . . . 63
6.1.3 Data exclusion . . . 63
6.1.4 Statistical software . . . 63
6.2 Sample representativeness . . . 63

7 Discussion 65
7.1 Main findings . . . 65
7.2 Normative choices in designing and valuing a measure . . . 66
7.3 A democratic paternalistic process (S1) . . . 66
7.3.1 Capabilities vs functionings . . . 68
7.4 Double counting . . . 68
7.5 The CALY-SWE weight scale . . . 69
7.5.1 Anchoring of weights relative to 0 (S2) . . . 69
7.5.2 Incorporating the sufficiency threshold (S1) . . . 70
7.5.3 Anchoring of weights relative to 1 (S2) . . . 71
7.5.4 TTO data quality (S2) . . . 71
7.6 Capability distribution, groups, prioritization and capability shortfalls (S3) . . . 72
7.7 Using CALY-SWE to evaluate the cost-effectiveness of the payroll tax cut (S4) . . . 73
7.7.1 Methodological limitations . . . 73
7.7.2 Policy implications . . . 74
7.8 Comparison to other capability outcome measures . . . 75
7.9 Would Amartya Sen object? . . . 78
7.10 CALY-SWE in the context of the Swedish welfare state . . . 78
7.11 Further limitations and missing pieces . . . 80

7.12 Outlook 81
7.13 Conclusions 82

References **83**

Epilogue **103**

Appendix 1: Additional results **104**

Appendix 2: Timeline CALY-SWE and PhD project activities **106**

Appendix 3: Contributions **107**

Appendix 4: Related resources **108**

8 Appendix 5: CALY-SWE questionnaire in Swedish **109**

List of Figures **111**

List of Tables **113**

List of Infoboxes **115**

Acknowledgements **117**

9 Previous PhD theses from the Department of Epidemiology and Global health **118**

English summary

Spending in social welfare areas such as healthcare, wider public health, education, and social care consumes a major part of the public budget. Cost-effective resource allocation is a moral obligation towards both taxpayers and beneficiaries: tax money should be used efficiently, and it should be transparently accounted for. After all, economical management of resources is important for sustaining future prosperity and for addressing fundamental challenges such as climate change and demographic shift.

In healthcare, cost-effectiveness using quality-adjusted life years (QALYs) is a well-established tool to inform policymakers. Using cost per QALY implies that health-related quality of life (QoL), and not money, is an end on its own. Moreover, cost per QALY allows one to compare unrelated interventions by measuring the effects on the common QALY scale. However, for actors concerned with broader social welfare, such as the Swedish municipalities, QALYs may be less useful because their measurement focus is largely limited to health. Comparable outcome measures for broader social welfare are still sparsely available and employed, and a context-specific measure for Sweden is lacking.

The aim of this thesis was to develop, value, and apply capability-adjusted life years Sweden (CALY-SWE), a QoL outcome measure conceptually based on the capability approach, for broader social welfare and specific for Sweden.

Within study 1, we organized a Delphi panel to select relevant capability attributes and then developed the phrasing for the questionnaire. The resulting questionnaire contains six attributes –health, social relations, financial situation & housing, security, occupation, and political & civil rights – each with three answer levels. The phrasing integrates an implicit threshold so that the sensitivity is focused on the lower range of the scale, thus incorporating equity considerations that relate to sufficientarianism and prioritarianism.

In study 2, we developed a value set consisting of all quality weights for the 729 possible CALY-SWE states. We relied on health economic outcome methodology, namely hybrid modelling of discrete choice and time trade-off data that we collected in a cross-sectional web survey with representative sampling. This value set allows to aggregate the CALY-SWE answers into a single quality weight that can be used in cost-effectiveness analysis to calculate CALYs.

In study 3, we applied the CALY-SWE questionnaire and value set to describe the capability distribution in a cross-sectional representative sample of the Swedish population. In a framework of group comparisons, we estimated capability inequalities and shortfalls for different population groups. The results showed that

there are capability inequalities for *disadvantage groups* as well as for groups with discriminative inequalities – *plurality groups*, for example between lower and higher education.

For study 4, we applied CALY-SWE in a cost-effectiveness application to model the effects of a payroll tax reduction in Sweden from 2007 to 2016 (during the financial crisis) on young people not in employment, education, or training (NEET). The intervention was likely cost-effective from a societal perspective, but only with limited probability from a fiscal perspective, although definite statements regarding cost-effectiveness are challenging because a threshold value for a CALY is still lacking.

The final chapter discusses the measure's development, including normative choices, in relation to the Swedish social welfare and policy context, the capability framework as suggested by Amartya Sen, distributive justice, and other outcome measures in cost-effectiveness evaluations.

Important work remains – for example, assessing psychometric properties, developing the conceptualization of the 0 to 1 anchor scale for capability weights, and assessing a threshold value for a CALY. In conclusion, with the questionnaire development, value set elicitation, and demonstration of applications, important steps for CALY-SWE were accomplished. Cost-effectiveness evaluations in wider social welfare and public health using CALY-SWE are now possible.

Svensk sammanfattning

Kostnader inom välfärdsområden som hälso- och sjukvård, socialt arbete, folkhälsa, utbildning och arbetsmarknad utgör en stor andel av den offentliga budgeten. En kostnadseffektiv resursallokering är en moralisk plikt mot både skattebetalare och de som använder välfärden. Skattepengar ska användas effektivt, och nyttan borde redovisas mer transparent. Till syvende och sist är ekonomisk hushållning viktigt för ett hållbart framtida välstånd och för att kunna möta grundläggande utmaningar som klimatpåverkan och demografiska förändringar.

I hälso- och sjukvården är kostnadseffektivitetsanalyser med kvalitetsjusterade levnadsår (QALY) ett etablerat verktyg för beslutsfattande. Med kostnader i förhållande till vunna QALYs blir hälsorelaterad livskvalitet - inte pengar - ett explicit självändamål. Dessutom möjliggör kostnad per QALY jämförelser av olika interventioner genom att effekterna mäts på en gemensam QALY skala.

För aktörer med verksamhet som berör välfärd i allmänhet, som svenska kommuner, är dock QALYs ett mindre användbart mått eftersom QALYs fokuserar på hälsa. Användbara mått, som beaktar bredare konsekvenser för livskvalitet och som liknar QALYs till konstruktionen, saknas i nuläget. Syftet med avhandlingen var att utveckla, värdera och tillämpa Goda år (CALY-SWE), ett mått som konceptuellt bygger på handlingsfrihetsansatsen (capability approach), för beskrivning och analys av välfärd i Sverige.

I delstudie 1 genomförde vi en Delfi-panel i syfte att välja ut relevanta handlingsfriheter för det nya måttet (capability-adjusted life years Sweden, CALY-SWE). Det resulterade i sex attribut – hälsa, nära relationer, ekonomi och boende, sysselsättning, säkerhet, politiska och medborgerliga rättigheter – som vardera har tre svarsnivåer. Formuleringarna inbegriper ett implicit tröskelvärde med syfte att fokusera sensitiviteten på nedre delen av skalan. På så sätt integreras effektivitets- och jämlikhetaspekter i samhällets prioriteringar.

I delstudie 2 utvecklade vi vikter för de 729 möjliga CALY-SWE tillstånden. Vi använde hälsoekonomiska metoder utvecklade i QALY-studier, nämligen hybridmodellering av DCE- och TTO-svar som vi samlade in genom en tvärsnittsenkät i ett representativt urval. På så vis kan CALY-SWE enkätsvar översättas i en enda *livskvalitetsvikt* som kan användas i kostnadseffektivitetsanalyser för att beräkna Goda År (capability-adjusted life years, CALYs).

I delstudie 3 använde vi CALY-SWE frågor (studie 1) och vikter (studie 2) för att beskriva fördelningen av handlingsfrihet i ett representativt tvärsnittsurval av den svenska befolkningen. I ett ramverk inspirerat av sjukvårdens etiska plattform jämfördes graden av handlingsfrihet och CALYs för olika befolkningsgrupper.

Resultaten visade att det finns brister i handlingsfrihet för grupper som har sämre objektiva förutsättningar för livskvalitet (*nackdelsgrupper*, t.ex. sjukskrivna och arbetslösa) samt för grupper som har ökad risk för diskriminering (*mångfaldsgrupper*, t.ex. kön och födelseland).

I delstudie 4 använde vi CALY-SWE i en kostnadseffektivitetsanalys för att modellera effekterna av sänkta arbetsgivaravgifter under finanskrisen från 2007 till 2016 bland unga som varken arbetar eller studerar (UVAS). Interventionen var säkerligen kostnadseffektiv från ett samhällsperspektiv men med mindre sannolikhet från ett fiskalt perspektiv. Resultaten bör dock tolkas med försiktighet då tröskelvärdet för en CALY ännu inte har beräknats.

Avslutningsvis förs ett resonemang om måttets utveckling i förhållande till normativa frågor som fördelningsrättvisa, det svenska välfärdssystemet inklusive beslutsfattande, samt till andra utfallsmått för ekonomisk utvärdering. Den sammantagna bedömningen är att CALY-SWE förhåller sig väl till handlingsfrihetsansatsen så som den ursprungligen föreslogs av Amartya Sen.

Det krävs fortsatt forskning om CALY-SWE som att utforska psykometriska egenskaper, utveckla konceptualiseringen av ankringsvärdena 1 och 0 för kapabilitetsvikter, och uppskatta ett tröskelvärde för en CALY. Denna avhandling utgör dock ett viktigt steg och kostnadseffektivitetsanalyser inom välfärd och folkhälsa med CALY-SWE är nu genomförbara.

Deutsche Zusammenfassung

Ausgaben im Zusammenhang mit sozialer Wohlfahrt, zum Beispiel für öffentliche Gesundheit, für Ausbildung und im Sozialwesen, beanspruchen einen großen Teil des öffentlichen Haushalts. Eine kosteneffektive Mittelverteilung ist eine moralische Verpflichtung sowohl gegenüber dem Steuerzahler als auch gegenüber den Begünstigten. Steuergelder sollten effizient und transparent genutzt werden. Letztendlich ist ein wirtschaftlicher Umgang mit Ressourcen wichtig, um zukünftig den Wohlstand aufrechtzuerhalten und um grundlegende Herausforderungen wie den Klimaschutz oder den demografischen Wandel anzugehen.

Im Gesundheitswesen ist Kosteneffizienz mittels qualitätskorrigierter Lebensjahre (quality-adjusted life years, QALYs) ein anerkanntes Werkzeug, um politische Entscheidungsträger zu informieren. Die Anwendung von QALYs impliziert, verglichen mit ausschließlich finanziellen Überlegungen, dass die Förderung gesundheitsbezogener Lebensqualität einen Selbstwert hat. Außerdem erlauben Kosten per QALY den Vergleich unterschiedlicher Interventionen, da die Effekte auf der gemeinsamen QALY-Skala gemessen werden. Für Akteure, die sich mit sozialer Wohlfahrt im weiteren Sinne befassen, wie zum Beispiel die schwedischen Gemeinden, sind QALYs hingegen weniger anwendbar, da der Messfokus mehrheitlich auf Gesundheit gerichtet ist. Vergleichbare Messinstrumente für breitere soziale Wohlfahrt sind noch relativ wenig verfügbar und ein kontextspezifisches Instrument für Schweden ist nicht vorhanden.

Die Zielsetzung dieser Doktorarbeit war die Entwicklung, Wertung, und Anwendung von befähigungskorrigierten Lebensjahren Schweden (capability-adjusted life years Sweden, CALY-SWE) für die weitere soziale Wohlfahrt und spezifisch für Schweden.

In Teilstudie 1 organisierten wir ein Delphi-Panel, um relevante Befähigungsattribute auszuwählen. Im Anschluss entwickelten wir die Formulierungen für den Fragebogen. Der resultierende Fragebogen enthält sechs Attribute: Gesundheit, soziale Beziehungen, finanzielle Situation und Wohnen, Sicherheit, Beschäftigung und politische und zivile Rechte, jeweils mit drei Antwortstufen. In die Frageformulierungen integriert sind implizite Schwellen, so dass die Empfindlichkeit auf den unteren Teil der Skala fokussiert ist. Auf diese Weise sind Verteilungsgerechtigkeitsaspekte im Sinne von Suffizientarismus und Prioritärianismus eingebaut.

In Teilstudie 2 entwickelten wir eine Wertemenge für die Gewichtungsfaktoren aller 729 CALY-SWE-Zustände. Wir haben gesundheitsökonomische Berechnungsmethoden angewendet, namentlich Hybridmodellierung von DCE- und

TTO-Daten, die wir in einer Querschnittstudie mit einer repräsentativen Stichprobe gesammelt haben. Diese Wertemenge erlaubt es, CALY-SWE-Antworten in einem einzelnen Gewichtungsfaktor zusammenzufassen, der dann in Kosteneffektivitätsanalysen angewendet werden kann, um befähigungskorrigierte Lebensjahre (CALYs) zu berechnen.

In Teilstudie 3 wandten wir den CALY-SWE-Fragebogen und die Wertemenge in einer querschnittlichen Bevölkerungsstichprobe an, um die Befähigungsverteilung zu untersuchen. Innerhalb eines Rahmenwerks von Gruppenvergleichen schätzten wir Befähigungsunterschiede und Mankos zwischen verschiedenen Gruppen. Die Resultate ließen auf Befähigungsunterschiede für Benachteiligungsgruppen und auch für Pluralitätsgruppen schließen, zum Beispiel zwischen niedriger und höherer Ausbildung.

In Teilstudie 4 wandten wir CALY-SWE an, um die Auswirkungen einer Ermäßigung der Sozialabgaben für Arbeitgeber während der Finanzkrise von 2007 bis 2016 auf das Vorkommen von NEET (junge Erwachsene, die weder arbeiten noch studieren oder in Ausbildung sind) in einer Kosteneffektivitätsanalyse zu untersuchen. Die Steuerermäßigung war wahrscheinlich kosteneffektiv in der Gesellschaftsperspektive, aber nur mit einer gewissen Wahrscheinlichkeit in der Fiskalperspektive, obwohl die Beurteilung wegen des fehlenden Kosteneffektivitäts-Schwellenwertes eines CALY schwierig ist.

Das letzte Kapitel diskutiert die CALY-SWE-Entwicklung und normative Entscheide im Zusammenhang mit dem schwedischen Wohlfahrts- und Entscheidungskontext, dem Befähigungsansatz wie vorgeschlagen von Amartya Sen, Verteilungsgerechtigkeit und anderen Messinstrumenten für Kosteneffektivitätsauswertungen.

Wichtige verbleibende Forschungsarbeit umfasst zum Beispiel die psychometrischen Eigenschaften, die weitere Konzeptualisierung der Ankerwerte 0 und 1 für Befähigungsgewichte und die Festlegung eines Schwellenwertes für ein CALY. Zusammenfassend sind mit der Entwicklung des Fragebogens, mit der Wertemenge und den dargestellten Anwendungen wichtige Entwicklungsschritte für CALY-SWE vollendet. Kosteneffektivitätsanalysen in breiter sozialer Wohlfahrt und öffentlicher Gesundheit mittels CALY-SWE sind nun möglich.

Supervisors and key roles

- **Anna Månsdotter**[1], Principal supervisor
- **Jan Hjelte**[1,2], Assistant supervisor
- **Lars Lindholm**[1], Assistant supervisor, PI CALY-SWE project

Examiner: Anna-Karin Hurtig[1]

Reference person: Henrik Holmberg[1]

[1]Department of Epidemiology and Global Health
Umeå University

[2]Department of Social Work
Umeå University

Studies in the PhD thesis

- **Study 1** (S1):
  An initiative to develop capability-adjusted life years in Sweden (CALY-SWE): Selecting capabilities with a Delphi panel and developing the questionnaire [1].
  (Published)

- **Study 2** (S2):
  Eliciting a value set for the Swedish Capability-Adjusted Life Years instrument (CALY-SWE) [2].
  (Published)

- **Study 3** (S3):
  Capability inequality: Does disadvantage or plurality matter more for policy? A Swedish cross-sectional study on the population distribution and group differences of capability using CALY-SWE.
  (Manuscript submitted)

- **Study 4**: (S4):
  NEET prevention through Sweden's youth payroll tax cut: Evaluating cost-effectiveness using CALY-SWE.
  (Manuscript draft)

Glossary and abbreviations

CALY-SWE (Capability-adjusted life years Sweden)	Name of the questionnaire developed in this thesis and of the research project.
CALY (Capability-adjusted life year)	Time lived adjusted for the capability of the life during this period.
Capability	All possible things individual can be or do; a set of functionings.
Capability-adjusted lifetime	Aggregation of lifetime weighted by capability.
Capability-related quality of life	Quality of life defined by the capability.
Commodities	An economic resource, usually a good, here also money.
CIS (combined inconsistency severity score)	A score to determine the quality of the TTO answers per participant.
DCE (Discrete choice experiment)	A question format where one picks the best of at least two alternatives with different characteristics between the alternatives so that the choice involves a trade-off.
Economic evaluations	Evaluations with the purpose of determining the economic value, for example in monetary terms.
Fair-minded	Acting in interest of all involved and for the greater cause.
Functionings	One particular, realizable being or doing that is part of a capability set.
Hybrid model	A statistical regression model that estimates a logit and linear component simultaneously.
Health-related quality of life	Quality of life defined by health.
Measure	An instrument to measure something, e.g. a questionnaire.
Not in education, training, or employment (NEET)	Concept for referring to youth that do not participate in economic activities or studies.

Instrument	Synonym of measure and questionnaire.
SDG (Sustainable Development Goals)	A worldwide initiative under the umbrella of the United Nations for sustainably developing human societies.
QALY (Quality-adjusted life year)	Time lived adjusted for the quality of the life during this period.
QoL (Quality of life)	The quality of one's life. More abstract, broader than wellbeing.
Quality-adjusted lifetime	Aggregated lifetime weighted by quality.
Questionnaire	Synonym of measure and instrument.
TTO (Time trade-off)	A question format where living time is traded against QoL to figure out how much time one is willing to trade-in for a QoL gain.
S1, S2, S3, S4	Designates the for studies included in the thesis.
Welfare	How well-off an individual or a group of individuals is. Often used in an economic context.
Wellbeing	How well of an individual is. Often used when then individual is in focus.

For references, please refer to the thesis text.

Chapter

1 The moral case for broader cost-effectiveness in social welfare

Solidarity for those in unfavourable circumstances and just conditions for all are important pillars in health [3] and social welfare [4,5]. However, upholding these principles can be costly for society [6]. Optimizing spending efficiency in social welfare is important so that scarce means can be used to improve the situation for more individuals, or to a greater extent for the single individual, or that the savings can be used in other areas or to avoid increasing taxes [7–10]. Moreover, for the sake of justice, transparency, and democratic accountability, it is important to spend money collected from taxpayers efficiently while accounting for what has been achieved [11] and under what ethical judgements it has been spent [12]. In other words, ineffective resource allocation in social welfare would be immoral vis-à-vis potential beneficiaries and funders of the public budget, and cost-effectiveness analyses should inform policy decisions in social welfare.

1.1 Cost-effectiveness in social welfare: shortcomings of the current logic

Measuring cost-effectiveness of social welfare programmes entails assessing the costs as well as the outcomes or benefits that are attributed to the programs [13]. Many financial components, both outcomes and costs, of social welfare programmes are relatively easy to assess, such as workers' salaries. Other factors, such as future incomes of clients, may be harder to assess, but they still fall within

the conceptionally clear categories of incurred costs, avoided costs, and increased future income. Nevertheless, this financially focused conceptional logic fails to account for perhaps the most important component in the calculus, the welfare of the individual, or in other terms, the improvement in quality of life (QoL).

Some of the main actors concerned with social welfare spending in Sweden, namely the municipalities, that are responsible for social care, elderly care, and schooling, among others, mostly follow a cost-focused logic [14,15], but presumably rather out of necessity than ill-practice. Appropriate tools that facilitate using general QoL as the main outcome in cost-effectiveness are largely lacking for Swedish municipalities. This deficiency introduces the risk that decision-makers focus on financial outcomes that are easier to assess than the QoL consequences for the affected individuals. For example, a project with a future decrease in spending but little QoL improvements may be preferred over one with a very large QoL gain but no projected savings. In addition, some areas, such as social care for which the municipalities are responsible, might deal with individuals for whom the relevant outcomes are heterogenous and complex compared with a health-only focus [16], necessitating broader outcome measures.

1.2 Cost-effectiveness in the healthcare context: cost per quality-adjusted life year (QALY)

The Swedish regions are responsible for healthcare and also are important actors in social welfare. In healthcare, cost-effectiveness has a fairly long tradition, and substantial methodological knowledge has accumulated in the field of health economics [13].

Simpler cost-effectiveness analyses assess costs per countable outcomes, such as cost per number of people vaccinated, cost per number of deaths averted, or costs per life-year gained thanks to a treatment. However, decades ago, health economics faced a similar challenge as social welfare does today: how can economic evaluations consider peoples' QoL [17]? Imagine two treatments for a disease that both cost $1,000. The first is a red pill, to be taken daily, with severe side effects, and the second is a green pill, to be taken once, with no side effects and thus better QoL. The QoL difference should matter and the rational choice would be the green pill as it must be taken only once and has no side effects.

The solution adopted for cost-effectiveness in health has been to weight the time lived by QoL and to use quality-adjusted lifetime as the outcome to measure the benefit in cost-effectiveness evaluations [18,19]. Continuing the above example, assume that both pills result in people living on average 10 additional years. People

who take the green pill will have the full QoL and thus accumulate the full 10 life years. Taking the red pill, on the other hand, would result in only 50% of full QoL. Then, living 10 more years would result in a quality-adjusted lifetime of only 5 additional years. Thus, quality-adjusted lifetime considers both length of time and the quality of the time lived. In a health economic cost-effectiveness analysis, we would compare costs in terms of monetary value with the benefits measured in QALYs by calculating the ratio of the two. The red pill would result in costs per QALY of \$1,000/5 QALYs = \$200 per QALY, while the green pill would result in \$1,000/10 QALYs = \$100 per QALY.

Crucially, cost-effectiveness analyses also allow one to compare options for healthcare spending with a seemingly unrelated outcome [13]. For example, the cost-effectiveness of vaccines for influenza, measured in costs per case averted, can only be compared with other types of vaccines against influenza, or possible alternative ways to avert the disease. However, a direct comparison with cost-effectiveness estimates for cardiovascular disease treatments is hardly possible. Using QALYS as the outcomes, the nature and cause of ill-health are irrelevant for comparability, at least as long as QALYS capture all relevant consequences adequately. Similarly can QALYs also be used for measuring effects for indirectly affected individuals such as informal carers [20]. Lastly, making health-related QoL the outcome of interest has ethical implications in the form of a value promise [21]: one might argue that health has a value on its own, and decisions concerning health should not be governed entirely based on financial metrics.

Given those arguments, policy makers faced with resource allocation decisions in healthcare routinely consider cost-effectiveness evidence. Healthcare can be maximized for a fixed budget by calculating the cost per QALY for all spending alternatives and then implementing the ones with the best cost-efficiency until the budget has been consumed [13].

1.3 Deriving the value of a QALY

In practice, one needs a method to derive a quality estimate for a specific living situation for which a QALY should be calculated, also called a state. This quality weight is conventionally defined on a scale from 0 to 1 so that it can be readily multiplied with a span of time lived. The living situations are described by using a number of attributes with a number of levels for each. Individuals usually self-report their living situation by using a questionnaire containing those attributes and answer levels [7,13,22,23].

Box 1.3.1: Questionnaire, Instrument, measure

Can be used interchangeably when referring to a series of questions, often with predetermined, selectable answer alternatives. Questionnaires are usually distributed to members of a sample with the purpose of collecting standardized data. However, measure and instrument are more general terms that can be used for data collection in other contexts, while questionnaire implies that the data is captured with questions.

Box 1.3.2: Attributes and answer levels

The *attributes* of a questionnaire are commonly understood to represent the different questions, and the answer levels represent the different answer options per question. *Dimension* is another term for attribute. The term *level* indicates that the answer alternatives are ordered, for example, a Likert scale.

Box 1.3.3: State, or config

State, or *config*(*uration*) refer to a specific combination of an attribute and levels that describes a living situation in terms of QoL.

Box 1.3.4: Weight, or quality score

For a specific state, this quantity defines QoL, usually on a scale from 0 to 1, with 1 being the best quality and 0 being the worst quality.

Box 1.3.5: Tarif, value set

The set of weights for all possible states [7,13]. For example, a questionnaire with two attributes with each three levels would result in 3^2 = 9 possible states that each would receive a quality weight on the scale of 0 to 1, and the entirety of these weights would form the value set.

Box 1.3.6: Value set elicitation, valuing

The process of deriving the weights.

To derive a value set, a considerable amount of methodological knowledge has been developed in the health economic context, and several methods exist.

They can be broadly categorized along a spectrum of two (partially intersecting) schools of thoughts on how to model and quantify choices.

1) Expected utility theory goes back to the conceptualization by von Neuman & Morgenstern of choices under uncertainty of a rational utility maximiser. This school has found widespread use in economics related to welfare, including health economics [7,24].
2) Random utility theory is rooted in psychology and has been adopted in market research. Lately, it has increasingly been taken up in health economics, including the elicitation of value sets [25,26].

Both schools of thought accommodate stated preferences in behavioural experiments. In such experiments, the participants are presented with tasks where they have to state a preference for one or rank multiple exclusive alternatives. The results can be analysed quantitatively, where the different alternatives or characteristics thereof can be assigned numerical values that denote their importance for guiding the choices of the participants. Similarly, both theories accommodate for the concept of a *latent* utility or variable that can be measured through the experiments.

Box 1.3.7: Latent variable

A internal indicator of individuals that indicates an extent of a trait, but is not directly observable [27]

The archetypical experiment under expected utility theory is the *standard Gamble* [7]. It typically involves a choice that is certain and a choice where two uncertain alternative outcomes can occur with probability p or $(1p)$, respectively. In health economics, the two uncertain outcomes usually involve full health and death, and the certain choice involves living in the health state below full health for which a quality weight is measured. The probability p^* that an individual is indifferent between the two choices (meaning that an individual has difficulties deciding between the choices and deems them equally worthy) corresponds to the quality weight of the health state. The better the evaluated health state is, the higher p^* needs to be to compensate for the risk of dying and to make the uncertain alternative attractive enough.

However, standard gamble has been found to be complex. Torrance et al. [24,28] developed a conceptually simpler alternative that has arguably replaced standard gamble as the quasi gold standard in the elicitation of value sets in health economics, namely time trade-off (TTO). In TTO, the choice is between living a maximum number of years in the impaired health state compared with a living x

number of years in full health that is smaller than the maximum number of years. Both scenarios are usually followed by immediate death. The maximum could be, for example, 10 years. The quality weight of the state under evaluation is then given by $x^*/10$, where x^* is the value where a participant is indifferent between the choices. The underlying assumption is that individuals are willing to trade-off time and live for a shorter period of time if they are compensated with higher QoL. There have also been variants adopted that allow negative valuations of health states – for example, by adding additional time in full health to the impaired state that can compensate quality weights below 0 until a participant becomes indifferent [7].

For both standard gamble and TTO, the value of indifference can be found iteratively by posing the same task repeatedly for an individual while varying p or x until the point of indifference is reached [7]. Alternatively, the value can be varied across different individuals in large samples and then the point of indifference is determined by the value where the choice proportions are equal [26].

In comparison, the random utility theory school of thought archetypically involves only discrete comparisons without dynamically varying numerical characteristics, called discrete choice experiments (DCEs). Arguably the most common form involves a choice between two alternatives. Other forms consist of picking the best alternative among more than two alternatives, picking the worst and the best, and ranking tasks with more than two alternatives. Discrete also entails that no indifference between the alternatives can be indicated by participants in the experiment. The value assigned to the alternatives and their characteristics is solely derived from statistically modelling the latent variable. One challenge with DCEs is that the resulting statistical values for strengths or quality weights of the alternatives are not anchored on a real-world scale in the way TTO is anchored on the time scale; rather, the values represent relative strengths of the involved alternatives. Thus, if DCEs are used to assign quality weights for health states, then additional considerations are needed to anchor the resulting weights – for example, by incorporating a duration as an additional attribute.

The EQ-5D family of QALY instruments has been used extensively to measure and research the preference values of health states. The initial three-level version was developed in the 1990s [22]. A five-level version of the instrument was developed and adopted in the 2010s [29]. Both variants are widely used in health economics, including by government institutions to make funding decisions for medical treatments. The process of developing value sets for the EQ-5D instruments is outlined in *valuation protocols* (EQ-VT). At the time of writing this thesis, the most current one is EQ-VT 2.1 [30,31].

The EQ-VT includes both TTO and DCE tasks and the possibility to combine their analysis in the form of *hybrid* modelling [32]. While TTO values can be used alone to derive a value set, in theory, hybrid modelling allows one to combine and to complement the advantages of both methodologies. For example, the time scale involved in TTO has been found to cause bias [7]. Furthermore, TTO is cognitively challenging. In contrast, DCE results need to be anchored using additional information, for example, via TTO and hybrid modelling.

1.4 Financial footprints in Swedish social welfare

A closer look into the responsibilities and financial consequences regarding social welfare within the Swedish context demonstrates the importance of social welfare spending. The Swedish municipalities are responsible for important areas within social welfare with far-reaching QoL consequences for individuals: schooling, social welfare, and elderly care [33]. On the other hand, healthcare is largely the responsibility of the regions.

Those responsibilities are reflected in the spending of the Swedish municipalities and regions. In 2022, the regions had around 320 billion SEK, corresponding to around 70% of their total costs, accrue in healthcare and dental care. Municipalities had 38% of costs, corresponding to around 293 billion SEK, accrue in healthcare and social care, while the share of education was 43% (332 billion SEK) [34,35]. When dissecting the municipal healthcare spending, the largest shares fall onto care services, namely care for the elderly (18.9%) and for people with disabilities according to the Support and Service to Persons with Certain Functional Disabilities and the Social Insurance Code laws [36,37] (9.2%).

In general, the spending by municipalities is largely in areas where health-related QoL is important but likely does not capture all relevant QoL aspects that go beyond health, such as education. The municipal healthcare and pedagogical spending items (roughly 600 million SEK) even exceed the total spending of the regions of roughly 480 million SEK (Figure 1.1). Thus, QALYs become a relatively "useless currency" [1], and their use within a municipal context comes with the danger of missing relevant non-health consequences and could potentially lead to suboptimal allocation of resources. These considerable amounts of money spent by the municipalities on healthcare, social care, and education, with potential broad QoL impacts, reinforce the arguments for more elaborate cost-effectiveness analyses that can capture the relevant QoL consequences. QoL should become an explicit, intrinsic target on its own that is subject to maximization in cost-effectiveness analyses of social welfare, similarly as in healthcare.

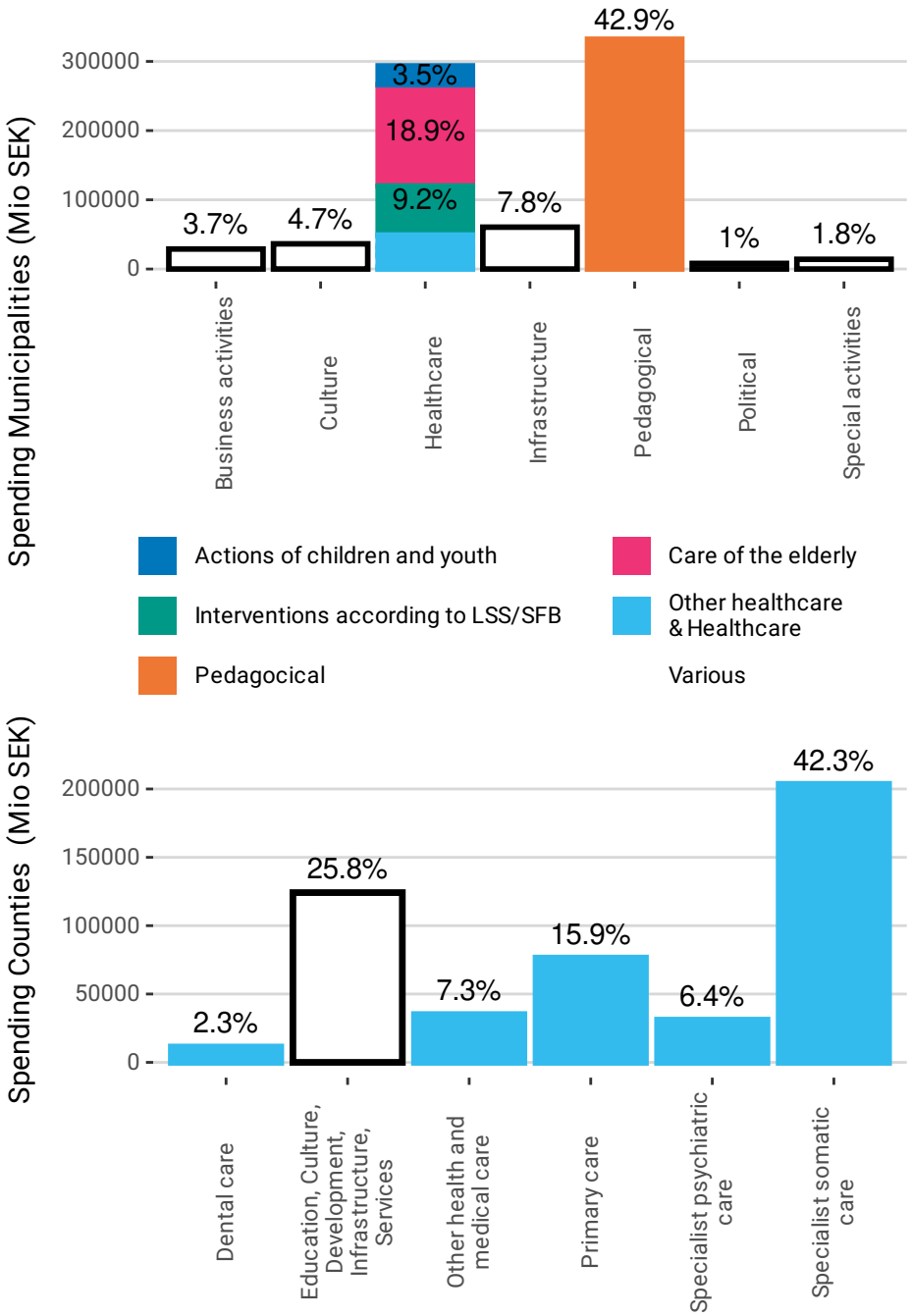

Figure 1.1: Spending of the Swedish municipalities and regions in 2022. Items closely relevant for social welfare are shown shaded. Less-relevant items are shaded white with a black border.

The situation that cost-effectiveness is more explicitly in focus in healthcare is also reflected in the Swedish legal landscape. For healthcare, the Swedish platform on priority setting incorporates three lexicographically ordered principles, namely human dignity, need and solidarity, and, ranked last, an explicit cost-effectiveness principle [9]. In other areas of social welfare [38], explicit cost-effectiveness priority considerations are lacking. The Social care Act that regulates social services instead focuses on individuals' needs [36], which fits well in a context where social care recipients can have complex needs [16], and where social care is the responsibility of municipalities [36,39].

In conclusion, there is a need for an outcome measure for diverse QoL consequences, beyond the health-focused QALY. Such a tool should deliver evidence on cost-effectiveness of existing spending and potential investments in social welfare in Sweden - for example in the municipal context – not only in terms of monetary consequences but also regarding QoL effects.

1.5 Theoretical and philosophical bases for measuring welfare in economic evaluations

Developing an outcome measure requires explicit considerations regarding the normative underpinnings of measurement, which is the subject of the next section. The theory behind economic evaluations is partially based on the field of welfare economics, which is concerned with the distribution of resources and the resulting wellbeing on a societal level [7,40]. It discusses problems such as what is the best distribution between two types of goods between two individuals when maximizing their welfare.

Box 1.5.1: Welfare, wellbeing, and QoL

Those terms have varied connotations and may be understood slightly differently in diverse contexts. Below are some clarifications and definitions for the purposes of the thesis.

(*Social*) *welfare*: In this used for societal-level welfare, as in wider welfare economics. Specifically, this thesis understands social welfare broadly. It encompasses important public affairs that have a direct influence on individual QoL and wellbeing, for example, healthcare, public health, social care, education, and unemployment measures [41]. In other contexts, social welfare may also be used similarly to social care assistance, for example, for public financial support (i.e., welfare payments) for financially deprived individuals [41].

Wellbeing: Describes the state of how well an individual is, implicitly regarding the absence of illness [42]. It is assumed to be largely defined by individuals' QoL, which is used as the preferred expression in this thesis. The reason is that thesis in particular, and capability-adjusted life years Sweden (CALY-SWE) in general, is concerned with individuals' QoL from a societal perspective. In other contexts, however, a distinction between wellbeing and QoL can be important [43,44].

QoL: Refers to how good the life of an individual seems for the individual themself [43,44].

Thus, welfare economics crucially depend on how individual welfare is defined, and how individuals' welfare is aggregated on a societal level.

Utilitarianism is one approach in welfare economics and defines value bases and aggregation rules for performing societal-welfare calculations. Utilitarianism has long been a prominent paradigm in welfare economics and economics in general, as manifested, for example, by the rational *homo economicus* [45]. The defining feature of utilitarianism is that individuals choose their course of action to increase their utility. Utility is valuable in itself, and utilitarianism is consequentialist, although the definition of utility can vary in different streams of the theory [46–48]. Some definitions place utility close to pleasure and are thus of a hedonistic nature [46,47]. Other accounts of utility introduce rules that interact with how valuable consequences are (rule-based utilitarianism) [47,48].

In the context of health economics, *welfarism* is closely related to utilitarianism. While the distinction between welfarism and *extra-welfarism* varies [49], welfarism is commonly described as a paradigm where the only valid information for mea-

suring welfare is an individual's utilities that they derive from their consumption of goods and services [49–51]. These utilities are not directly comparable (interpersonal comparisons of utility levels are disallowed). Instead, *pareto efficiency* should be used to judge whether resources are efficiently allocated: an allocation can only be made more efficient if a change in resource allocation leads exclusively to improvements in an individual's utilities, but no worsening in utility for any individual. If such pareto improvement is not possible, then pareto efficiency is achieved. However, in most real-life economic evaluations, it seems necessary to accept *potential pareto improvements*, where individuals are compensated for accepting a utility loss. Otherwise, changing the resource allocation may prove to be difficult because a lot of possible allocations would produce at least some individuals that are worse off. Thus, accepting potential pareto improvements allows for cost-benefit analyses where the costs and benefits of involved individuals are compared in monetary terms [7,13] and can be maximized. However, critics see this approach as a violation of the normative foundations of welfarism, as the potential pareto payments correspond de facto to interpersonal comparisons [50].

Extra-welfarism, the currently predominant paradigm in health economics, embraces the consideration of 'extra' information beyond individual utility in the equation. This extra information may include an individual's preferences regarding the distribution of welfare, or non-consequential, procedural aspects, among other factors. The accumulation of an individual's utility or other relevant outcomes is explicitly allowed. Other sources of information than an individual for informing welfare are valid [49,50,52]. Cost-effectiveness analyses using QALYs are dependent on accepting extra-welfarism, given that the relevant outcome is not utility but rather health, which is directly compared and aggregated between individuals.

1.6 The capability approach

The capability approach is an alternative paradigm for measuring welfare; it was proposed by Nobel prize laureate Amartya Sen [53–55]. Sen developed the capability approach as an answer to the predominant paradigm of welfarism [52–54]. Sen has criticised utilitarianism and welfarism regarding the limited scope of information that only gives relevance to individual utility, the problems of focusing on maximizing either marginal or total utility, and the difficulties of interpersonal utility comparisons when relying on rank orderings or scaling individual utilities to reflect moral claims. He exemplifies his critique of the falsely luring promise of utility equality with a "pleasure wizard" that is easier to please than a cripple despite already having more utility [54]. Utility maximization would result in the

pleasure wizard receiving more resources, which would surely be unfair.

The capability approach aims to address these shortcomings by making *capability* the end to pursue instead of utility. Capabilities comprise all possible functionings for an individual. Functionings are all achievable beings and doings of an individual and are objectively and interpersonally comparable. Higher capability entails a larger set of opportunities that individuals can freely choose to shape their lives according to their own values. For a given limit of resources and capability, *conversion factors* determine how well individuals can turn commodities into functionings [53,55].

For example, consider a country where the retirement age is 65 years, and another country where the retirement is flexibly regulated, say between 60 and 70 years. In the second country, the capability regarding retirement age is higher ceterus paribus because there are more possibilities from which to choose. Importantly, this assumes that there is an effective freedom to retire early that is not overly hindered by financial circumstances.

The capability approach, whose development was partially paralleled by the development of the notion of extra-welfarism [50], can be interpreted in extra-welfarist terms [51,52], where health capability and not utility becomes a maximization end on its own, and interpersonal comparisons are more readily embraced.

What exact capabilities should factor into the calculus is an important aspect when considering an implementation. Measuring all existing capabilities would be impractical, and many capabilities likely are irrelevant in specific contexts. Hence, the list of capabilities to consider has to be shortened. Some scholars argue that it is possible to derive universal lists of capabilities that are relevant for all humans. A prominent example is Nussbaum's list of 10 central human capabilities (life, bodily health, bodily integrity, senses & imagination & thoughts, emotions, practical reason, affiliation, other species, play, and control over one's environment) [56]. Sen himself has not suggested a specific list: he has argued that the relevant capabilities depend on the context [54,55,57]. For example, it is likely that access to basic necessities such as clean drinking water is less policy-relevant in Sweden compared with some low- and middle-income countries or countries where fresh water is not as readily available.

1.7 Point of departure: The CALY-SWE proposal

In light of the shortcomings of economic evaluations in Swedish municipalities and the philosophical advancements in welfare economics in the form of the capability approach, Månsdotter et al. [58] proposed a capability measure, in the form

of a self-report questionnaire, to measure cost-effectiveness in social welfare and public health. Conceptually, the CALY-SWE proposal highlights the roots in health economics and the similarity to QALYs.

Box 1.7.1: CALY-SWE naming

The initial proposal from 2016 [58] referred to the questionnaire only as CALY. However, a similar term has been used elsewhere previously [59–61], and in hindsight we agree that the initial term is too broad. It better suits the general notion of CALYs, similarly to the usage of the term QALY.

The focus of CALY-SWE is on policy relevance in the Swedish context. For ease of use, an initial design goal was to limit the instrument to the most important few key capabilities. The proposal argues that the selection of capabilities should be context specific; potentially, each country should undertake an effort to develop its specific capability measure. Crucially, the proposal envisioned for Sweden an approach where fair-minded civil-society stakeholders engage in a partially paternalistic, deliberative process in the form of a *Delphi panel* [62]. The starting point for this process was based on a list of capabilities propagated in a 2015 Swedish governmental investigation [63].

Box 1.7.2: Delphi panel

A Delphi panel (also referred to as the Delphi process or the Delphi method) is a process where subject-matter experts iteratively deliberate until they agree upon a solution. In each iteration, organized feedback is given to the participants based on the previous deliberation results. The deliberation should not be focused on confrontation, where participants try to convince others of their viewpoint, but rather on finding the best common resolution. Moreover, differences in opinion should be explicitly discussed. The deliberative Delphi process can span over an extended period of time [64].

The proposal also included plans for deriving quality weights by using pairwise comparison questions, TTO questions, or alternatively visual analogue scales (VAS). Furthermore, the proposal envisioned a trade-off between efficiency and equity and rejected pure maximization, inspired by Alan Williams' fair innings [65]. Moreover, the proposal suggested having an upper ceiling and focusing on the sensitivity of the measure at the lower end of the scale, because the focus of societal policy should be to help those below this threshold to achieve capability for a good life rather than enabling some individuals to have a perfect life.

Box 1.7.3: Equity efficiency trade-off

When allocating resources subject to maximizing, for example, health or capability, situations may arise where the most efficient allocation does not correspond to the most equal allocation. For example, it might be more efficient to target higher educated people with health literacy campaigns as for the same cost a better uptake could be achieved compared to targeting people with less education. Because higher educated people are tend to have better health, this might increase health inequity.

Another area where CALY-SWE could add value and facilitate evaluation – for example, for health impact assessment – is intersectoral public health work, where the challenge is to coordinate across different policy sectors to promote and improve population health beyond what is achievable in single sectors [66,67].

This proposal formed the point of departure for my thesis. While some of the initial parameters of CALY-SWE have been outlined, much work lied ahead before the concept could be operationalized in an economic evaluation.

1.8 CALY-SWE as an alternative to describe societal welfare

Beyond using CALY-SWE as an outcome for economic evaluations, the proposal also envisioned measuring the distribution of capability in the Swedish population. This vision aligned well with an international push towards more diverse measures of societal welfare than the predominant, financially oriented monetary measures like the gross domestic product (GDP) and was the topic of a report by Stiglitz, Sen and Fitoussi [68]. In fact, the Swedish government's investigation, which was mentioned in the CALY-SWE proposal as the starting point for selecting capabilities, was part of the Swedish effort towards such a diversification in welfare measures [63].

The resulting distribution of capability would likely differ between population groups, meaning that CALY-SWE could be used to assess capability inequalities. Of relevance in this regard there are the three concepts of *equalitarianism*, *prioritarianism*, and *sufficientarianism*. These concepts stand for different ways to view inequal distributions of welfare, which is the topic of the field of distributive justice [69]. Egalitarianism postulates that to achieve an equal distribution, it is permitted to decrease the welfare of those who are better off. This "levelling-down" [69] is not particularly attractive when considering that taking away all welfare from all

individuals would result in a just distribution, an issue that prioritarianism and sufficientarianism try to address. Prioritarianism holds that the welfare of those worst-off should be prioritized, while sufficientarianism holds that any welfare improvement for those below a certain threshold should be prioritized before improvements above the threshold. These three concepts can be contrasted with the maximization principle, which does not have any equality in distribution as a restraint. For sufficientarianism, however, prioritizing the maximal improvement below the threshold may be a compatible strategy. For all, maximization can still be pursued if the available alternatives would not change the relative distribution.

In summary, using CALY-SWE to measure the distribution of capability within the frame of the thesis is valuable to collect evidence on the level of Swedish societal welfare and to assess capability inequalities in the Swedish population. Demonstrating the usability of CALY-SWE in this fashion would also help to pave the way to use it for its main purpose: to conduct cost-effectiveness evaluations of interventions with broad social welfare outcomes.

1.9 Applying CALY-SWE in an economic cost-effectiveness evaluation

To apply CALY-SWE in an economic evaluation, we needed to find an appropriate intervention to evaluate. I decided to use secondary data to perform an economic evaluation, coupled with capability-related QoL data collected for S3. One such opportunity arose through the Swedish payroll tax cut from 2007 to 2016 during the great recession for employees aged $\leq$ 26 years, which attracted a substantial amount of research that relied on Swedish population registry data [70] to analyse the policy. The payroll tax was reduced from 31.4% to minimally 15.5%, with the goal of increasing occupation levels of young workers [70–72].

The concept of youth unemployment has shortcomings. In particular, it relies on individuals being in the labour force. In addition, it includes students who look for work, a situation in which the unemployment status may be less problematic compared to those not finding jobs for longer periods. To address these issues, the concept of young people *not being in employment, education or training* (NEET) has been increasingly used since the turn of the millennium [73,74]. The Swedish Agency for Youth and Civil Society (MUCF) [75] publishes registry-based statistics on NEET prevalence. It defines NEET as not having earned one prise base amount (roughly 50,000 SEK in 2023 [76]), not receiving study support, and not being registered as a student for at least 1 year. The prevalence in 2021 was 5.7% among young people aged 16-24 years and 11.2% among young people aged 25-29 years

[77].

When using MUCF's NEET definition [77], one can conceptualize lengthy unemployment as a precursor to being NEET. The effects on youth unemployment due to the payroll tax reduction, as a hypothesis, would thus logically also affect NEET prevalence. This view is partially supported by existing evidence of risk factors for NEET, where economic circumstances such as job shortage have also been assessed to be associated with NEET occurrence [78,79], among a multitude of other factors, including socioeconomic circumstances related to the family and education [80].

Regarding consequences, having been NEET earlier in life has been debated to result in scarring effects later in life [81]. Specifically, there is evidence for lower job satisfaction [82], increased risk for economic inactivity [83], social exclusion [84], mental health problems [85,86], or prolonged negative effects on income [87].

CALY-SWE should be well suited to evaluate the consequences of the payroll tax cut because the QoL consequences are broad and potentially accumulate over prolonged periods over the life course. Most of the existing research seems to have been conducted from an economic and labour market perspective [70] Hence, besides proving its applicability, using CALY-SWE for evaluation may help to improve the evidence on the consequences of the payroll tax cut.. However, to do so, monetary cost and benefits as well as CALY-SWE QoL consequences must be quantified.

1.10 Existing capability measures for economic evaluations

Besides CALY-SWE, several instruments that ideologically ground themselves at least in part on the capability approach, and whose intended use case is in economic evaluations, have been developed (Table 1.1). However, their characteristics differ from those specified by the CALY-SWE proposal.

The most established capability questionnaire for use in economic evalution is arguably the English ICEpop CAPability (ICECAP) family of measures targeted at economic evaluation in the general population (ICECAP-A) [88]; elderly care (ICECAP-O) [91]; end-of-life care (ICECAP-SCM [95] and ICECAP-CPM [89]); and, recently, a version for youth, children, and their parents (ICECAP-CYP) [94]. Another stream of instruments, also developed in the UK, incorporates the OXford CAPability (OXCAP) [96], OXCAP-18 [97], and OXCAP-MP [98]

	ICECAP-A, ICECAP-O, -SCM, -CPM, CYP	OxCAP, OCAP-18, OxCAP-MH	ASCOT	Questionnaire by Anand and van Hees 2006	CALY-SWE
Purpose	Economic evaluation in the general population	•General population •(OxCAP-MH: Mental illness)	Adult social care	General population capability measurement	•Economic evaluation •Tracking living conditions over time, between areas
References	[88-95]	[96-98]	[99]	[100,101]	[1,2,58]

Table 1.1: Overview of some prominent capability outcome measures. Adapted from S1. ICEpop CAPability (ICECAP) – measure for Adults (A) – measure for Older people (O) – Supportive Care Measure (SCM) – Close Person Measure (CPM) – Children and young people (CYP). Oxford capability (OxCAP). Capabilty-adjusted life years Sweden (CALY-SWE).

measures and features attributes that are more closely related to Nussbaum's list of central capabilities. One of the involved authors in OXCAP, Anand [98], had previously also proposed a questionnaire with seven capability dimensions [100]. Additionally, there is a UK-based initiative building on the capability approach, the Adult Outcome for Social Care Toolkit (ASCOT), that focuses on social care. The presented list is not comprehensive, and other efforts to develop capability instruments have been undertaken and are ongoing [102,103].

The need for broader outcome measures for economic cost-effectiveness evaluations of social care, healthcare, and public health has also been recognized in a non-explicit capability context. A recent well-supported endeavour is the development of the EQ Health and Wellbeing (EQ-HWB) instrument, which was developed in the UK and part of the EuroQol family of measures. In addition to health-related attributes, it includes more general wellbeing items such a question on feeling lonely or feeling good about oneself. In total, there are 25 items in the full version and 9 in the short version [104].

Sen [55,57] emphasizes the context-specificity of the capability approach, but vis-à-vis the CALY-SWE proposal, none of the other capability measures were designed specifically for the Swedish context. A Swedish attribute selection process would still need to be conducted even if an identical attribute list as in another measure would result, just for confirmation. Furthermore, the suggested approach of selecting capabilities by using a fair-minded Delphi panel differs from the approach of participatory research that was, for example, used to develop the ICECAP-A instrument [88]. Moreover, the suggested development of the value set using pairwise comparison was distinctively inspired by health economic

methodology [58].

1.11 Thesis aim, goals, and objectives

The overall aim of this thesis was to develop, value, and apply a broad QoL questionnaire and a corresponding value set, built on the capability approach, that is tailored to the Swedish social welfare context, as outlined in the CALY-SWE proposal [58]. This gives rise to the following goals and objectives for my thesis:
Study 1: *Select* the CALY-SWE questionnaire attributes and develop the phrasing of statements and answer levels

- Conduct a Delphi panel with fair-minded stakeholders from civil society to select relevant capabilities for the Swedish context
- Incorporate inequality considerations in the questionnaire

Study 2: *Elicit* a value set of quality weights for the CALY-SWE questionnaire to enable quality weighting of lifetime and subsequent use in cost-effectiveness evaluations in social welfare

- Develop appropriate methods, considering the nature of the questionnaire and practical constraints
- Conduct a valuation study to elicit weights

Study 3: *Apply* the CALY-SWE questionnaire to *measure the capability distribution* in a Swedish population sample

- Describe the distribution in terms of absolute capability attainment
- Characterize the distribution in terms of relative differences, inequalities linked relevance of contributing factors

Study 4: *Apply* the CALY-SWE questionnaire to assess the *cost-effectiveness* of the payroll tax reduction in 2007 to 2016 for ages below 27 to avert individuals becoming NEET

- Model and establish the association of the impact of the tax reduction on NEET occurrence
- Estimate the impact of having been NEET on capability-related QoL, and cost consequences connected to the tax cut, to calculate costs per CALY

1.11.1 Outline of the thesis

After this general introduction, where I covered relevant concepts and literature, and presented and motivated the research aims, there follow four chapters, one for each research goal and study. Each of these for study-specific chapters contains a methods and results sections. In the final chapter, I synthesize and discuss the findings. After the bibliography chapter, I finish with an Epilogue of some more personal reflections. The thesis also contains appendices to provide additional context. Appendix 1 contains additional results, Appendix 2 a timeline of relevant events in the CALY-SWE development, Appendix 3 my contributions to the individual studies, and Appendix 4 additional resources related to the research project. Lists of Figures, Tables, and Infoboxes follow, before an Acknowledgements section. The final section is comprised of the traditional list of all previous PhD theses at the Umeå University Department of Epidemiology and Global Health.

Chapter

2 Capability of what: selecting capability attributes and phrasing the statements (S1)

2.1 Overview

S1 concerned the development of the questionnaire as outlined in the CALY-SWE proposal [62] and contained two main parts:

(1) Select a list of capabilities to be included as attributes. The proposal suggested five as a reasonable number of attributes, because keeping the questionnaire short eases practical usability and limits cognitive demands in valuation tasks [62,105].
(2) Adopt phrasing for the selected capability attributes that implies a ceiling above which capability should not be measured. We interpret this as a combination of both prioritarianism and sufficientarianism [69]. The goal is to increase equity in society by prioritizing those under the sufficiency threshold.

2.2 Methods

2.2.1 Selection of capabilities using a Delphi panel

The initial consideration for the selection of capabilities was policy relevance. For CALY-SWE to be useful in policymaking, it should include areas that are policy relevant and show sufficient variation between individuals in present-day Sweden,

so individual differences are relevant. The questionnaire should also be easy to administer, implying that there is a limited number of attributes and items.

We conducted a Delphi panel to select the relevant capabilities to be included in the CALY-SWE instrument. The overarching aim of this Delphi panel was to implement a version of the *fair-minded people* concept, which was introduced by Norman Daniels [106]. The intention was to arrive at a consensus between relevant Swedish stakeholders regarding a list of capabilities to be included.

Box 2.2.1: Fair-minded

Fair-minded people want follow accepted rules and prefer rules that further the common cause [106]. In other words, we aimed for the Delphi process members to also have the best interest of society at large in mind, and not particular interests of limited population subgroups or even self-interest.

We invited a man and a woman of several Swedish non-profit organizations to participate. Our intention was to ensure that the participants would genuinely represent the interests of their interest group as well as the entire Swedish population, while keeping gender balance in mind. For this reason, we also refrained from inviting political parties to the process as we feared that it might be difficult for political representants to adopt this overarching perspective while genuinely collaborating across political boundaries. The selection of organizations was also guided by the Sustainable Development Goals (SDGs) [107]. However, not all SDGs are equally relevant at the individual level. Some of the goals, such as combating climate change, are still important at an individual level, but have greater collective importance. We selected organizations so that the more individual SDGs were covered by at least one organization, namely SDG 1 (no poverty), SDG 2 (no hunger), SDG 3 (good health and wellbeing), SDG 4 (quality education), SDG 8 (decent working conditions and economic growth), and SDG 16 (peace, justice, and stronger institutions). Table 2.1 shows a list of the invited organizations.

2.2.2 Delphi panel procedure

We organized a start-up day in Stockholm in January 2020 and used an online survey for the subsequent deliberation rounds. We presented the Delphi panel with the list of 10 basic capabilities suggested in the Swedish governmental investigation [63]: financial resources; mental & physical health; political resources & civil rights; knowledge & skills; occupation; social relations; security of life and property; housing; balance of time; and living environment. We also asked the participants to agree to act fair-minded.

Figure 2.1: Lars Lindholm at the Delphi panel day in Stockholm, 23. of January, in Stockholm.

After the start-up day, we asked the Delphi panel to rank the capabilities by importance in a web survey. In addition, the Delphi panel members could leave comments regarding whether all relevant capabilities had been included, on the number of capabilities in the final instrument, and regarding the motivation for the ranking. Afterwards, we employed web surveys for subsequent rounds, based on a synthesis of the statements collected in the previous rounds.

2.2.3 Empirical capability distribution to inform statement phrasing

To inform where and how to set the ceiling in the phrasing, we measured the distribution of capabilities to be included as attributes in the questionnaire in a Swedish population sample. We created three different versions of wording for each of the selected capabilities. The A version used the wording *always*, the B version used the wording *almost always*, and the C version used the wording *mostly*. If necessary, the wording could be changed to other similar frequency statements for grammatical reasons. In addition, for the B and C versions, we added exemplifying clauses that specified the frequency in more detail.

If the differences in wording were to influence the resulting answer proportions, then the final questionnaire phrasing could be informed by those differences so that

we could choose an appropriate wording that would result in a sensible answer distribution. Specifically, with a decrease in strictness from A over B to C, we expected the share of *Agree completely* answers to rise, and the shares of *Agree partially* and *Not at all* answers to decrease. The statements for the security attribute follow as an example:

- ”Risk for violence or other crime *never* affects me during my work or my leisure time.”
- ”Risk for violence or other crime *almost never* affects me during my work or my leisure time. I have changed my plans an estimated 10 or fever times due to unsafety (for example chosen another route for walking) during the last year.”
- ”Risk for violence or other crime *rarely* affects me during my work or my leisure time. I have changed my plans an estimated 25 or fever times due to unsafety (for example chosen another route for walking) during the last year.”

We also wanted to avoid too low or too high of a proportion for an answer level to maintain the sensitivity of the measure to changes in capability. If, for example, the proportion of *Completely agree* were to reach 1 for an attribute, then this phrasing could not measure improvements in a population sample.

We decided on a sample size of 500 for each version considering that a two-sample t-test for proportions has 80% power to detect a 0.1 difference in proportions at a 5% significance level. We also used a t-test with those specifications to assess the differences between versions A, B, and C.

In addition, we included three questions on inequality aversion in the survey for capability attributes where we managed to develop a plausible question that could be used in an online survey, namely for life expectancy as a proxy for the health attribute, and income as a proxy for financial situation & housing and education. We included education even though it did not conform closely to one of the selected attributes: we intended for it to serve as additional data point on the feasibility of the concept and to be able to contrast the results with the other two questions. We included identical forms of those questions in the three versions, presenting the participants with a statement on the distribution, for example, for income: “In Sweden the average salary for full-time workers is 34,600 and 10% earn more than 49,200. What is your opinion about the salary difference?” The participants could state if they found this situation acceptable, or if the presented inequality was too large. See the supplementary information to S1 [1] for the complete questionnaire, including all statements for the A, B, and C versions and inequality aversion questions.

Organization name	SDG(s)	Start-up day	Delphi process
Amnesty	16	No	-
Children's Rights in Society	1, 5	Yes	Yes
Crime Victim Support Sweden	16	Yes	Yes
The National Association for Cancer Patients	1, 3	No	-
Disability Human Rights	1, 3, 8, 16	Yes	Yes
The Swedish Network of Refugee Support Groups	1, 16	No	-
Islamic Association in Sweden	1, 16	No	-
National Organization for Pensioners	1, 3, 16	No	-
Swedish Pensioners' Association	1, 3, 16	No	-
Comrade Association of Former Criminals	4, 16	No	-
Fryshuset Global (supporting young people)	16	No	-
RFSL	3, 16	Yes	Yes
Red Cross	1, 2	Yes	Yes
Save the Children	1, 2, 4	No	-
The Swedish Trade Union Confederation	8	Yes	Yes
The National Organization for White Collar Workers	8	Yes	No
The Swedish Confederation of Professional Associations	8	Yes	Yes
Confederation of Swedish Enterprise	8	No	-
Salvation Army	1, 2	Yes	Yes
The Swedish Council for Information on Alcohol and Other Drugs	3, 16	No	-

Table 2.1: Delphi panel participants. Columns one to four contain the organization name, the SDG(s) related to the organization, if the organization representants participated in the start-up day, and if the representatives participated in the Delphi process, in order. Adapted from S1. SDG numbers: 1. No poverty, 3. Good health and wellbeing, 4. Quality education, 5. Gender equality. 8. Decent work and economic growth, 16. Peace, justice, and stronger institutions. Sustainable Development Goals (SDG). The Swedish Federation for Lesbian, Gay, Bisexual, Transgender, Queer, and Intersex Rights (RFSL).

2.3 Results

2.3.1 Delphi panel

Twenty-two organizations were invited, and nine participated in the start-up day. Eighteen survey invitations to the first-round survey were sent out (to the two representatives for each organization), and 16 representatives answered (Table 2.1). The rankings indicated strong agreement for including the capabilities of health, social relations, financial resources, and housing. Knowledge & skills, time balance, and living environment ranked lowest; thus, we removed them from the list of potential candidates. There was disagreement for occupation, security of life and property, and political resources & civil rights.

In the second round, we presented the representatives with a synthesis of the first-round rankings and asked them to rank occupation, security of life and property, and political resources & civil rights again. In addition, we proposed merging the capabilities of housing with financial resources, with the argument

that they are related to a large degree.

In this round, 12 of the 18 representatives answered. The ranking remained largely unchanged compared with the first round, but there was broad support for the merger of financial resources and housing. However, the ranking still did not provide a clear indication regarding whether occupation, security of life and property, and political resources & civil rights should be included as the fourth and fifth attributes. We did not conduct a third round because we assessed that a consensus was unlikely to be achieved. Instead, we decided to include six capability attributes instead of the five initially envisioned so that the three remaining attributes could be included.

For the final questionnaire, we made slight changes to the wording of the attribute headings to make them more descriptive and concise. The final list of included attributes was: health, financial situation & housing, social relations, occupation, security, and political & civil rights.

2.3.2 Empirical capability distribution to inform statement phrasing

In the population sample web survey (version A, n = 497; version B, n = 503; version C, n = 505), for all attributes except political & civil rights, the highest answer proportions occurred for *Completely agree*; it ranged from 37% to 67%. The proportions for *Not at all* ranged from 4% to 15% and was of a similar low magnitude for all attributes, except for social relations, which showed a slightly lower proportion, and with small differences between the versions. There were marked differences in answer proportions for *Completely agree* between versions A and C for health. *Partially agree* and *Completely agree* showed both differences between versions B and C for financial situation & housing; between versions A and C and between versions B and C for occupation; and between versions A and B and between versions A and C for security. For *Partially agree*, marked differences occurred between versions A and B for political & civil rights (2.2)).

The inequality aversion question indicated markedly higher inequality aversion for health and education compared with income (Figure 2.3).

2.3.3 Phrasing the statements

Next, we developed the phrasing for the capability attributes based on the results from the quantitative study on the impact of different wordings; the Delphi panel; and other arguments of which we were aware, such as contextual knowledge of Sweden, Swedish laws and regulations, and normative and ethical aspects. We also changed the wording for the lowest answer level from *Not at all* to *Do not agree*

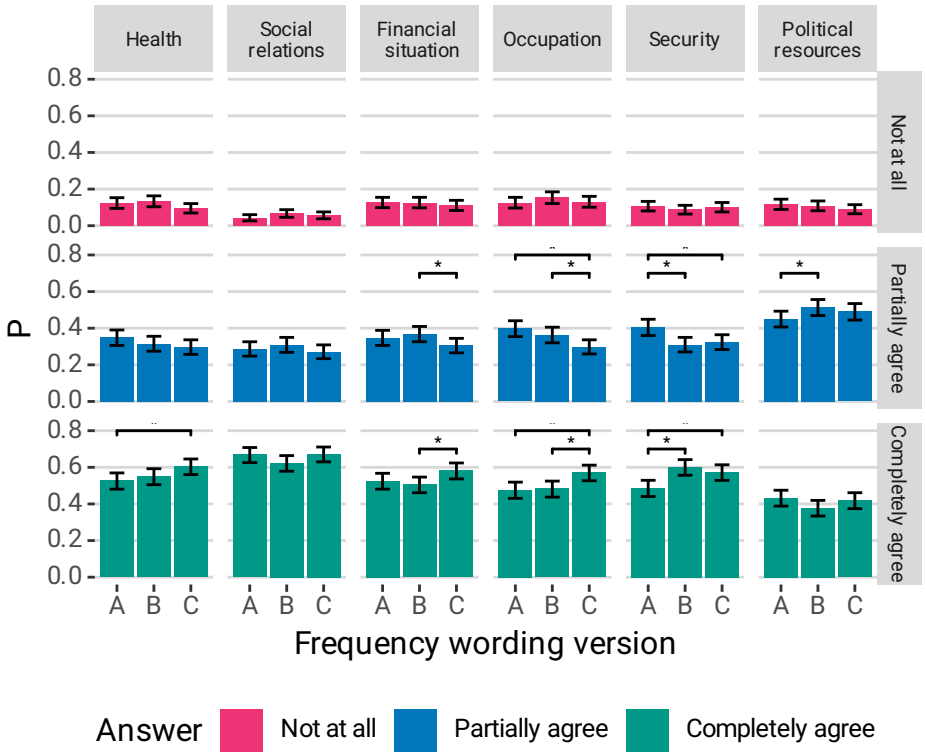

Figure 2.2: Distribution of attribute-level answers per frequency for each version (A, B, or C). *Indicates a significant difference in the proportions between the versions, using a t-test for proportions with a significance level of 0.05.

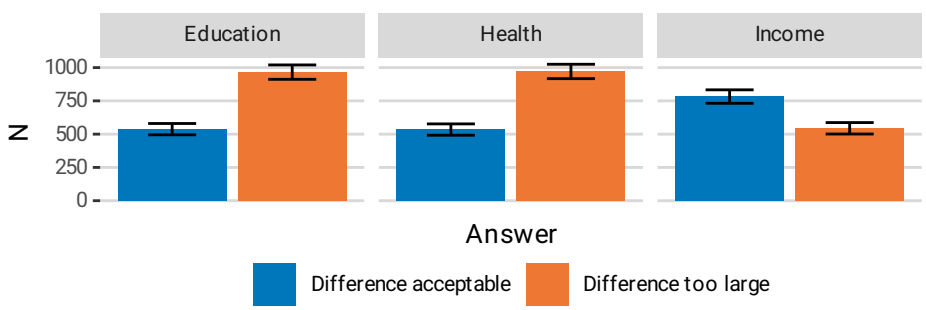

Figure 2.3: Inequality aversion answers.

Informal english	Swedish
Health: In general, I have good health (mental and physical) that almost always (more than 95% of days) allows me to work or do what I want.	**Hälsa**: Jag har ett bra allmänt hälsotillstånd (psykiskt och fysiskt) som nästan alltid (säg minst 95% av alla dagar) tillåter mig att arbeta eller ägna mig åt det jag vill.
Social relations: I have access to close relations (family, friends, or acquaintances) who give me advice and support when I need.	**Nära relationer**: Jag har tillgång till nära relationer (familj, vänner eller bekanta) som ger mig råd och stöd när jag behöver.
Financial situation & housing: I have a financial situation (salary, other incomes, savings) that always allows me to have a permanent housing, and mostly (at least 8 out of 10 times) buy what I think I need.	**Ekonomi & boende**: Jag har en ekonomi (lön, annan inkomst eller besparingar) som alltid tillåter mig att ha en fast bostad och för det mesta (minst 8 gånger av 10) tillåter mig att köpa det jag tycker mig behöva.
Occupation: I have employment or another occupation (study, internship, homework, care of relative) that I am mostly satisfied with. During the last 5 years, I have been satisfied with my occupation at least 75% of the time.	**Sysselsättning**: Jag har ett arbete eller annan sysselsättning (studier, praktik, hemarbete, vård av anhörig, etc.) som jag för det mesta är nöjd med. De senaste fem åren har jag minst 75% av tiden varit nöjd med det jag ägnat mig åt.
Security: I have a sense of security, and the risk of violence or other crime does not influence me in my work or leisure activities.	**Säkerhet**: Jag har en trygghet som gör att risk för våld eller annan kriminalitet inte påverkar mig i mitt arbete eller i min fritid.
Political & civil rights: I trust that our rights as citizens, and our active participation in the democratic processes make it possible to shape our common living conditions.	**Politiska & medborgerliga rättligheter**: Jag litar på att våra rättigheter som medborgare, och vårt aktiva deltagande i det demokratiska systemet, gör det möjligt att påverka våra gemensamma livsvillkor
Level answers	
Do not agree	Instämmer inte
Partially agree	Instämmer delvis
Completely agree	Instämmer helt

Table 2.2: Swedish phrasing and informal English translation. Adapted from the supplementary information to S2.

to make it more similar to the other levels. Table 2.2 provides the final adopted phrasings.

2.3.3.1 Health

We decided to use the phrasing "almost always". While the Delphi process underlined the importance of heath, based on a high inequality aversion, and Swedish policy values good and equitable health highly [108,109], it may be unreasonable to "always" have full capability as this would be too demanding on resources in present-day Sweden, and it might compete with the realization of other capabilities.

2.3.3.2 Financial situation & housing

We decided to use "always" as a modifier for housing and "mostly" for the financial situation. The reason for a lower threshold for the financial situation lies in the less marked inequality aversion and the already high general economic standard in Sweden relative to the global and even European scales. In combination, the proportions for *Agree* and *Agree partially* differed between the B and C version, evidencing an impact of the wording differences. On the other hand, we chose "always" for housing because we understand access to proper housing to be a more fundamentally accepted right in the Swedish context.

2.3.3.3 Occupation

We decided to use the "mostly" variant of the phrasing. Gainful employment or other economically productive activities can, to some extent, be considered a duty to solidarily contribute to the resources of the welfare state, and thus motivate setting a lower threshold. Similarly, the biggest concern in this area is unemployment, and furnishing basic occupation capability for unemployment should have priority before providing a high employment capability for all. In addition, for *Partially agree* and *Completely agree*, the wording differences between versions A and C and between versions B and C seemed to be meaningful, as indicated by the t-test.

2.3.3.4 Security, social relations, and political & civil rights

We discarded the version wording modifiers because of normative arguments, even though the empirical data indicated some differences between the versions for security and political & civil rights. For security, tolerating degrees of unsafety seems to be immoral, in line with Article 3 of the Universal Declaration of Human Rights [5], Chapters 1 and 2 of the Swedish instrument of government [110], which proclaim the equality and bodily integrity of individuals.

Security and political & civil rights are not purely individual; rather, they have shared characteristics of a *public good* [111], where individuals cannot be excluded from consuming the good, and where the good is not reduced by individual consumption. The provision of those by an institution should be unconstrained and available for all. Social relations have similar characteristics as a public good, although in a more reduced scope that, for example, includes family and friends. Thus, for these three capability attributes, we did not include frequency modifier words.

2.3.4 Meaning of the highest state (333333)

The conceptualization of the highest CALY-SWE state builds on two normative concepts: *flourishing* and *sufficiency*.

Flourishing concerns the normative description of the state of a "good" life. For this description, we followed Martha Nussbaum and conceptualized the CALY-SWE state of 333333 to be sufficient to lead a flourishing life [56]. Human flourishing, as a comprehensive view of wellbeing, crucially encompasses more than just the current state of physical and mental wellbeing, for example, virtue. It can be understood in longitudinal dimensions, where the entire can be assessed regarding its degree of flourishing [112].

The second concept concerns the in-build equity consideration, namely sufficiency for flourishing. CALY-SWE does not aim to measure the full extent of capability; instead, it focuses on the lower end of the scale until the sufficiency threshold that is operationalized in the frequency clause of the questionnaire phrasing – for example, for health: "... that almost always allows me to (more than 95% of days)...". This approach is akin to the distributive justice principle of sufficientarianism. Applied to the capability context, sufficientarianism postulates that the priority is to ensure a sufficient capability for all. Sufficient capability is a normative threshold above which the capability level is judged to be sufficient. Such a threshold can be established – for example – by means of a democratic process. Improvements below the sufficiency threshold should be prioritized before attempting improvements that go beyond the threshold or that lie above the threshold. [69].

Chapter

3 For what it is worth: eliciting capability weights (S2)

3.1 Overview

S2 was concerned with developing a value set of CALY-SWE quality weights situated on the scale of 0 to 1. This chapter describes how we developed the survey and then collected and analysed the data to produce the value set.

3.2 Methods

For CALY-SWE, we decided to rely on the methodological experience gained in the health-economics QALY context, because cost-effectiveness evaluations using QALYs as an outcome are widely established. We opted to derive the methods for eliciting the value set in large part from the EQ-5D elicitation methodology and to use TTO and binary DCE tasks. However, in contrast to EQ-VT, we did not consider the possibility of negative TTO valuations. We focused on eliciting an initial value set for CALY-SWE given that this study was one of the first to value a capability measure by using TTO. In addition, allowing negative weights would have increased the complexity of the survey valuation task. Given the similar use cases in economic evaluations and the conceptual origin of CALY-SWE in health economics, building on health economics methodology seemed to be appropriate. Yet, other capability instruments have taken a different route. For example, the ICECAP family of measures relies on best-worst scaling [91,92], which goes back

to the random utility school.

CALY-SWE has six attributes with three levels each, resulting in 3^6 = 279 different situations that the instrument describes. One such situation is called a configuration or capability state. We denote a capability state as xxxxxx, with x ranging from 1 to 3 and corresponding to the level of the attribute. The lowest level of capability (*Do not agree*) is encoded as 1, the second level (*Partially agree*) is encoded as 2, and the highest level (*Completely agree*) is encoded as 3. For example, 333111 stands for level 3 for health, social relations, and financial situation & housing, and level 1 for occupation, security, and political & civil rights.

3.2.1 Planning and timeline

The overall planning involved the following partially independent steps.

1. Determining the cornerstones
    a) Methodology
    b) Survey mode
2. Web survey development
    a) Qualitative interviews
    b) Staged roll-out
3. Experimental design development
4. Sample size simulation
5. Data collection
6. Data quality assessment
7. Value set modelling

The schedule spread over a prolonged period with the development of the web survey starting in summer 2020 and manuscript writing until summer 2023 (Table 3.1).

3.2.2 Determining the cornerstones

As described above, we decided to rely on DCE and TTO questions. In the EQ-5D context, the data for TTO and DCE questions is normally collected in face-to-face interviews during which individual interviewers guide participants through the questions. Recent studies have also conducted interviews virtually [114,115], but studies relying on unsupervised surveys without interview guidance are rare. In contrast, for CALY-SWE we decided to rely on unsupervised online surveys.

Activity	Approximate period
Methodology decision for DCE and TTO questions	2018
Decision for online web survey mode	Summer 2020
Web survey development	Summer 2020 to January 2022
Experimental design development	Fall 2020 to Spring 2021
Qualitative interviews	Fall 2021
Sample size simulation	Fall to winter 2021
Data collection staged roll-out	Spring 2022
Data quality assessment	Summer to Fall 2022
Value set modelling	Fall 2022 to Spring 2023
Manuscript writing and publication	Fall 2022 to Summer 2023

Table 3.1: Timeline value set study. Time trade-off (TTO). Discrete choice experiment (DCE).

Given the project's resources and challenges related to the COVID-19 pandemic, we deemed face-to-face interviews to be a less appropriate choice.

An important focus of the value set study was to ensure data quality. Online unsupervised data collection comes with its own challenges regarding data quality that are especially relevant for TTO questions. Because there is no direct contact with the participants and the amount of data collected per participant is limited, it is more challenging to determine whether the participants understood the survey in accordance with the exceptions and assumptions that underlie the original design. The participants may also speed through the survey with low engagement. TTO questions are complex, and there are known challenges with data quality [28,116].

To address these challenges as thoroughly as feasible, we adopted an integrated, holistic development approach [113] to produce a user-friendly web survey with a focus on collecting adequate TTO data, subsequent data quality assessment, and statistical analysis. As part of this approach, the experimental design and larger sample size should compensate for the lower amount of data collected per participant compared with traditional, physical face-to-face interviews. In the analysis stage, we aimed to consider data quality and to select an appropriate model to generate the final value set. An overview of the study process is provided in Figure 3.1 (See also the timeline in Table 3.1).

3.2.3 Technical infrastructure for the survey

At the onset of the study, externally available survey software seemed unsuited to properly accommodate the complex TTO questions. Starting with pilot versions [117,118], we developed a custom survey using a frontend based on HTML, CSS, and JavaScript; a PHP [119] backend with Twigg templating [120]; and MariaDB [120] and SQLite as storage solutions [120], hosted on Umeå University servers. We

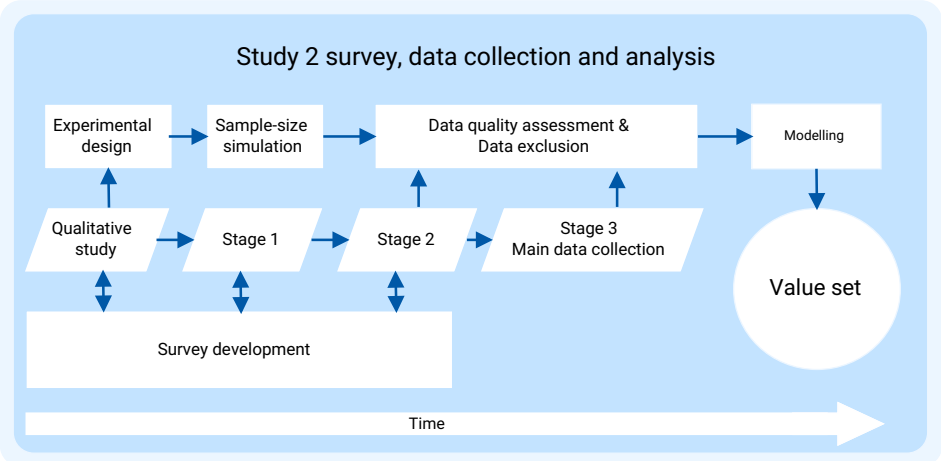

Figure 3.1: Overview of the weight study. Adapted from [113].

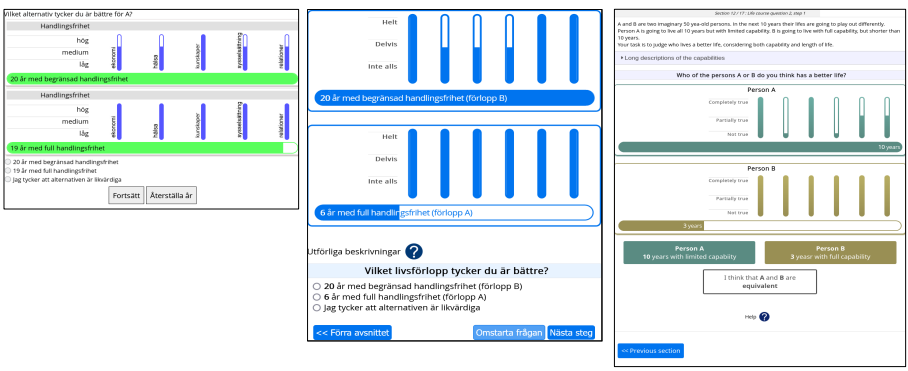

Figure 3.2: Left: survey for the qualitative interviews. Middle: survey for stage 1. Right: survey for stages 2 and 3.

adopted a responsive design mobile-first development strategywhere the website was first developed for mobile screens before larger devices [121].

3.2.4 Survey evolution, changes, and improvements in multiple stages

Before web panel sampling, we conducted a qualitative study with 16 participants. Quantitative researchers (Jan Hjelte and Kerstin Edin) asked the participants questions on ease of use, question ordering, impressions and more. Based on the feedback, we fixed bugs and improved the visual design and the instructions in the survey (Figure 3.2).

Afterwards, we started testing with panel-sampled participants in three stages. The initial stage with 111 participants resulted in limited data quality. We made changes to the survey: we reduced the time horizon in the TTO question from 20 to 10 years, tuned the iteration procedure so that choosing equivalence of the two states was no longer directly possible at the beginning of the iteration procedure, and introduced colour coding for the two choices. The goal was to reduce the average number of iterations per TTO answer to counteract participant fatigue while still ensuring a minimal number of iterations to reduce incentives for answering the TTO question too hastily.

We also introduced a *learning state* for the TTO questions that was displayed as the first TTO state for each participant. In this learning state, the participant was presented with a dominated choice with two alternatives: 9 years in a state with 3 attributes on level 2 and 3 attributes on level 3, and 10 years in the full capability state 333333 with all attributes on level 3 (For example 9 years in 333222 vs 10 years in 333333). The first two iterations of the learning state were explained to the participant to guide and familiarize them with the task, but the following iterations proceeded unguided.

We collected again data from 199 participants in the second stage and used a bootstrapping approach to test if data quality had improved. Specifically, we used a linear regression model for the TTO data and a logit regression model for the DCE data and randomly resampled with repetition the same number of observations from each stage and assessed the model fits.

After testing the updated survey version in stage 2 with 199 participants, we found that the TTO data quality had improved while the DCE data quality had remained largely unchanged. We launched the final data collection in the third stage (Figure 3.1) and pooled stage 2 and 3 data for analysis.

The final survey contained the following blocks, in order: informed consent, the CALY-SWE questionnaire, a visual analogue scale task, six DCE tasks, and five TTO tasks. More details on the integrated survey development are available in Meili et al. [113].

3.2.5 TTO iteration procedure

The final version of the iteration procedure used a bisection procedure with a starting interval of 1 to 10. We chose a random value between the upper and lower limit of the current interval for each iteration, depending on the choice of the participant. The aim of randomization was to avoid directional bias by randomizing the bisection point.

It worked roughly as follows (details available in [113]): choosing the impaired state with the maximum time associated would result in a new value for x in the next iteration that was between x and the current upper interval limit. Choosing the full capability with tradeable time x would result in a new value for x in the next iteration that was located above the current lower interval limit and below the current value of x. The bisection interval was updated accordingly, with x becoming the new lower or new upper interval limit, respectively, thus decreasing the width of the interval with each iteration.

In other words, the full capability state with tradeable time was made more or made less attractive to iteratively approach the point of indifference. The participant could also state that both choices were equivalent instead of choosing either. This choice would reduce the width of the current interval by half by shrinking it around the current value of tradeable time, and choose a new random value of x. If the current interval reached a width of 1 in an iteration, then the participant was instead prompted to confirm that they agree that for this number of years, the choices were equivalent. The lowest tradeable units was 1 year.

The wording used in the TTO question did not explicitly mention death because we conceptualized the meaning of 0 as 0 lifetime, in contrast to QALYs, where 0 is commonly anchored to death [7].

3.2.6 Experimental design and coefficient coding

For the value set study, the experimental design should ensure that the value set could be estimated without bias and with maximum precision for a given sample size.

Box 3.2.1: Experimental design

Experimental design is concerned with optimizing the parameters of an experiment so that the maximum amount of information and precision can be obtained. For a survey, this means optimizing which questions get asked [122].

To generate the experimental design and the sample size simulation, we decided to rely on a basic additive main effects model, where a coefficient encodes the value of one attribute level. To elicit the value set, it is necessary to generalize the statistical information obtained from DCE and TTO questions by using a statistical model so that a quality weight for each possible state can be calculated. TTO data are naturally suited for analysis with linear regression because the outcome is on a

continuous scale, while DCE data are suitable for analysis using logistic regression due to the binary nature of the outcome.

For TTO, we used the skipr R package [123] that tries to find an optimal design with an algorithm that maximizes the amount of information by probing different configurations. For DCE, we employed the orthogonal array approach that is outlined by Street et al. [124]. In this approach, linear algebra properties of a special form of matrix columns, the orthogonal arrays, are exploited to develop a design that maximizes the amount of information in the experiment.

3.2.7 Sample size simulation

We used the mean time per question type and position from the qualitative study to inform the expected required time. Waimed for each participant to spend a maximum of 20 minutes on the survey to maintain engagement and to keep participant recruitment feasible. We calculated different configurations of numbers of DCE and TTO questions that would fit, together with the remaining questions, within the 20-minute limit.

For TTO, we developed different designs where the number of TTO questions per participant corresponded to the different configurations of the DCE and TTO questions. For DCE, we used random blocking, meaning that we developed just one experimental design with 43 different choice tasks that we randomly assigned to participants with different block sizes depending on the configuration.

Box 3.2.2: Credible interval

In a Bayesian context, the equivalent to a *confidence interval* (CI) is often termed a *credible interval* (CRI) to emphasize conceptual differences compared with the frequentist CI. In short, CRIs indicate the range for the true estimate for a given probability and the given prior information. On the other hand, CIs indicate the range that would be narrow enough to contain the true population parameter with a given frequency, often 95%, if the same analysis including sampling was repeated many times [125,126].

For each configuration, we then simulated a basic additive effect hybrid model and a linear regression TTO-only model for different numbers of participants, ranging from 100 to 2,000. We used generative parameters for weights and uncertainty based on earlier pilot results [118], and we assessed model performance based on the 95% CRI width and the mean absolute error of the weight posteriors compared to the simulation parameters.

Based on the sample size simulation, we decided for a minimum target sample size of 1,500 participants with four TTO and six TTO questions per participant, based on the consideration that this sample size would allow us to only rely on TTO data if necessary and to include a safety margin that would allow us to exclude some lower-quality data. We included the state 111111 in all TTO blocks. After the first stage roll-out, we revised the number of DCE questions to five as we increased the number of TTO questions by one by adding the TTO learning state. We launched the third and final stage with a sample size of 1,500 participants. We decided to include the data from the second stage, but not from the first stage, for the final analysis. The changes between the first and second stage were major, but the changes in the third stage were mainly limited to spelling corrections. Including the 200 participants from the second stage, the effective sample size was around 1,700 participants.

3.2.8 TTO and DCE data quality assessment

To assess data quality for DCE we included a dominated choice set in the DCE task that was always included in the randomly generated DCE block for each participant.

For TTO, we needed to develop a new strategy compared with EQ-VT, which mainly relies on quality indicators that are calculated for each interviewer. We tried different ways to quantify the quality of the collected TTO data by calculating different metrics per participant or per question and to estimate a main effects linear regression model for including different percentiles of the better data. Specifically, we ordered the observations by the quality metric so that the data quality decreased as the percentile of included data increased. We then looked at how the mean t-score per coefficient, the mean value, adjusted R^2, and the value of the coefficients changed for each additional percentile of lower quality data included. We also analysed the impact of data exclusions on the final hybrid model by including 50% to 100% in 10% increments of the best quality TTO data in a sensitivity analysis.

Box 3.2.3: Level sum score

The *level sum score* (LSS) is the sum of all attribute levels and a proxy for the value of the state. For example, 111111 has an LSS of 6.

The best performing quality metric was a score that we developed to take into account the severity of inconsistencies in the TTO question per participant, the *combined inconsistency severity score* (*CIS*) . In brief, a TTO inconsistency occurs if a participant answers a higher TTO weight for a state A than for state B, if also state

A is dominated in terms of attribute levels by B. The severity is approximated by the normalized distance between the level sum score (LSS) plus the normalized difference between the two TTO answers involved in the inconsistency. For example, if inconsistency involved 232323 and 121212 with respective TTO answers of 0.2 and 0.5, the level sum score difference was 15-9 = 6 and the weight difference was 0.3. This inconsistency is for example more severe than another inconsistency involving the states 333332 and 333333 with answers 0.8 and 0.9, respectively. The score was again normalized for all participants per TTO block to accommodate for different numbers of theoretically possible inconsistencies per block.

More details on web survey development, experimental design, sample size simulations, and data quality assessment are available in [113].

3.2.9 Model comparison and value set modeling

To implement the hybrid model, we adopted a Bayesian perspective and relied on the Stan statistical software platform [127] with the cmdstanr [128] R interface, because this approach and software readily allows to incorporate complex models. We used a basic additive main effects specification: The constant encoded the weight for the state 111111, an attribute specific coefficient encoded the distance between the constant and level two for each attribute, and another attribute specific coefficient encoded the difference between level two and level three for each attribute. We restricted the coefficients to be positive.

In hybrid modeling, a separate model component is estimated for the DCE and TTO data [32] simultaneously. For the TTO data, we specified a linear model and for the DCE data a logit model. We linked the attribute-level coefficients between the linear and the logit model using a multiplicative constant.

In addition, we tried model features to improve the fit: Letting the error term depend on the magnitude of predicted weight to capture heterogeneity [129], and censored regression to accommodate that the TTO answers were limited by 0.1 and 1. In censored regression, observations at the interval limits are considered to possibly originate from beyond the interval limits, but to be only observable at the interval limits.

Further, the TTO tasks implies a weight of 1 for 333333 while the predicted weight for 333333 depends on the best fit of the TTO answers and can be lower than 1 which might be an issue for the face validity of the value set. To accommodate for this issue, we first attempted to model discapability from 1 instead of increasing capability attainment, and leaving out the constant, effectively binding 333333 to 1:

$$1 - y_i \sim X_i\beta + \epsilon_i$$

Second, we included an additional scaling factor in the model to rescale the predicted weight to 1 using a very strong prior, leaving the weight of 111111 encoded by the constant unchanged, resulting in the following specification (s being the anchor scale factor, and $\vec{x}_{333333} \cdot \beta$ the prediction for the state 333333):

$$y_i \sim \mu + X_i\beta \cdot s + \epsilon_i$$

$$1 \sim \mu + \vec{x}_{333333} \cdot \beta + \epsilon_i$$

To compare the models, we performed a k-fold cross validation with k=10 where we divided the data into 10 equally sized parts and used nine parts for estimating the model and the remaining part for assessing the model performance, repeating the same procedure 10 times for the 10 possible combinations. For DCE, we assessed accuracy in predicting the choices. For TTO, we looked at the mean error, mean absolute error, and mean squared error of predicted vs observed answers. Moreover, the TTO task involves additional assumptions regarding timing and scale that may introduce additional bias [28] regarding the relative magnitudes of the resulting coefficients in comparison to the DCE logit coefficients [130]. These on the other hand are unanchored and only depend on the binary choice answers. For this reason, we included a measure for how well the estimated TTO coefficients correspond to the DCE coefficients estimated with a pure logit model. To do so, we normalized the coefficients for the DCE logit reference model and the comparator model so that their sum each corresponded to 100% and then calculated the sum of the absolute differences per coefficient between the models. Thus, if coefficients were proportionally very different between a model and the DCE logit model, the score would increase and indicate a worse match.

More details on value set modeling and model comparison methods are available in the publication for S2 and an accompanying preprint [2,113].

3.3 Results

3.3.1 TTO and DCE answers

The resulting TTO answer distribution showed a large proportion of answers at 0.1, especially for 111111. The LSS was positively correlated with the mean TTO

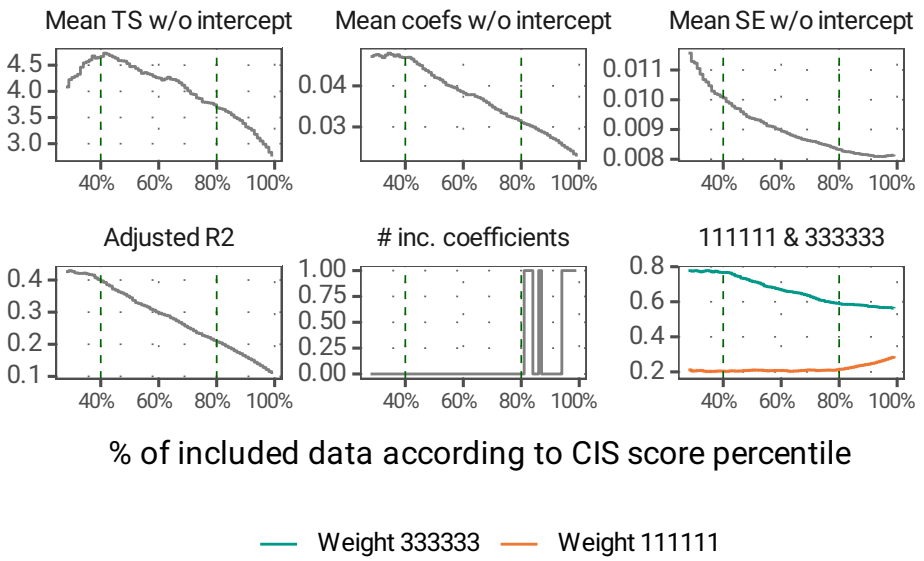

Figure 3.3: Data quality indicators by the included percentile according to the combined inconsistency score (CIS).

answer, and for DCE the choice proportion of a state correlated positively with the distance in LSS to the comparison state.

3.3.2 Data quality assessment

The participants failed the DCE consistency check and picked the dominated state for 17% of the questions.

For TTO, we found that the combined inconsistency severity score offered the best performance to discriminate lower-quality data and decided to include 80% of the better quality TTO data. The model fit became worse if we included more data, as indicated by a decrease in mean t-score of the coefficients, a decrease in adjusted R^2, and a decrease in the magnitude of the coefficients, the number of inconsistent coefficients, and an increase of the constant (Figure 3.3).

In the final dataset, 50% of the TTO data of all participants had at least one inconsistency. This share decreased to 38.3% after excluding 20% of the lower-quality TTO data. Nevertheless, the TTO data included in the final model still showed partial signs of limited quality, specifically clustering around the lowest value of 0.1, even for states other than 111111 (Figure 3.4).

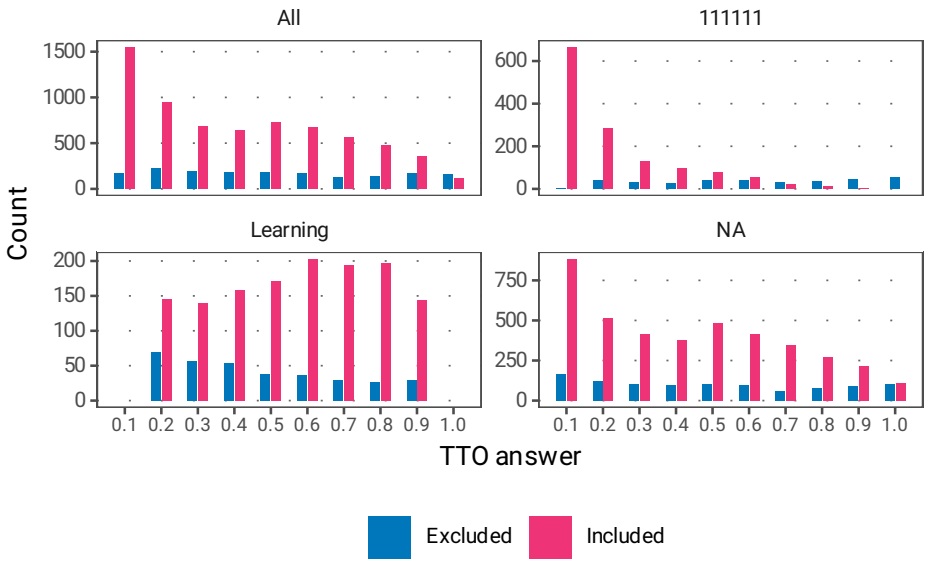

Figure 3.4: Distribution of the time trade-off (TTO) answers. "Other states" refers to all other states in the design besides 111111 and the learning state.

3.3.3 Model comparison and value set modelling

We compared different combinations of model features for hybrid models and TTO-data-only models – namely heterogeneity specification, anchor scale attainment, or discapability specifications; models with and without varying intercept; and models with and without censoring – and ranked them according to these metrics. We did not find significant improvement in the model fit for the heterogeneity specification or when we included an individual-varying intercept.

Models using the discapability specification where capability weights were modelled without constant resulted in stretched level 2 and 3 coefficients. This effectively distorted and reversed the order of magnitude of the coefficients compared with a DCE-logit model and hybrid models with a constant (Figure 3.5). In addition, the DCE accuracy decreased and the mean absolute error for predicting TTO answers increased. Compared with the hybrid models, the TTO-data-only models resulted in wider 95% CRIs for the weights. The hybrid model with censoring extended the range of the resulting value set and seemed to slightly improve the mean absolute error and to reduce the mean CRI width of the weights compared with the hybrid model without censoring.

We decided to generate the value set by using the hybrid model with an anchor scale and censoring, but without a varying intercept and heterogeneity specifica-

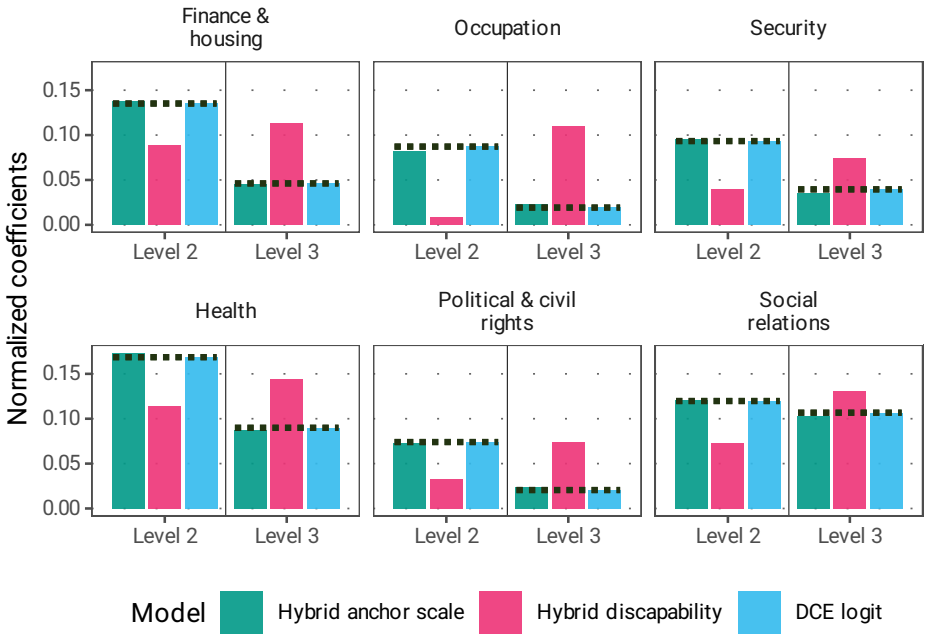

Figure 3.5: Comparison of the anchor scale, the discrete choice experiment (DCE) logit model, and the discapability model. The bars represent the magnitude of coefficients, normalized so that the coefficient sum per model without a constant corresponds to 100%.

tion. The estimated weight was 0.114 for 111111 and 1.001 for 333333 (Table 3.2). According to the decreasing order of magnitude of the summed coefficient per attribute, we found that health was most important, followed by social relations, financial situation & housing, security, occupation, and political & civil rights. The sensitivity analysis of including 50%-100% of TTO data in 10% increments revealed decreased model performance when including 90%-100% of TTO data compared with including 80%. Including 50%, 60%, or 70% of the highest quality resulted in a minor improvement in the model fit.

Attribute (Importance rank)	Coefficient name	Value (95% credible interval)
	Don't agree (constant)	0.114 (0.10, 0.13)
Health (1)	Agree partially	0.154 (0.14, 0.17)
	Agree completely	0.078 (0.07, 0.09)
Social relations (2)	Agree partially	0.108 (0.10, 0.12)
	Agree completely	0.091 (0.08, 0.10)
Finance & housing (3)	Agree partially	0.122 (0.11, 0.13)
	Agree completely	0.040 (0.03, 0.05)
Occupation (6)	Agree partially	0.073 (0.06, 0.08)
	Agree completely	0.021 (0.01, 0.03)
Security (4)	Agree partially	0.085 (0.07, 0.10)
	Agree completely	0.031 (0.02, 0.04)
Political & civil rights (5)	Agree partially	0.065 (0.05, 0.08)
	Agree completely	0.021 (0.01, 0.03)
	Anchor scale	0.545 (0.52, 0.57)
Other model properties	MAE	0.195
	MSE	0.06
	Accuracy	73.345%
	Number of observations	27,099
	333333	1.001

Table 3.2: Results of the main model. Mean square error (MSE), Mean absolute error (MAE), and accuracy refer to the results of the cross validation. The 27,099 observations consisted of 20,364 discrete choice experiment (DCE) answers and 6,735 time trade-off (TTO) answers.

Chapter

4 How do we fare? Applying CALY-SWE to measure the distribution and inequalities of capability (S3)

4.1 Overview

In S3, we applied CALY-SWE in a representative population sample to measure the capability distribution and assess capability inequalities between groups.

4.2 Methods

4.2.1 Questionnaire

To better understand how capability is distributed along relevant covariates, the survey asked a series of background questions after the CALY-SWE questionnaire. The selection and phrasing of these questions was inspired by the social determinants of health frameworks, materialism, and life course theories in social epidemiology [131]. For example, we included questions on materialistic circumstances such as employment and income, and on social background factors such as education, family situation, payment difficulties, and migration background. We also included a question on whether participants had involuntarily spent at least a year without studying or working between the ages of 16 and 29 years (i.e., were not in employment, education, or training - NEET).

4.2.2 Equity and comparing groups

We adopted an equity perspective and focused the analysis on comparing capability as measured by CALY-SWE between population groups, for the following reasons. First, information on inequalities between groups has the potential to inform policy and direct policy measures so that undue inequalities can be tackled. Second, applying CALY-SWE by comparing group capability might deliver evidence as to how appropriate CALY-SWE is to pick up hypothesized differences.

Box 4.2.1: Inequality, inequity, differences

We relied on a definition of *inequality* – "differences between individuals or population groups" [132] – applied to capability. A common distinction when assessing differences in outcomes is to categorize differences as *inequalities* or *inequities*. However, we refrained from doing so and used only the term inequality to avoid taking an a priori stance on what is and is not unfair so that we could discuss the extent of unfairness explicitly.

The group formation was informed by the Swedish policy context. Inspired by the Swedish platform on priority setting in healthcare [9], we formed two groupings: *plurality* and *disadvantage* (Table 4.1). Plurality groups are connected to the first principle in the platform, namely human dignity. Disadvantage groups are related to the second principle, need and solidarity. The decisive difference between the groups concerns the normative context of the grouping factor. For groups with a disadvantage factor, a capability shortfall would be a logical result because the disadvantage directly impedes an individual's capability, or indirectly impedes it by affecting a conversion factor. In comparison, no such a priori reason for a capability difference would exist for plurality groups. An example of a disadvantage group is individuals who were NEET between the ages of 16 and 29 years.

4.2.3 Statistical analysis

We described the distribution of capability by using tables and figures and assessed the attribute level and weight distributions for the overall sample and for groups. We analysed inequalities for the groups for the aggregated weight and for the attribute-specific weight components by calculating the mean capability difference per group member.

We further assessed the relevance of individual contributing factors. First, we estimated a multinomial logistic regression with the three answer levels as the outcome, and a selection of relevant covariates that that kept the variance influence

	Disadvantage	Plurality
Linked principle of the healthcare prioritization platform	Need and solidarity	Human dignity
Reason for inequality	Disadvantage in capability and/or conversion factor	No normatively acceptable reason
Acceptance of inequality	Unfair	Potentially unfair
Included groups	•Ever NEET •Unemployed •Poor self-rated health •Long-term sick leave •Payment difficulties •Homeless	•Men vs women •Lower education vs other education •Birthplace Global south vs other birthplace •Not Swedish born-parents vs Swedish born parents •Rural vs urban •Informal work vs other occupations •Age categories
Comparison mode	Against sample mean	Against comparison group

Table 4.1: Disadvantage and plurality group categorizations.

factor below 5. Next, we compared the influence of the covariates in contributing to the inequality between the two compared groups. We counterfactually altered the proportions of the first group to the proportions of the second group, one covariate at a time, and looked at how much the expected retrodicted weight of the first group changed. Then, we interpreted the resulting difference as the marginal counterfactual contribution (of the respective covariate) to between-group inequality.

To illustrate the magnitude of the inequalities for the entire Swedish population, we rescaled the per-individual averaged capability inequality up to the population level by multiplying by the inverse fraction of the sample size relative to the population size of those aged 18 or more for the year 2022 [133].

4.3 Results

4.3.1 Attribute level answers

Most participants answered with level 3 (*Completely agree*) for all attributes. Social relations showed the highest proportion of level 3 answers, followed by security (71% and 67%, respectively), while occupation received the lowest proportion,

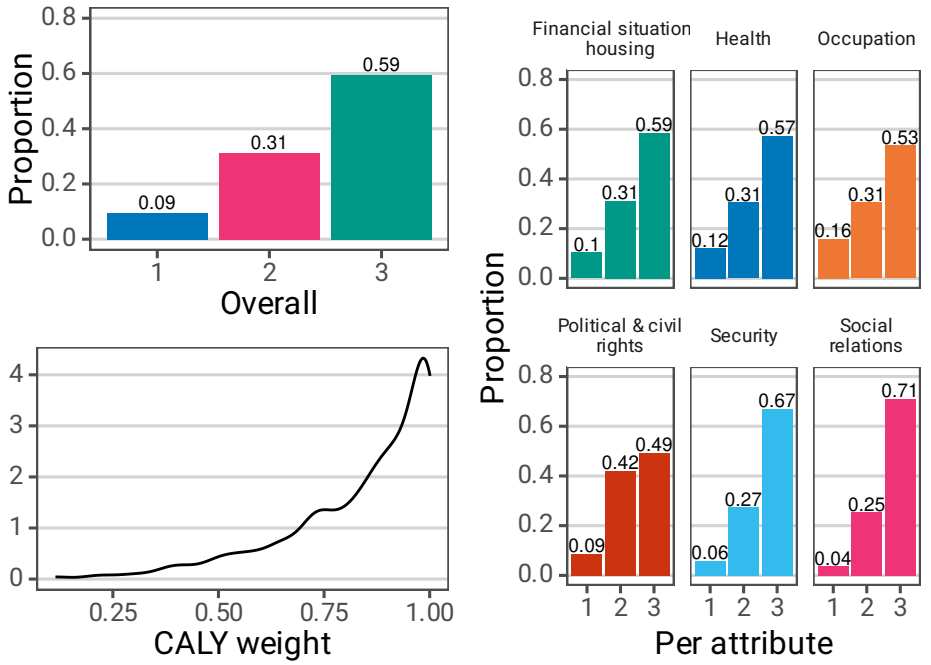

Figure 4.1: Top left: overall answer level proportions. Right: answer level proportions by attribute. Bottom left: overall weight distribution. Capability-adjusted life year (CALY).

resulting in a range of 53% to 71% across all attributes. The highest proportion of level 1 (*Do not agree*) answers occurred for occupation, followed by health (16% and 12%, respectively), with a range of 4% to 16% for all attributes; social relations had the lowest proportion. Overall, 59% of the participants answered with level 3, 31% answered with level 2, and 9% answered with level 1 (Fig 4.1).

4.3.2 Aggregated weights

The average CALY-SWE capability weight was 0.837. All disadvantage groups had a lower average weight than the sample; the lowest observed weight of 0.465 occurred for poor self-rated health. The inequalities ranged from -0.075 to -0.135, and none of the 95% CIs included 0 (Table 7.2).

The plurality groups had lower inequalities, ranging from -0.064 for lower education vs other education to -0.004 for a birthplace in the Global South vs being born in another place. The 95% CI for informal work vs other occupation and for birthplace in the Global South vs other birthplaces did contain 0 (Table 7.2).For age groups, capability differences were generally small, ranging from -0.02 (45-64 years) to +0.03 (65-105 years).

4.3.3 Weight components

The magnitude of weight differences per weight component varied for the disadvantage groups. Generally, health was important (especially for the poor self-rated health and long-term sick leave groups), as was financial situation & housing. Social relations was also important, albeit to a lesser extent for the unemployed and the payment difficulties groups (Figure 4.2).

For the plurality groups, differences in weight components were generally less consistent, more diverse, of a lower magnitude, and with more associated uncertainty. For women vs men, women reported higher capability in social relations, but lower capability in health and slightly lower capability in financial situation & housing. The reverse picture with lower capability in social relations but higher health capability emerged for being born in the Global South vs other birthplaces. Lower education vs other education showed lower capability in health, occupation, financial situation & housing, and social relations. Rural vs urban showed a similar picture, but with less pronounced differences (Figure 4.3).

The 18-29-year-old age group had a lower financial situation & housing capability and a higher occupation capability than the population average. A similar picture emerged for the second lowest age group, 30-44 years old, with the difference that occupation capability no longer was above population average. The 45-64-year-old age category had the lowest capability, with all attributes being below the population average and the largest shortfall for social relations. The highest age category of 65-105 years showed a positive capability difference that was mainly driven by the markedly better financial situation & housing and social relations attribute weight components (Figure 4.4).

4.3.4 Counterfactual contributing covariables

For the disadvantage groups, the most important covariables were generally the employment situation, financial difficulties, and monthly income. Unsurprisingly, Ever NEET status was important for the Ever NEET group, as was housing type for the homeless group.

The employment situation and financial difficulties were also important covariates in all plurality groups, albeit to a lesser extent for rural versus urban and the 45-64-year-old age group versus the other age groups. Monthly income also contributed to differences for women versus men, lower education versus other education, informal work versus other occupation, and the 18-29-year-old age group versus the other age groups. Unsurprisingly, the age group covariate was relatively important for the age groups, the gender covariate was important for

women vs men, and the rural urban categorization covariate was important for rural vs urban. More detailed results on counterfactual contributing covariables are available in the supplementary information to S3 (available upon request).

4.3.5 Population scale CALY consequences

The population scale confirmed the general picture of higher inequalities in the disadvantage groups. The payment difficulties and ever NEET groups showed a capability shortage of $>$ 100,000 CALYs. For the plurality groups, women versus men, the 45-64-year-old age group versus the population average, and lower education versus higher education had the highest capability inequalities, which stayed below 100,000 CALYs (Figure 4.5.

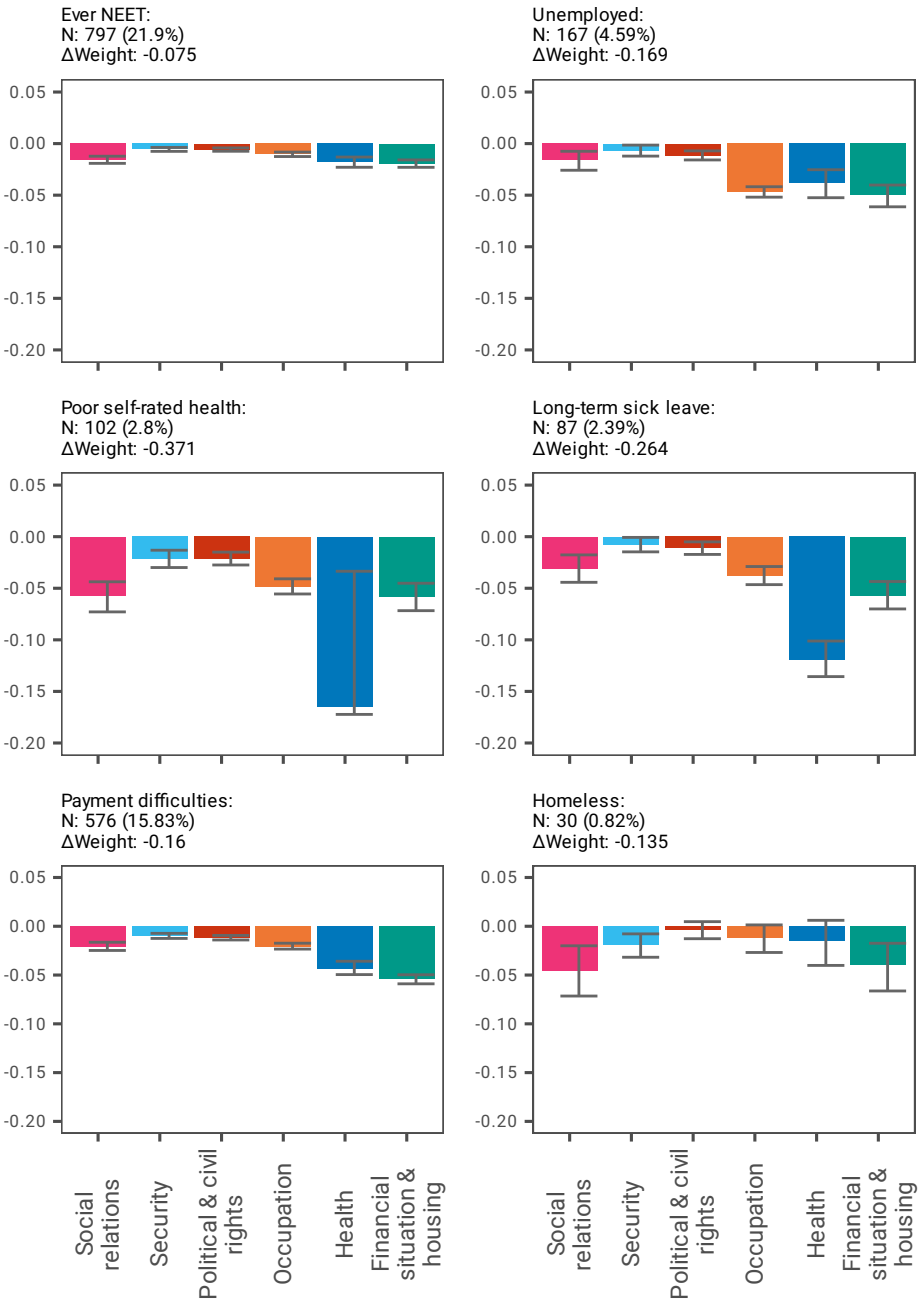

Figure 4.2: Weight component differences for the disadvantage groups. Adapted from S3. Not in employment, education, or training (NEET).

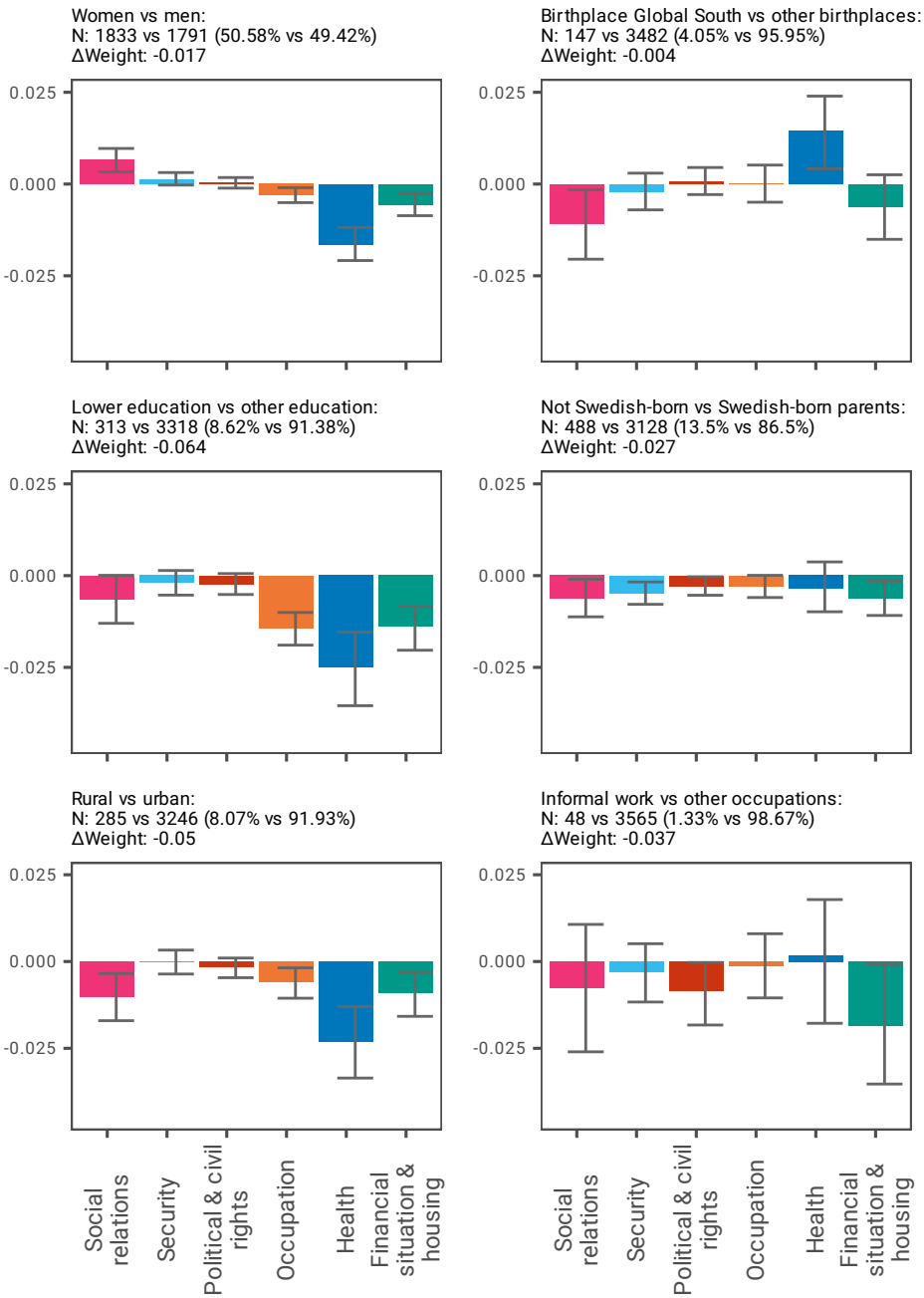

Figure 4.3: Weight component differences for the plurality groups. Adapted from S3.

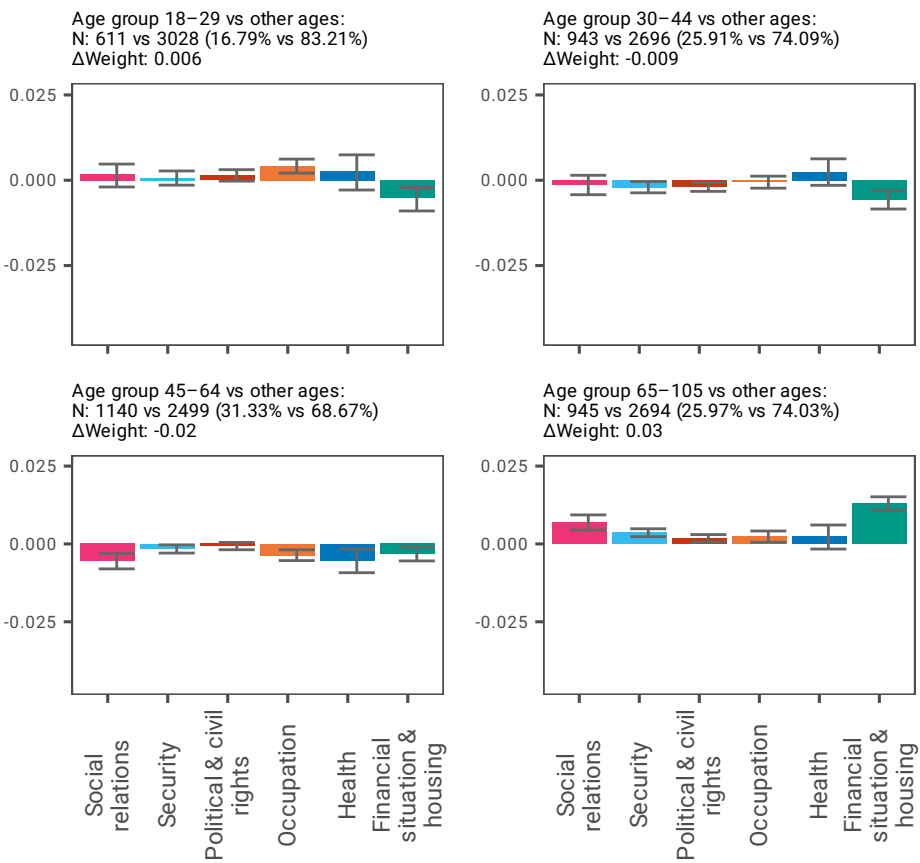

Figure 4.4: Weight component differences for the age groups. Adapted from S3.

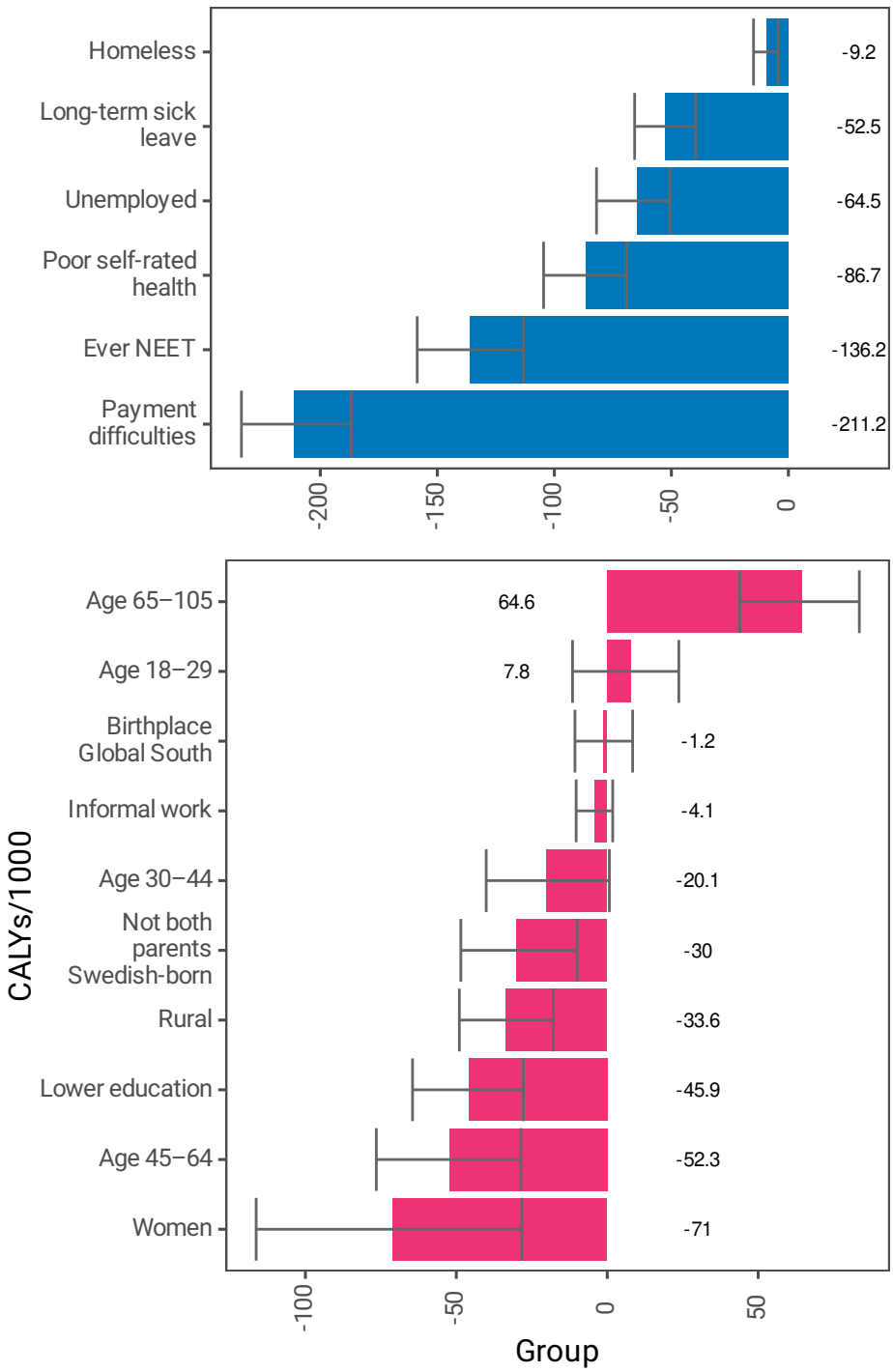

Figure 4.5: Capability-adjusted life year (CALY) consequences on the population scale. Adapted from S3. Not in employment, education, or training (NEET)

Chapter

5 Is rolling back the payroll tax paying back? Applying CALY-SWE in a cost-effectiveness study of NEET prevention (S4)

5.1 Overview

S4 was concerned with applying CALY-SWE in a cost-effectiveness analysis of the payroll tax cut from 2007 to 2016 for Swedish employees aged ≤ 26 years to calculate cost per CALY. The tax cut increased employment among young employees and we hypothesized that also NEET rates were affected.

5.2 Methods

5.2.1 Data sources

We used the dataset collected on the distribution of capabilities in the Swedish population (S3) that contained a question on the whether the participants had been involuntarily out of employment, education, or training during the ages of 16-29 years. This question defined the NEET outcome that we called *former NEET*.

Saez et al. [71] quantified the effect of the payroll tax cut on employment rates and found an increase in employment of around three percentage points for the affected age group, based on population-registry data. Importantly, they found evidence for a continuously heightened employment after the cohorts aged out of the affected age range ("spill-over") and even persistent effects for treated age

groups after the tax cut was undone (a "hysteresis effect"). We digitized these effect percentage point effect estimates that were provided in figures for different ages and years and converted them to unemployment rates.

We also obtained aggregated-level data from Statistics Sweden, namely nationwide levels of unemployment for the ages of 20-24 and 25-29 years from 1987 onwards [134,135], and data on the income distribution in deciles in 2019 [136] and median and the 10th percentile salary by age in 2019 [137]. We also obtained the population distribution by age in 2022 [133]

5.2.2 Modelling NEET proportions

To estimate the effect of the intervention, the available effects on unemployment first need to be translated into an estimate on the impact of NEET occurrence. The cross-sectional data on the capability distribution contains self-reported former NEET occurrence and was representatively sampled according to age. Thus, we assumed a longitudinal perspective and hypothesized that the longitudinal exposure to macroeconomic unemployment is linked to the probability of self-reported former NEET. The unemployment rates and effects reported by Saez e al. 2021 [71] are only reported with a 2-year frequency for 2003-2018. To have a longer time range, we applied the reported effect to the official unemployment series from Statistics Sweden that are available from 1987 for the ages of 20-24 and 25-29 years. This resulted in a factual scenario and a counterfactual scenario with higher unemployment if the intervention did not happen. We extrapolated figures until 2034 but assumed that the intervention effect on unemployment, starting from 2018, would eventually fade-out to zero by 2024.

Still, the NEET question and unemployment data were not readily usable to infer a statistical relationship. The unemployment series report yearly estimates and not the cumulative incidence for the 16-29-year-old age group as in the our cross-sectional data. Hence, we conceptualized the exposure to macroeconomic unemployment levels to accumulate on a yearly basis and calculated for each year the aggregated *influence value*. We defined the influence value as the mean of all macroeconomic unemployment rates that an individual passes through during the 10 years throughout the ages 20 to 29. For example, during the ages of 22-26 years, the unemployment rate of the ages 20-24 years would be included in the mean calculation for 3 years, and then for the next two years the unemployment series of age 25-29 years would be included in the mean calculation.

Next, we used the influence value as an outcome in a logistic regression on NEET occurrence, limiting the participants to those who reached at least 30 years of age in 2022. We then retrodicted and predicted former NEET occurrence for the

factual and counterfactual scenarios until 2034, including individuals turning 30 in the future. We included three possible confounders: gender, birthplace, and education. We chose the year 2034 because this was the first year when the cohort in the 10 years leading up to 30 years of age would no longer be affected by the intervention effect, which we assumed would fade-out by 2024. We calculated the number of NEET averted by summing the total predicted probability for NEET in the sample and taking the difference between the counterfactual and factual scenarios.

5.2.3 Cost-effectiveness

We used two perspectives (social and fiscal) and three scenarios (unemployment only, cautious, and optimistic). Here we report only the cautious and optimistic scenarios. The fiscal perspective only considered tax consequences for the Swedish state, while the societal perspective approximated full productivity effects. The optimistic scenario used unadjusted QoL and monetary effects, while the cautious scenario employed a factor of 0.75 for NEET numbers, monetary increases in tax and productivity, and 0.5 for QoL consequences. We calculated in 2019 dollars, with conversion from SEK to USD taking into account purchasing power parity [138].

For QoL consequences during the life course of former NEET, we calculated the observed QoL difference between former NEET and those not having been NEET, while adjusting for gender, place of birth, age category, and education, and then multiplied with the number of averted NEET. For unemployment QoL consequences, we used the unadjusted CALY-SWE weight difference between reporting being unemployed and those not reporting being unemployed, for participants in the working age (18-64 years old).

We included four monetary costs and saving components:

1. For the societal and fiscal perspectives, we calculated the cost of the intervention by multiplying the payroll tax loss by 9 corresponding to the 9 years when the cut was in effect (from 2007 to 2016), the same approach used by Saez et al. [71].
2. We calculated savings though increased employment by assuming a salary corresponding to the 10th salary percentile in the private sector, conservatively assuming that newly created jobs would be at the lower end of the pay scale.
3. For the societal perspective, we also assumed an additional cost through a substitution effect, such as if older employees with a higher salary get replaced with younger employees affected by the tax cut. We approximated

the substitution effect through the difference between the 50th and 10th salary percentiles in the private sector.

4. We assumed savings through higher income throughout the life course for averted former NEET, determined by taking the mean salary difference by self-reported former NEET status but only for participants that reported the lowest three income categories up to a monthly income of 29,000 SEK before taxes.

For the fiscal perspective, we calculated with tax consequences only, restricted to the employer payroll tax and the individual income tax. For 2., averted unemployment salary, we did not include individual income tax as we assumed that similarly high taxes would be paid on unemployment benefits instead. We adjusted the payroll tax rate according to the intervention. For 4., we simplified and assumed a 60% total tax rate for income and payroll taxes, as both income (for this lower income range) and payroll taxes are around 30% [71,72,139].

For the societal perspective, we applied a multiplicative factor of 1.6 to approximate productivity costs, considering around 50% extra costs through legal requirements per employee, for example, for holiday salary, payroll tax, and insurance [140,141], and assumed 10% extra costs representing overhead such as office space, non-salary benefits, and management.

We used a discounting rate of 3% for both costs and CALYs and bootstrapped the entire analysis to model uncertainty. To arrive at population-scale numbers, we multiplied estimates by the ratio of sample participants and population individuals aged 18-56 years in 2022. We conducted all analyses stratified by gender, although we mainly report sum-aggregated results for both men and women.

5.3 Results

5.3.1 Modelling NEET proportions

The unemployment effect was 389,000 and 542,000 job years created in the cautious and optimistic scenarios, respectively. Logistic regression for predicting self-reported former NEET resulted in 93,500 former NEET averted for the cautious scenario and 167,700 former NEET averted for the optimistic scenario from the societal perspective (Figure 5.1, Table 5.1).

5.3.2 Cost-effectiveness

The estimated underlying capability-related QoL improvements for averting former NEET were 0.074 and 0.086 for men and women, respectively. For averted

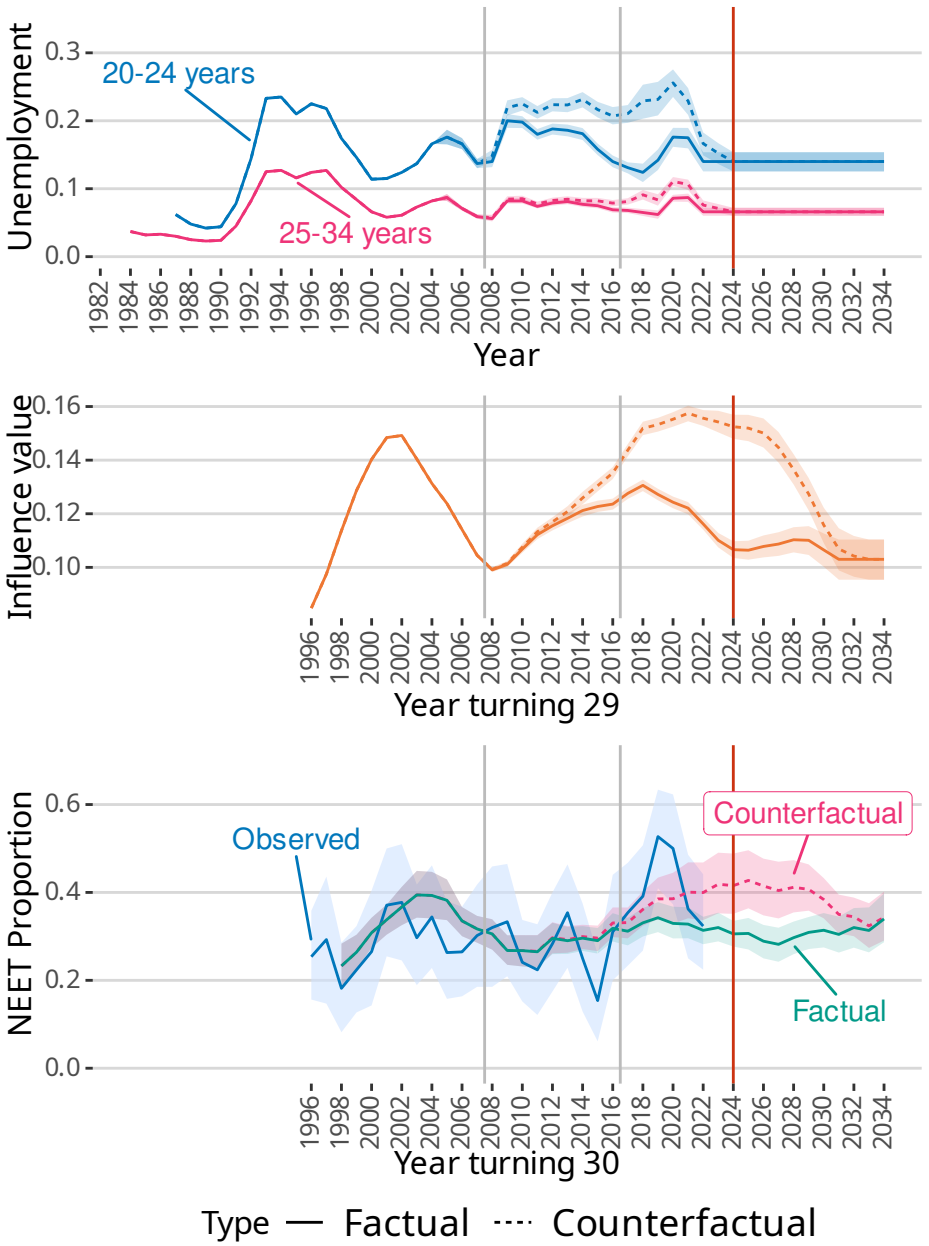

Figure 5.1: Unemployment (top), influence value (middle), and not in employment, education, or training (NEET, bottom) proportions. The factual (solid lines) and counterfactual (dashed lines) scenarios are shown. The grey vertical lines represent the duration of the payroll tax-cut intervention. Adapted from S4.

Metric	Cautious scenario	Optimistic scenario
CALY gain unemployment	49,347	91,692
CALY gain former NEET	123,290	366,268
Gross cost fiscal*	13,471	13,471
Tax averted unemployment*	−2,655	−3,696
Tax former NEET salary increase*	−1,280	−4,937
Net costs fiscal*	9,536	4,838
Gross cost societal*	18,614	20,639
Productivity increase averted unemployment*	−19,203	−26,760
Productivity increase former NEET averted*	−3,413	−13,165
Net costs societal*	−4,001	−19,287
Cost/CALY fiscal	55,236	10,565
Cost/CALY societal	−23,177	−42,115

Table 5.1: Cost-effectiveness results. Costs in 2019 Dollars. *in millions. Capability-adjusted life years (CALY). Not in employment, education, or training (NEET). Adapted from S4.

unemployment, they were 0.184 (men) and 0.157 (women). Yearly salary for avoided unemployment was estimated to be $31,688 (men) and $29,621 (women), and salary differences for averted former NEET landed at $589 (men) and $2,710 (women), with all salary effects showing large uncertainty.

Table 5.1 contains the cost-effectiveness results. Cost per CALY was $55,236 for the cautious scenario and $10,565 for the optimistic scenario from the fiscal perspective. From the societal perspective, both scenarios were cost saving. CALY gains were mostly driven by averted former NEET through the life course, while tax and productivity gains were more pronounced through averted unemployment.

The cost acceptability curve showed that a cost-effectiveness in at least 50% can be expected for a threshold above approximately $56,000, and that the tax cut would have a 95% chance of being cost-effective for a threshold above roughly $108,000 (Figure 5.2.

Crucially, judging cost-effectiveness is only possible to a limited extent. A cost-effectiveness threshold for CALY-SWE would need to be established first. On the other hand, some inference may be possible by taking the National Board of Health and Welfare's moderate cost threshold of around $50,000 for a QALY [142] as a lower limit: in theory, CALY-SWE should be valued higher because it measures more broadly than QALYs. However, other factors may play into the comparability of the scales, such as the range, given that CALY-SWE weights are restricted to be positive while QALY weights negative. Other methodological aspects such as phrasing, the number of levels, and weight valuation methods could also limit

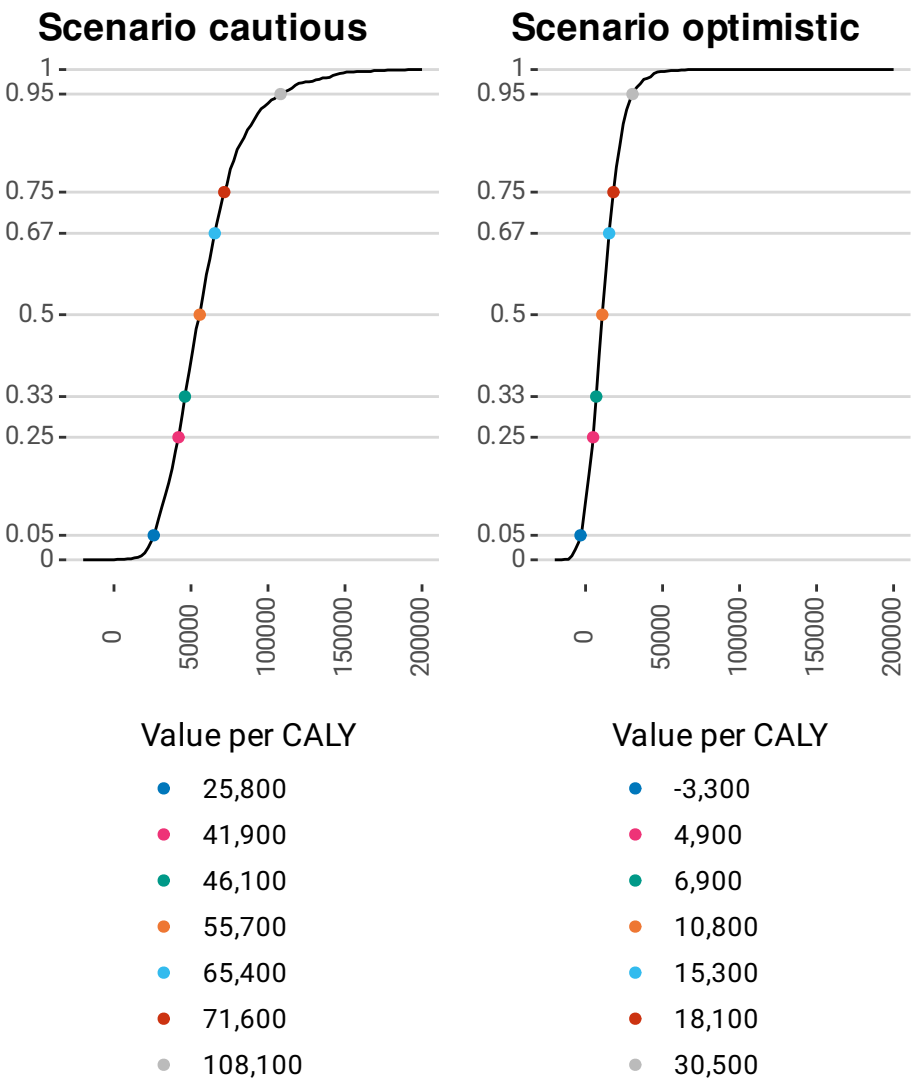

Figure 5.2: Cost-acceptability curves for the cautious and optimistic scenarios from the fiscal perspective. Capability-adjusted life years (CALY). Adapted From S4.

comparability.

Chapter

6 Data collection and ethical approval

6.1 Methods

All studies are at least partly based on individual cross-sectional data that were collected using an online panel [144] where the participants were invited and subsequentially redirect to the web survey in their web browser. After having completed the survey, the participants were automatically redirected back to the panel company.The participants were sampled to be representative of the Swedish population by using age, gender, and region of residence as criteria.

6.1.1 Datasets

Table 6.1 shows the different datasets used in the studies. The dataset used in S1 contained three variations of the same questionnaire (see S1 more details). The value set dataset (S2) was collected in three stages, where the first stage corresponded to a pilot stage. After changes to optimize the survey format for stage 2, we used the data from stages 2 and 3 for S2. We collected the data for S3 and S4 in six waves in the second half of 2022 to preclude potential seasonal bias. After collection, we merged the data. The size of the collected sample tended to vary slightly from the target sample size due to technicalities regarding the survey sampling procedure: automatic redirection seemed to fail sometimes, or the panel company sampler deviated slightly from the set target sample size. We used LimeSurvey [145] to collect the data for S1, S2, and S4, but not for S2.

Dataset name ***Used in***	**Sample size (target)** ***Sampling and population***	**Collection period** ***Mode***	**Content**
Delphi surveys *S1*	•1st round: 16 (18) •2nd round: 12 (18) *Non-profit civil society stakeholders purposive sampling*	January-February 2020 *Online survey*	•Ranking of capabilities and free text comment
Frequency wording (Version A, B & C) *S1*	Version (500) •A: 497 •B: 503 •C: 505 *Web panel general population quota sampling*	June 2020 *Online survey*	•CALY-SWE preliminary statements with different frequency wordings (versions A, B, & C) •Inequality aversion questions •Sociodemographic background questions
Value set *S2*	Stage •1: 111 (100, pilot) •2: 199 (200) •3: 1,498 (1,500) *Web panel general population quota sampling*	Stage •1:November-December 2021 •2: January 2022 •3: March-April 2022 *Online survey*	•CALY-SWE questionnaire •DCE questions •TTO questions •Sociodemographic background questions
Capability distribution *S3 & S4*	Wave (600) •1: 612 •2: 607 •3: 596 •4: 612 •5: 605 •6: 607 *Web panel general population quota sampling*	Wave (all in 2022): •1: July •2: August •3: September •4: October •5: November •6: December *Online survey*	•CALY-SWE questionnaire •Extended sociodemographic background questions

Qualitative web survey study *S2 **	16 *Purposive sampling local associations*	September– December 2020 *Qualitative interview*	•Semi-structured interview guide
Pilot study weights *S2 **	57 *Purposive sampling department staff*	February-August 2018 *Online survey*	•DCE questions •TTO questions

Table 6.1: Overview of the data sources used in the thesis studies. For the sample size column, the target sample size is displayed in brackets. *Data sources that were only used for methodological development, but do not contribute to the results of the studies directly. Third-party data sources such as Statistics Sweden [143] are not listed. Capability-adjusted life years Sweden (CALY-SWE). Discrete choice experiment (DCE). Time trade-off (TTO).

6.1.2 Ethical approval

We applied for ethical approval from the Swedish Ethical Review Authority [146] in two applications. The first application concerned the frequency wording data collection and the Delphi panel in S1, while the second application concerned both the value set and capability distribution data collections in S2, S3, and S4. For S1, the Swedish Ethical Review Agency issued an advisory statement that ethical assessment was not necessary because it deemed the data to not be sensitive. In contrast, the Swedish Ethical Review Agency judged the weight and capability distribution data for S2, S3, and S4 to be sensitive and granted ethical permission to collect the data and to conduct the studies. All participants provided informed consent, either electronically or on paper, and we followed the General Data Protection Regulation, other relevant Swedish legislation, and university guidelines for data management and storage.

6.1.3 Data exclusion

For all studies, we excluded participants with a self-reported age < 18 years for ethical reasons and with a self-reported age > 99 years due to data quality concerns.

6.1.4 Statistical software

We used R [147] for all statistical analyses.

6.2 Sample representativeness

We did not analyse sample representativeness for S1; rather, we relied on the representative quota sampling. For S2, after exclusions due to a self-reported age $<$

18 years and technical issues, the sample comprised 1,697 participants. Compared with the population, the sample was biased towards being Swedish born and with higher education, and it contained fewer people from municipalities with less than 100,000 inhabitants and people aged ≥ 80 years. In the dataset with totally 3639 participants for S3 and S4, there was no evidence for significantly different population proportions for age, gender, and region of residence that were used for quota sampling. On the other hand, there was evidence for bias towards higher education, being born outside of Sweden, and living in larger cities.

Chapter

7 Discussion

7.1 Main findings

Study 1:

- The Delphi panel selected health, financial resources & housing, social relations, occupation, security, and political & civil rights as the most relevant attributes for CALY-SWE
- CALY-SWE implements a sufficiency threshold, implying that a sufficientarianism distributive justice approach follows through its use

Study 2:

- The derived value set enables cost-effectiveness analyses using CALY-SWE, with weights spanning from 0.114 to 1.001
- The Swedish population valued health highest, followed by social relations, financial situation & housing, occupation, security, and political & civil rights

Study 3:

- The overall mean capability weight was 0.837. Measuring the population distribution of capability revealed shortfalls in capability for disadvantage

groups, in particular for poor self-rated health and unemployment, and to a lesser extent inequalities along plurality axes, such as between urban and rural population groups
- CALY-SWE can be used to identify disadvantage groups – and capability inequalities – who could be targeted in policymaking within a contextualized and nuanced framework for inequality

Study 4:

- Past cumulative exposure to macro-level unemployment seems to be linked to the risk of NEET (S4)
- The payroll tax cut from 2007 to 2016 that targeted young employees was likely cost-effective or cost-saving in preventing NEET in terms of cost per CALY from a societal perspective, but with lesser probability for the fiscal perspective, if inferring the monetary threshold value from national guidance for QALYs, and under the assumption of prolonged effects

In summary, this thesis has successfully moved CALY-SWE forward by developing the questionnaire and eliciting a value set. The thesis also demonstrated CALY-SWE's applicability in measuring capability and use in broad cost-effectiveness analyses of social welfare.

7.2 Normative choices in designing and valuing a measure

I believe that the most relevant way to discuss the development of CALY-SWE is in relation to normative aspects. The methodological decisions we took as researchers ultimately shaped what CALY-SWE measures. Indeed, the normative choices taken during the process of developing the measure potentially have far-reaching consequences and could influence resource allocation decisions, most importantly the selection of the attributes and phrasing, but also the elicitation of the value set. Hence, I have structured the discussion to reflect these normative aspects, before discussing the findings of the thesis and CALY-SWE from a broader perspective.

7.3 A democratic paternalistic process (S1)

The number of perceivable capabilities is large and necessitates a restriction of the measure to only the most relevant capabilities. The number of relevant capabilities

and what the most relevant capabilities are is inherently subjective. To ensure that CALY-SWE could measure QoL appropriately and fairly and could be useful in the Swedish policy context, we used an approach for selecting capability attributes that constituted a middle ground on three scales: global and local policy relevance, subjective and objective wellbeing, and democratic and paternalistic selection of capabilities.

While the Delphi panel itself is largely a democratic process, the idea that a panel selects the capabilities for the whole population but without direct democratic input of the population is to some degree paternalistic. The main motivation for this choice was that the Delphi panel participants are representatives of interest groups in Sweden with vast knowledge of the subjective wellbeing of those that they represent. On the other hand, we also took as starting point a Swedish list derived by perhaps more objective, subject-matter experts [63]. This initial list was based on a literature review of initiatives to measure general capability of life – for example, "How's Life?" from the Organisation for Economic Co-ordination and Develop (OECD) [148] and the Swedish level of living survey from 1968 [149]. The paternalistic aspect was also motivated by the intended focus of CALY-SWE on those worse-off [1] and the related invitation of Delphi panel participants with a core interest in worse-off groups, such as the Red Cross. The marginalized in a society experience bad QoL conditions, while the population average may be to some degree agnostic to their struggle. Thus, it is important that this perspective is given a strong voice.

One limitation of this approach was that not all invited participants participated and contributed to the selection of the capabilities, effectively foregoing the imposed duty to represent the voice of the represented interest group. Missing perspectives included migrants, patients with cancer, pensioners, business owners, and people who misuse alcohol or drugs (Table 2.1). Thus, the representativeness of the resulting Delphi panel was reduced, although achieving and even defining representativeness in the context a Delphi panel with civil society participants seems very challenging. For example, should the size of population groups be reflected or not? Relatedly, we adopted the fair-minded perspective partly to overcome the need for representativeness. An alternative Delphi panel composition might have resulted in a different list. Nevertheless, the rankings in S1 and S2, in addition to previous results based on rankings [150] by public health researchers for the same original list of 10 capabilities [63] as used in S1, indicated strong evidence that health, social relations, and finance & housing are most important.

The decreased representativeness may also be detrimental to the outfacing legitimacy of CALY-SWE. Even if we argue that fair-mindedness can compensate

for a lack of representativeness, population groups might disagree and consider that CALY-SWE does not appropriately capture their capability-related QoL. This might also undermine their trust in using the measure to inform policy. One particularly relevant group in this regard might include migrants who currently live in adverse socioeconomic conditions. On the other hand, it was probably beneficial that the more interested organizations voluntarily participated in the Delphi panel. They were willing to adopt a fair-minded outlook to consider the best interests of the overall Swedish society and to participate with good faith in the process.

7.3.1 Capabilities vs functionings

To what extent does CALY-SWE measure capability rather than functioning? For example, the occupation attribute statement asks whether one has been satisfied at least 75% of the time with their occupation during the last 5 years. While this wording may be more related to functioning, the wording for the health attribute – "Health that allows me to do work with or do what I like 95% of the days" – may be more related to capability due to the use of "allow". One practical assumption was that if there is an accepted normative ordering of functioning, then that capability may be approximated on average by assessing the highest achieved functioning. Thus, this phrasing is a compromise that considers other factors such as how long a relevant time horizon should be and how plausible the statement is.

7.4 Double counting

One interesting issue that arises through the choice of attributes concerns double counting. In the QALY context, QALYs have been argued to double count if better income leads to better health-related QoL and if, at the same time, it is counted as a productivity gain from a societal perspective, although there has been little empirical evidence for this phenomenon [13,151]. For CALY-SWE, this issue is likely more relevant, as financial resources are explicitly included. For example, S3 revealed correlation of financial resources & housing capability and income and financial difficulties, while at the same time S4 measured both salary effects and CALY-SWE weight differences to calculate cost-effectiveness. However, I would argue that the capability approach helps to conceptually dissect the issue: it is important to differentiate between capability and commodities [53] for a number of reasons. Money by itself can be seen as only a value promise and a means to organize society's economic activity [152], while capability is the end to maximize. Not all individuals have the same ability (conversion factor) to convert monetary commodities into capability, putting them at a disadvantage.

An illustrative example could be the higher cost of living in urban areas. A similar argument could be made that income is correlated to job satisfaction, which might be captured by the occupation attribute; however, some jobs might be more menial or risky and thus deserve higher renumeration. The answer to these issues surely cannot omit the attributes where monetary individual gains have a direct impact, but it is necessary to measure both and to keep them distinctly separated by the fraction bar (cost/CALY).

We also have not yet derived experience-based weights. Experience-based capability weights could be derived by analysing how the perceived overall capability QoL reported on a quantitative scale would vary with respect to the self-reported capability attribute-levels. Doing so would be valuable to offer an additional normative perspective on what capability value set is relevant. Another potential research interest is to compare the hypothetical value set derived in S2, as well as similar assessments of differences between hypothetical and experience-based value sets in the QALY context [153,154]

7.5 The CALY-SWE weight scale

The scale on which the CALY-SWE weights are measured implies additional important normative choices, specifically the conceptualization of the 0 and 1 anchoring points. In the health QALY context, 1 is commonly defined as full health and 0 as death, and negative weights are frequently allowed, such as for EQ-5D. For CALY-SWE, we defined 1 to correspond to a threshold that is sufficient for a flourishing life in present-day Sweden and 0 to correspond to 0 lifetime (Figure 7.1). The 0 anchoring is implicitly reflected in the TTO task phrasing where death was not explicitly mentioned.

7.5.1 Anchoring of weights relative to 0 (S2)

There were two reasons for not considering negative capability weights. First, negative capability is intuitively difficult to conceptualize, and it is unclear how it would fit into the broader capability framework, not least for the end-users of the value set. Similarly, this would possibly require a more explicit concept for the 0 reference point. In comparison, death as a reference point is quite conceivable in the health context given that it is a natural consequence of serious ill health, but it seems further away in the capability context. A similar argument was brought forward in the ICECAP-A valuation study: while 0 capability logically follows from death, it is unclear if 0 capability would automatically entail death [92]. Negative values for human life represented by QALYs have also been criticized

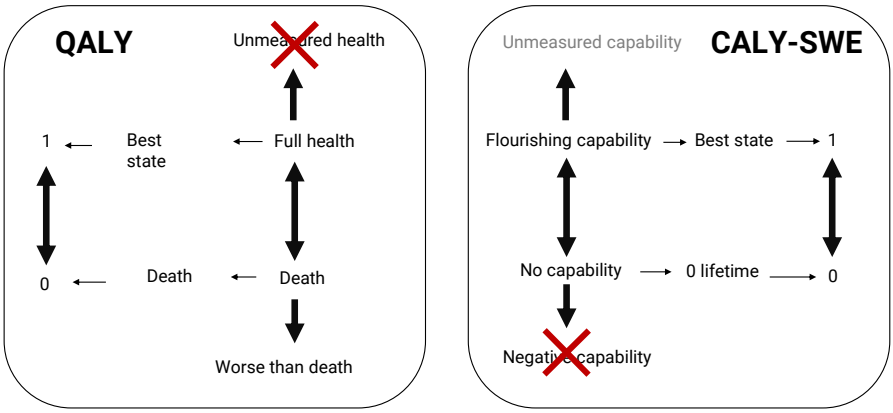

Figure 7.1: Comparison of quality-adjusted life year (QALY) and capability-adjusted life years Sweden (CALY-SWE) anchoring on the 0 to 1 scale.

in health economics. They might, for example, imply unethical consequences if negative QALYs based on social value sets are aggregated over time while costs remain the same, resulting in a greater loss. The results may then inform decisions against financing life-prolonging treatments. However, affected individuals in the respective health states might value their health-related QoL more than death [155]. Second, allowing negative weight would entail more complex TTO tasks [156–158] with presumably unfavourable effects for face validity and data quality.

7.5.2 Incorporating the sufficiency threshold (S1)

Constructing CALY-SWE so that the measuring range is focused below the sufficiency threshold implies a strong normative stance in that every application of CALY-SWE is inherently bound to this account of distributive justice. Instead, it could be argued that this should be an explicit consideration by policymakers. One advantage of using an implicit threshold is the inherent simplicity that allows uncomplicated usage of CALY-SWE while still incorporating equity considerations towards a more just distribution of capability. In health economics, methodological efforts to consider equity when using QALYs [159] have long been ongoing. It seems that the additional analytical effort required to estimate the efficiency-equity trade-off had been an obstacle for wider proliferation until recently [160–162]. On the other hand, policymakers cannot pursue a pure maximization principle by using CALY-SWE.

7.5.3 Anchoring of weights relative to 1 (S2)

The TTO answer distribution in S2 implied a weight for 333333 that assigned a value of less than 1 for 333333 using a linear model, which goes against the assumption of the TTO task where the highest capability state is anchored at 1. Furthermore, a value set with 333333 not receiving 1 might have limited face validity both normatively and practically: normatively, 1 year of lifetime in "full" quality should be counted fully, and practically, the meaning of 1 as the upper end of the scale is firmly established in the health-economics context. We counteracted this issue by introducing the additional anchoring scale factor into the model, so that the 333333 was anchored at 1. This is a limitation insofar as the resulting weights do not strictly correspond to the observed TTO answers. However, the lowest state 111111 remained unaffected by the anchoring and simultaneously had most TTO observations as it was included in all TTO blocks, so the lower anchoring should be precise and correspond to the TTO observations. The other attempted solution of modelling discapability without a constant would have distorted the relative magnitude of the coefficients.

7.5.4 TTO data quality (S2)

Using an unsupervised online survey was challenging in connection with the resulting data quality in the TTO question; perhaps this issue contributed to the discrepancy of the implied weight for 333333. While we tried to increase the data quality in several ways, it is likely that person-to-person interviews would have resulted in more valid and better-quality data. However, our resources were limited and conducting face-to-face interviews likely would have been less feasible, also considering that the value set study took place during the COVID-19 pandemic. In addition, not being accompanied by an interviewer left the participants free to reason independently regarding how to answer the DCE and TTO questions, lessening the risk for observation or desirability bias.

The clustering around 0.1, together with the limited range between the predicted weight for 111111 and 333333 in the ordinary linear model, might be an indicator of limited data quality. Clustering around the lowest reachable TTO answer might mean that the participants were quite willing to trade in lifetime to maintain the capability-related QoL level, that the 0 anchoring and the scale were limiting the effective answer range, or that there were artefacts in the TTO question design. Other possible explanations for the limited data quality could be that the TTO task was not well understood, there were genuine preferences for low values, or the participants implicitly confounded the sufficiency threshold internally and assumed a higher latent anchoring value for 333333, thus compressing the range

of observed values. However, other measures of face validity, such as the average TTO answer per state that correlated well with the LSS, indicated that the data collected corresponded reasonably well to the task assumptions.

The data quality assessment and data exclusion criterion developed in S2 can also be connected to the assumption of fair-minded thinking. Part of the justification for excluding the 20% of TTO data of participants with the lowest TTO data quality was that those may have given less consideration to the TTO valuation task. By implication, this could mean that those contributing lower-quality data were less fair-minded, and we argue that giving them an equal say would lead to a less appropriate value set. On the other hand, this approach risks imposing the prior expectations of the researchers, specifically the rationality of choosing the dominant alternative, on the data.

7.6 Capability distribution, groups, prioritization and capability shortfalls (S3)

One interesting result in S3 was that an older age did not per se automatically result in lower capability; possibly an indication for adoption of relative expectations to change life circumstances throughout the life course. Adaptive preferences in welfarism where individuals "level down" to adverse conditions and still experience high utility has been one of the motivating arguments for the capability approach. However, not all preference adaptions may be inherently negative [163,164]; they can also be seen as a "human ability to respond" [165] to change. Age is unavoidable, at least to some extent, affects humans universally, and ultimately leads to death. If this inevitability cannot be avoided, surely adapting to it in a sustainable and positive fashion should be legitimate.

While CALY-SWE is sufficientarian due to the inbuilt threshold, its intended usage at the same time resonates with prioritarianism. We designed CALY-SWE to find groups that are worse-off in terms of capability and to evaluate cost-effectiveness for interventions targeted at improving the situation of disadvantaged groups, as S3 and S4 demonstrate.

I argue that the by introducing cost-effectiveness evaluations to broader welfare, one should also be prepared for ethical challenges similar to those that are present in the healthcare context regarding prioritization. One particular question is whether it is always a good idea to choose the most cost-effective measure to gain CALYs. For example, in the healthcare context, the Swedish platform for priority settings stipulates that cost-effectiveness should be considered if doing so does not discriminate individuals based on their socioeconomic position, and after

considering their need for healthcare [9]. Thus, prioritization in healthcare based on gender is not opportune in Sweden, and the same should be true for capability in general.

A clear-cut case can be made for capability inequalities that follow from involuntarily constituted factors that directly negatively impact capabilities or conversion factors, for example a disability. In accordance with the principles of need and solidarity and human dignity [9], societies should prioritize those that are disadvantaged in this way, towards a more level-playing in terms of opportunities for all members of society.

However, in some cases inequalities occur along axes of plurality, namely factors that should normatively not result in inequalities, such as gender. Then, it may be necessary to lead a democratic debate if a prioritization, or positive discrimination, of a particular population group is just and warranted [166]. For example, in S3 we found a shortfall of capability in women compared with men, and affirmative action towards increasing the capability of women, directly or indirectly at the expense of men, may be legitimate. However, this approach may not hold for all plurality factors. For example, a society may find it unnecessary to prioritize artists with lower financial capability as long as they choose their profession freely, or if there is not high enough aversion against financial inequalities.

7.7 Using CALY-SWE to evaluate the cost-effectiveness of the payroll tax cut (S4)

7.7.1 Methodological limitations

We found the intervention to likely to be cost-effective, with high likelihood from a societal perspective but only with some probability from the fiscal perspective. However, this finding is subject to large uncertainty and several methodological limitations.

Importantly, we only established the association between NEET unemployment exposure by using cross-sectional data in an observational fashion and without a control group. True causality is difficult to estimate, although Saez et al. [71] argued that the clear delineation of the effects along age and year cut-offs constitutes supporting evidence. For similar reasons, it is complex to estimate true QoL effects as a consequence of the tax cuts. While it seems clear that having been NEET is associated with lower self-reported capability, it is unclear which individuals were actually treated by the intervention, and it is likely that those that profited the most were less vulnerable. Other socioeconomic risk factors correlated with becoming

NEET and QoL during the life course might have exacerbated the QoL impact of the intervention. However, we did control for confounding to some extent, namely gender, education, and place of birth, while also generally being conservative in parameter choices and only counting parts of the QoL consequences in the cautious scenario. Similar limitations apply to the estimation of salary consequences based on the cross-sectional data.

Another limitation concerns the ability to estimate the true costs from the societal perspective. We focused mainly on the payroll tax forgone through the tax reduction, assuming that this could capture the value consequences of having to finance the tax cut. In addition, we assumed that additional costs could be captured by modelling a substitution effect. However, this is only an approximation. A more accurate picture may need to consider how the tax cut was financed – for example, if other state spending was displaced or through the costs of issuing debt.

7.7.2 Policy implications

S4 demonstrated that CALY-SWE can be used to estimate QoL effects and to conduct cost-effectiveness evaluations of large scale macro-economic policies with broad social welfare consequences. The likely cost-effectiveness of the payroll tax cut reduction is by itself positive, mainly because it constitutes some evidence that the tax cut prevented NEET status and increased an individual's capability-related QoL, while not being overly costly. However, it is perhaps not right to conclude that, in general, lowering taxes for the sake of easing unemployment is a sensible choice. The analysis depends on the assumption of prolonged and marked effects of the tax cut on unemployment even after the tax was increased back to normal levels. Thus, this effect may not be repeated in the same way by cutting payroll taxes again, as an identical effect would be less likely. Moreover, it can be argued that reintroducing the original tax level was the right choice to secure tax income, if the positive effects of "ending youth discrimination" [71] are a one-off, permanent achievement.

One advantage of having established a cost-effectiveness estimate is that of future comparability. Seeing that the payroll tax reduction was mostly untargeted – the only criterion was being ≤ 26 years old – other more targeted interventions that manage to avert NEET status might be even more cost-effective. The results of such interventions could then be compared with the results of S4 to yield evidence on cost-effectiveness that is useful for policymaking.

Similarly to healthcare prioritization [9], pure cost-effectiveness should neither be the sole criterion for decision-making. Decision-makers need to consider different factors before adopting a policy. In the case of a tax reduction, it seems

	ICECAP-A, ICECAP-O, -SCM, -CPM, CYP	OxCAP, OCAP-18, OxCAP-MH	ASCOT	Questionnaire by Anand and van Hees 2006	CALY-SWE
Selection	Qualitative interviews	Indicators mapped to Nussbaum's list	Previous measures, qualitative interviews, literature review, Delphi process	Custom objective list partially related to earlier utilitarianist objective list [167]	Delphi process among civil society representatives
Attributes	•Stability •Attachment •Autonomy •Achievement •Enjoyment (For ICECAP-A)	•Life •Bodily health •Bodily integrity •Senses imagination and thought •Emotions •Affiliation •Other species •Play •Control over one's environment	•Accommodation cleanliness and comfort •Safety •Food and drink •Personal care •Control over daily life •Social participation and involvement •Dignity •Control over daily life Occupation and employment	•Happiness •Sense of achievement •Health •Intellectual stimulation •Social relations •Environment •Personal integrity •Capabilities overall	•Financial situation & housing •Health •Social relations •Occupation •Security •Political & civil rights
Valuation	BWS		BWS and TTO		TTO and DCE

Table 7.1: Overview of some prominent capability outcome measures, attributes, and selection thereof, and valuation method. Adapted from S1. Refer to table c1t1 for key references. Best-worst scaling (BWS). Time trade-off (TTO). Discrete choice experiments (DCE). ICEpop CAPability (ICECAP) – measure for Adults (A) – measure for Older people (O) – Supportive Care Measure (SCM) – Close Person Measure (CPM) – Children and young people (CYP). Oxford capability (OxCAP). Capability-adjusted life years Sweden (CALY-SWE).

relevant to consider, for example, the overall macroeconomic situation.

Another interesting discussion involves the concepts brought forward in S3. While NEET represents a disadvantaged group, reducing payroll taxes to all people aged ≤ 26 years potentially involves aspects that could be argued to be unethical, making these people more attractive to employ. While this factor is less problematic for those that would end up outside of employment, education, or training, individuals who are not disadvantaged could also potentially profit. Those could now be advantaged compared with older employees, leading indirectly to discrimination ("ageism") and substitution effects.

7.8 | Comparison to other capability outcome measures

All of the capabilities included in CALY-SWE can be related to capabilities on Nussbaum's [56] more abstract list that also underlies the OxCAP family of measures (Table 7.1), Table 3 in Meili et al. [1]). For example, social relations in CALY-SWE may be related attachment in ICECAP-A. Similarly, financial situation & housing and occupation overlap with autonomy, achievement, and progress in ICECAP-A. Achievement as well as social relations are also included in the questionnaire by Anand and Hees [100], and similar additional relations could be identified for the other attributes in CALY-SWE and other comparable capability measures. All in all, there seems to be agreement for the importance of health and social relations across the compared measures.

Compared with CALY-SWE that used a Delphi panel, the process of selecting capabilities for capability questionnaires usually involved participatory research with informants sampled from patient groups or their carers, the general public, or professional care workers [102]. An empirical or qualitative selection procedure with general population participants for CALY-SWE might have resulted in a different list of capability attributes that are more relevant from an individual wellbeing perspective. However, if one is willing to accept that paternalism, or at least a democratically deliberated common ground, may increase the degree of objectivity and in turn interpersonal comparability of capability-related QoL, the Delphi process offers an advantage over purely relying on an individual's view regarding what is important to them. Arguably, the Delphi panel participants had more experience in reflecting on the QoL of others because they are formal representants of various societal groups that are partially disadvantaged.

I argue that one distinction of CALY-SWE compared with other capability lists pertains to the policy relevance of the included attributes. The original list of capabilities presented to the Delphi panel purposely did not include more personal wellbeing aspects of happiness and life satisfaction: *"[translated from Swedish] Investigations assessment: Measures on life satisfaction and wellbeing should not be included in official follow-up/monitoring ..."*. This form of policy relevance manifests itself in CALY-SWE attributes and the phrasing. For example, social relations in CALY-SWE mentions "I have access to advice and support", while for the attachment attribute ICECAP-A includes the phrasing "I can have ... love, friendship and support". Arguably, the romantic aspect of social connections is one that is furthest away from state involvement and lies largely within an individual's sphere: while interventions to prevent loneliness are perceivable social care or public health policy measures, similar measures for fostering romantic relationships would probably be less socially accepted.

In addition, the information content of the capability attributes themselves

may be more specific and meaningful for policy for CALY-SWE than for some of the other measures. For example, the effects of poor working conditions would likely be captured distinctly by the financial situation and occupation attributes. Compared with Nussbaum's central capabilities, besides 10.B, "...being able to work as human beings, exercising practical reasoning and entering into meaningful relationships of mutual recognition with other workers ...", it seems that all of 4. Senses, imagination, and thought; 5. Emotions; 6. Practical reasoning; and 7. Affiliation could also capture aspects of working conditions. Similarly, for ICECAP-A, the effects of poor working conditions could be mostly reflected by 4. Achievement and progress, but also by all other attributes. At least for the Swedish context, it appears that the CALY-SWE questionnaire is more concrete and relevant for policy.

We could identify [102] one other capability measure that also seems to have involved a Delphi process for selecting attributes, namely the social-care focused ASCOT [99]. It is in part based on a predecessor that used a "Delphi exercise" [99,168] with a reference group consisting of authority and care managers, academics, officials, and voluntary sector participants. Although comparisons are not straightforward due to the different scope and context, the reference groups also seemed to deem physical safety as relevant, similarly to the CALY-SWE Delphi panel.

Regarding value set elicitation, there have been few valuation studies [102] of economic evaluation capability questionnaires. Up to now, the ICECAP family of measures seems to have relied on best-worst scaling, anchoring the highest and lowest state to 1 and 0, respectively [91,92,95]. Generally, not much is known about the TTO scale in a capability context, and this work represents one of the first applications. To our knowledge, only the ASCOT valuation also employed TTO for a capability scale, albeit using face-to-face interviews, and in combination with DCE questions but without a hybrid model [99].

A sufficiency threshold has also been conceptualized for ICECAP [169–171], but differently than for CALY-SWE. While we incorporated it implicitly in the wording, in ICECAP-A it was defined explicitly as having all attributes at level 3 of 4.

Regarding the broader picture of furthering the use of cost-effectiveness analysis, introducing an additional option in the form of the CALY-SWE questionnaire might increase fracturing and limit comparability if cost-effectiveness analyses were done with different outcome measures within a payer's responsibility. This would limit one main advantage of using CALYs, namely comparability of effects in unrelated fields. However, potentially different outcome measures can still be

compared – for example, by using mapping algorithms to translate between different outcome measures [172], or simply by using monetary thresholds. On the other hand, it could be important to have measures that are specifically apt for a given use case, and evaluators themselves should be able to choose the most relevant one, with the most suitable outcome measures asserting themselves eventually.

7.9 Would Amartya Sen object?

In *Inequality Reexamined*, Sen highlighted the importance of not taking the capability approach as an "all or nothing" exercise [55], where for any two individuals without uncertainty one individual can be categorized as better-off than the other individual. In other words, Sen stressed that no complete ordering of capability should be required, and that the capability approach inherently needs to allow incomplete, partial orderings. For example, this notion of the capability approach might allow two individuals to rank capabilities differently, or some capability might simply not be comparable between two individuals

In comparison, selecting capabilities and assigning a weight score to each possible combination of capabilities, the approach used in this thesis, directly assigns a complete ordering and thus violates this ambiguity. While this might be an important limitation of the adoption of the capability approach, practical considerations and usability may justify deviations. Furthermore, CALY-SWE weights represent population averages rather than individual observations. Orderings that result from their usage are only complete orderings if they are used in this way; where they represent aggregated averages from multiple individuals. For individual observations of attribute levels without the weight score, incomplete orderings are likely to occur as many pairs of CALY-SWE configurations are not logically ordered, for example 222333 vs 333222.

7.10 CALY-SWE in the context of the Swedish welfare state

I argue that the context-specificity for the Swedish is one principal argument behind the development of CALY-SWE. Then, a natural question arises: what part of CALY-SWE is specific to the Swedish context? However, this argument should not be taken too far. There might not be an actual need for some specific characteristic, as in theory CALY-SWE could be equally relevant to other contexts and in similar circumstances a similar capability measure might have emerged elsewhere. Still, context specificity here could at least mean that the relevance for

the Swedish context needs to be established by some democratically legitimate practice, such as the Delphi panel.

It nevertheless is likely that there are context-specific aspects regarding CALY-SWE and its potential applications. One is the policy context as discussed for S3. For example, the degree of equity-efficiency trade-off a society is willing to make might vary: it might go further in some countries compared with others [173].

Sweden is classified to be part of the *Nordic welfare model*, where equality, solidarity, the role of the state, and universal rights have traditionally been characterized as more relevant than for other welfare regimes [174]. However, these characteristics may have lost some strength since the 1990s [175], and the concept of the Nordic welfare model itself has been interpreted differently but still maintains it relevance [176].

Incorporating an implicit sufficiency threshold into CALY-SWE with the goal of achieving a more equitable capability distribution might be more acceptable within a Nordic welfare state. Similarly, perhaps the level of the sufficiency threshold itself would vary compared with other contexts, such as if less inequality aversion for health were to occur. The comparatively large role of the state could also increase the potential to adopt CALY-SWE because a more centralized, top-down approach to structuring and integrating cost-effectiveness evaluations into social welfare would be pushed. From a pragmatic perspective, the higher social spending is, the more relevant cost-effectiveness becomes, and Sweden has traditionally been above the OECD average regarding social spending as a share of GDP [177]. However, it is worth mentioning that the state does not necessarily need to be the direct provider of social welfare measures for CALY-SWE to be useful for resource allocation. For example, governments could mandate cost-effectiveness analyses to inform decisions and to design legally binding policies for private-sector actors.

On the other hand, one might also argue that the CALY-SWE attributes are rather general, so they could be relevant for most west and central European welfare states. This argument has also been made for the initial list of 10 capabilities from the Swedish governmental investigation [63].

The different responsibilities for social welfare areas are also important to consider regarding CALY-SWE. For example, regions are responsible for healthcare [109], while municipalities are responsible for social care and elderly care [109,178]. Currently, cost-effectiveness in terms of QALYs largely concerns the areas for which regions are responsible, while we developed CALY-SWE with the explicit intent to enable similar usage for areas that are relevant for municipalities.

Even other actors could use CALY-SWE, including the regions or national level authorities, actors in intersectoral public health, or non-profit organizations. If cost-effectiveness analyses via CALY-SWE were to be done by actors in different sectors, then the results would be comparable, as intended. By extension, if cost-effectiveness estimates are compared between sectors, then the distribution of financial resources between sectors might need to be reconsidered to equalize for the sake of maximizing gained CALYs. Or, thresholds could be allowed to vary by sectors. In practice, differential cost thresholds by sector already occur when comparing value for a statistical life used in traffic policymaking and willingness to pay per QALY [179]. In other words, achieving allocative efficiency new mutates into a problem with an additional dimension of production efficiency of maximizing CALYs [180].

Furthermore, there might be geographical variation in cost-effectiveness. For example, municipalities with larger shares of vulnerable populations might only be able to afford cheaper interventions in terms of cost per CALY. While variations by sector might be acceptable, geographical variation in cost-effectiveness might be perceived as unfair and motivate policy action. In any case, broader outcome measures such as CALY-SWE could help to shed light on such ethical issues.

7.11 Further limitations and missing pieces

While we used population quotas for age, gender, and region of residence for the survey data that underly all studies in this thesis, the resulting samples were only partially representative of the Swedish population. Specifically, we observed a bias towards higher education and Swedish-born participants. These biases are undesirable, for example, regarding how much CALY-SWE weights represent population norms. However, while representativeness was reduced, I still deem CALY-SWE to be acceptable.

While online web panels offer an affordable and easily manageable way to collect data, they have some disadvantages. Representativeness may be reduced by selection into the panel, and the transparency regarding the characteristics of panel members is limited compared with other methods of sampling study participants.

Furthermore, a general limitation that affects all studies is that the CALY-SWE attributes are restricted to three answer levels, instead of, for example, five answer levels, which might have offered additional sensitivity. We made this early design choice to ensure that the development of the questionnaire remained feasible with three levels, given several conceptual challenges such as the selection of attributes, statement phrasing, and the cognitive burden in the DCE and TTO valuation tasks.

Deciding on what research could be conducted during the CALY-SWE project and this thesis required a compromise. We did not address relevant questions to leave room for what we deemed the most relevant and beneficial research goals. Importantly, we have not yet explicitly investigated the psychometric properties and validity of CALY-SWE, which many would consider to be a fundamental requirement for establishing the validity of the measure. Nevertheless, the results of the quantitative part of S1, further work on perceived capability during the COVID-19 pandemic [181], and the results of S2, including the qualitative part, and S3 indicate to some extent reliability and validity: the importance of health, social relations, and financial situation and housing was reflected both in the Delphi panel results, and the TTO and DCE task results as well as the qualitative interviews in S2. Even the earlier ranking study of the initial list with 10 capabilities revealed that health, social relations, and financial situation were deemed most important by a sample of researchers and doctoral students [150]. Additionally, the results from S3 confirm some known inequality patterns in public health – for example, for the axis of financial difficulties and lower education [182].

We also have not yet derived experience-based weights. Experience-based capability weights could be derived by analysing how the perceived overall capability QoL reported on a quantitative scale would vary with respect to the self-reported capability attribute-levels. Doing so would be valuable to offer an additional normative perspective on what capability value set is relevant. Another potential research interest is to compare the hypothetical value set derived in S2, as well as similar assessments of differences between hypothetical and experience-based value sets in the QALY context [153,154]

7.12 Outlook

Future research on CALY-SWE should focus on establishing psychometric properties in relation to validity and reliability, to elicit experience-based weights, and to extend CALY-SWE to five levels. Furthermore, the conceptualization of the 0 anchoring should be developed to unravel its meaning in a capability concept. In the medium to long term, what capability attributes are relevant for the Swedish context might change, and it could become opportune to confirm or adopt the included dimensions. Similar research in other context than the Swedish one could also be of interest. Applying comparable Delphi panel approaches in other countries might result in better knowledge on the relevance and selection of capabilities in different contexts. Similarly, repeating the weight study S2 in other countries would deliver insights on the relative value of capabilities in other contexts.

Ideally, improved versions of the TTO task might be developed to elicit capability value sets in similar circumstances, namely the online unsupervised TTO task, to address the resulting gap between the weight of 333333 and 1. Alternatively, future research could derive value sets by using other methods such as face-to-face interviews and then compare them to the results of S2.

Importantly, CALY-SWE should also be tested and assessed in more classical cost-effectiveness studies of localized interventions of trials in observational studies. Moreover, the relation of CALY-SWE attributes and to what extent one capability enables other capabilities should be analysed, both in isolation and in connection with the overall observed distribution of capabilities in population groups.

In the medium to long run, CALY-SWE and other measures can hopefully contribute to an increase in to a more equitable, transparent, and efficient resource allocation in broader social welfare. Beyond that, a moral case should not only be made for improving individuals QoL through social welfare while keeping the costs in check for taxpayers, but also for freeing up resources for allocation in other areas. Additional economical resources are needed to help humanity to address other pressing needs, such as combating climate change, sustaining long-term prosperity in the face of demographic shifts and rapidly advancing technological progress, and addressing global inequality.

7.13 Conclusions

All in all, thesis provides a basis for establishing CALY-SWE as a capability-based outcome measure in larger social welfare and public health by developing the questionnaire and a corresponding value set. The thesis demonstrates two applications: economic cost-effectiveness evaluations of interventions with broad consequences and measuring capability distributions, providing reference points for future works. The development of CALY-SWE also revealed the importance of methodological and normative underpinnings. Methodological and normative choices potentially come with high stakes due to the large scale of economic resources within CALY-SWE's broad measurement scope. Nevertheless, CALY-SWE may help in striving for better resource allocation decisions.

References

1. Meili KW, Månsdotter A, Sundberg LR, Hjelte J, Lindholm L. An initiative to develop capability-adjusted life years in Sweden (CALY-SWE): Selecting capabilities with a Delphi panel and developing the questionnaire. PLOS ONE [Internet]. 2022 Feb 8 [cited 2022 Feb 14];17(2):e0263231. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0263231
2. Meili KW, Mulhern B, Ssegonja R, Norström F, Feldman I, Månsdotter A, et al. Eliciting a value set for the Swedish Capability-Adjusted Life Years instrument (CALY-SWE). Qual Life Res [Internet]. 2023 Sep 11 [cited 2023 Sep 12]; Available from: https://link.springer.com/10.1007/s11136-023-03507-w
3. International Conference on Primary Health Care. Declaration of Alma-Ata. WHO Chron. 1978 Nov;32(11):428–30.
4. Meulen R ter. Solidarity and Justice in Health Care. A Critical Analysis of their Relationship. Diametros [Internet]. 2015 Mar 14 [cited 2023 Nov 1];(43, 43):1–20. Available from: https://diametros.uj.edu.pl/diametros/article/view/710
5. Assembly UNG. The Universal Declaration of Human Rights [Internet]. U.N. UDHR, Res 217 A Dec 10, 1948. Available from: https://www.un.org/en/universal-declaration-human-rights/index.html
6. van Till SAL, Smids J, Bunnik EM. Access to effective but expensive treatments: An analysis of the solidarity argument in discussions on funding of medical treatments. Bioethics [Internet]. 2023 Feb [cited 2023 Nov 1];37(2):111–9. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10098599/
7. Brazier J, Ratcliffe J, Salomon J, Tsuchiya A. Measuring and Valuing Health Benefits for Economic Evaluation. Second Edition. Oxford, New York: Oxford University Press; 2016. 372 p.

8. Brülde B. Folkhälsoarbetets etik [Public health work ethics] [Internet]. Studentlitteratur; 2011. 307 p. Available from: https://books.google.com?id=QqoZzQEACAAJ

9. Health and Social Affairs M of. Prioriteringar inom hälso- och sjukvården [Prioritization in healthcare] [Internet]. Prop. 1996/97:60 1997. Available from: https://lagen.nu/prop/1996/97:60?attachment=index.wpd&repo=proptrips&dir=downloaded

10. Ord T. The Moral Imperative Toward Cost-Effectiveness in Global Health. In: Effective Altruism [Internet]. Oxford University Press; 2019 [cited 2023 Nov 1]. p. 29–36. Available from: https://academic.oup.com/book/32430/chapter/268751816

11. Budgetlag. Budgetlag 2011:203 [Budget act] [Internet]. 2011:203 2011. Available from: https://www.riksdagen.se/sv/dokument-och-lagar/dokument/svensk-forfattningssamling/budgetlag-2011203_sfs-2011-203/

12. Williams A. Cost-effectiveness analysis: Is it ethical? Journal of Medical Ethics [Internet]. 1992 Mar 1 [cited 2023 Nov 1];18(1):7–11. Available from: https://jme.bmj.com/lookup/doi/10.1136/jme.18.1.7

13. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. 4th ed. Oxford, New York: Oxford University Press; 2015. 464 p.

14. Nilsson I, Wadeskog A. Focus on the individual. 2008.

15. Nordesjö K. The constitutive effects of social investment evaluation. Evaluation [Internet]. 2021 Apr [cited 2022 Sep 15];27(2):210–28. Available from: http://journals.sagepub.com/doi/10.1177/1356389020969712

16. Grell P, Ahmadi N, Blom B. "Sometimes it's really complicated!" – clients with complex needs on their encounters with specialised personal social service organisations in Sweden. Nordic Social Work Research [Internet]. 2016 Sep 1 [cited 2023 Dec 5];6(3):188–200. Available from: https://doi.org/10.1080/2156857X.2016.1156017

17. MacKillop E, Sheard S. Quantifying life: Understanding the history of Quality-Adjusted Life-Years (QALYs). Soc Sci Med [Internet]. 2018 Aug 1 [cited 2020 Oct 26];211:359–66. Available from: http://www.sciencedirect.com/science/article/pii/S0277953618303605

18. Rand LZ, Kesselheim AS. An International Review of Health Technology Assessment Approaches to Prescription Drugs and Their Ethical Principles. Journal of Law, Medicine & Ethics [Internet]. 2020 Oct [cited 2023 Nov 28];48(3):583–94. Available from: https://doi.org/10.1177/1073110520958885

19. Rios-Diaz AJ, Lam J, Ramos MS, Moscoso AV, Vaughn P, Zogg CK, et al. Global Patterns of QALY and DALY Use in Surgical Cost-Utility Analyses: A Systematic Review. PLOS ONE [Internet]. 2016 Feb 10 [cited 2023 Nov 28];11(2):e0148304. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0148304

20. Mott DJ, Schirrmacher H, Al-Janabi H, Guest S, Pennington B, Scheuer N, et al. Modelling Spillover Effects on Informal Carers: The Carer QALY Trap. Pharmacoeconomics [Internet]. 2023 [cited 2024 Jan 4];41(12):1557–61. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10635951/

21. Lindholm L. Health economic evaluation of community-based cardiovascular disease prevention : Some theoretical aspects and empirical results. 1996 [cited 2023 Nov 28]; Available from: https://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-7539

22. Dolan P. Modeling valuations for EuroQol health states. Med Care. 1997 Nov;35(11):1095–108.

23. Furlong WJ, Feeny DH, Torrance GW, Barr RD. The Health Utilities Index (HUI®) system for assessing health-related quality of life in clinical studies. Ann Med [Internet]. 2001 Jan 1 [cited 2020 Oct 27];33(5):375–84. Available from: https://doi.org/10.3109/07853890109002092

24. Torrance GW. Social preferences for health states: An empirical evaluation of three measurement techniques. Socio-Economic Planning Sciences [Internet]. 1976 Jan 1 [cited 2022 Oct 5];10(3):129–36. Available from: https://www.sciencedirect.com/science/article/pii/0038012176900367

25. Clark MD, Determann D, Petrou S, Moro D, de Bekker-Grob EW. Discrete Choice Experiments in Health Economics: A Review of the Literature. PharmacoEconomics [Internet]. 2014 Sep [cited 2022 Nov 19];32(9):883–902. Available from: http://link.springer.com/10.1007/s40273-014-0170-x

26. Salomon JA, Vos T, Hogan DR, Gagnon M, Naghavi M, Mokdad A, et al. Common values in assessing health outcomes from disease and injury: Disability weights measurement study for the Global Burden of Disease Study 2010. The Lancet [Internet]. 2012 Dec 15 [cited 2021 Sep 27];380(9859):2129–43. Available from: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(12)61680-8/fulltext

27. Borsboom D, Mellenbergh GJ, van Heerden J. The theoretical status of latent variables. Psychological Review. 2003;110(2):203–19.

28. Lugnér AK, Krabbe PFM. An overview of the time trade-off method: Concept, foundation, and the evaluation of distorting factors in putting a value on health. Expert Review of Pharmacoeconomics & Outcomes Research [Internet]. 2020 Jul 3 [cited 2022 Oct 24];20(4):331–42. Available from: https://doi.org/10.1080/14737167.2020.1779062

29. Herdman M, Gudex C, Lloyd A, Janssen Mf, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res [Internet]. 2011 Dec [cited 2022 Aug 10];20(10):1727–36. Available from: http://link.springer.com/10.1007/s11136-011-9903-x

30. Stolk E, Ludwig K, Rand K, van Hout B, Ramos-Goñi JM. Overview, Update, and Lessons Learned From the International EQ-5D-5L Valuation Work: Version 2 of the EQ-5D-5L Valuation Protocol. Value Health [Internet]. 2019 Jan [cited 2019 Mar 22];22(1):23–30. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1098301518361643

31. Stolk E, Ramos-Goñi JM, Ludwig K, Oppe M, Norman R. The Development and Strengthening of Methods for Valuing EQ-5D-5L – An Overview. In: Devlin N, Roudijk B, Ludwig K, editors. Value Sets for EQ-5D-5L: A Compendium, Comparative Review & User Guide [Internet]. Cham: Springer International Publishing; 2022 [cited 2023 Jun 12]. p. 13–27. Available from: https://doi.org/10.1007/978-3-030-89289-0_2

32. Ramos-Goñi JM, Pinto-Prades JL, Oppe M, Cabasés JM, Serrano-Aguilar P, Rivero-Arias O. Valuation and Modeling of EQ-5D-5L Health States Using a Hybrid Approach. Med Care [Internet]. 2017 Jul [cited 2018 Sep 11];55(7):e51. Available from: https://journals.lww.com/lww-medicalcare/fulltext/2017/07000/Valuation_and_Modeling_of_EQ_5D_5L_Health_States.13.aspx

33. Ministry of Finance] F [Swedish. Kommunallag 2017:725 [Municipality Act] [Internet]. 2017:725, 2017:725 2017. Available from: https://lagen.nu/2017:725

34. Sweden S. Costs and incomes for municipalities by region and activity. Year 2011 - 2022 [Internet]. Statistikdatabasen; 2022 [cited 2023 Nov 28]. Available from: http://www.statistikdatabasen.scb.se/pxweb/en/ssd/START__OE__OE0107__OE0107B/KostnDR/

35. Sweden S. Expenses and revenues for regions by activity. Year 2007 - 2022 [Internet]. Statistikdatabasen; 2022 [cited 2023 Nov 28]. Available from: http://www.statistikdatabasen.scb.se/pxweb/en/ssd/START__OE__OE0107__OE0107D/KostnDRLT/

36. Health and Social Affairs] S [Ministry of. Lag om stöd och service till vissa funktionshindrade (LSS) 1993:387 [Law regulating Support and Service to Persons with Certain Functional Disabilities] [Internet]. 1993:387 1993. Available from: https://lagen.nu/1993:387

37. Health and Social Affairs] S [Ministry of. Socialförsäkringsbalk SFB 2010:110 [Social Insurance Code] [Internet]. 2010. Available from: https://lagen.nu/2010:110

38. Krevers B, Bäckman K. Modell för prioriteringsarbete i kommuner : Processer och verktyg vid resursfördelning i en övergripande ledningsprocess [Internet]. Linköping University Electronic Press; 2023 [cited 2023 Dec 16]. Available from: https://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-198143

39. Health and Social Affairs M of. Social service act [Internet]. 2001:453 2001. Available from: https://lagen.nu/2001:453

40. Baujard A. Welfare economics [Internet]. 2013 [cited 2023 Dec 24]. Available from: https://shs.hal.science/halshs-00906907

41. Dictionary C. Social welfare [Internet]. 2023 [cited 2023 Nov 28]. Available from: https://www.collinsdictionary.com/dictionary/english/social-welfare

42. International Health Conference. Constitution of the World Health Organization. 1946. Bull World Health Organ. 2002;80(12):983–4.

43. Salvador-Carulla L, Lucas R, Ayuso-Mateos JL, Miret M. Use of the terms ”Wellbeing” and ”Quality of Life” in health sciences: A conceptual framework. The European Journal of Psychiatry [Internet]. 2014 Mar [cited 2023 Oct 28];28(1):50–65. Available from: https://scielo.isciii.es/scielo.php?script=sci_abstract&pid=S0213-61632014000100005&lng=es&nrm=iso&tlng=en

44. Upton D, Upton P. Quality of Life and Well-Being. In: Upton D, Upton P, editors. Psychology of Wounds and Wound Care in Clinical Practice [Internet]. Cham: Springer International Publishing; 2015 [cited 2023 Oct 28]. p. 85–111. Available from: https://doi.org/10.1007/978-3-319-09653-7_4

45. Urbina DA, Ruiz-Villaverde A. A Critical Review of Homo Economicus from Five Approaches. The American Journal of Economics and Sociology [Internet]. 2019 [cited 2023 Oct 28];78(1):63–93. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/ajes.12258

46. Driver J. The History of Utilitarianism. In: Zalta EN, Nodelman U, editors. The Stanford Encyclopedia of Philosophy [Internet]. Winter 2022. Metaphysics Research Lab, Stanford University; 2022 [cited 2023 Oct 28]. Available from: https://plato.stanford.edu/archives/win2022/entries/utilitarianism-history/

47. Kymlicka W. Contemporary Political Philosophy: An Introduction. Second Edition. Oxford, New York: Oxford University Press; 2001. 512 p.

48. Sen A. Utilitarianism and Welfarism. The Journal of Philosophy [Internet]. 1979 [cited 2023 Aug 18];76(9):463–89. Available from: https://www.jstor.org/stable/2025934

49. Sakowsky RA. Disentangling the welfarism/extra-welfarism distinction: Towards a more fine-grained categorization. Health Economics [Internet]. 2021 [cited 2023 Aug 31];30(9):2307–11. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/hec.4382

50. Brouwer WBF, Culyer AJ, van Exel NJA, Rutten FFH. Welfarism vs. Extra-welfarism. J Health Econ [Internet]. 2008 Mar [cited 2019 Feb 27];27(2):325–38. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0167629607000902

51. Coast J, Smith RD, Lorgelly P. Welfarism, extra-welfarism and capability: The spread of ideas in health economics. Soc Sci Med. 2008 Oct;67(7):1190–8.

52. Cookson R. QALYs and the capability approach. Health Econ [Internet]. 2005 Aug [cited 2020 Jan 20];14(8):817–29. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/hec.975

53. Sen A. Commodities and Capabilities. Amsterdam: North-Holland; 1985.

54. Sen A. Equality of What? In: McMurrin SM, editor. The Tanner lectures on human values. 1st ed. Cambridge: Univ. Press; 1980. p. 195–220.

55. Sen A. Inequality reexamined. New York; Cambridge, MA: Russell Sage Foundation ; Harvard University Press; 1992.

56. Nussbaum MC. Women and Human Development: The Capabilities Approach. New York: Cambridge University Press; 2000.

57. Sen A. Capabilities, Lists, and Public Reason: Continuing the Conversation. Fem Econ [Internet]. 2004 Nov 1 [cited 2020 Oct 2];10(3):77–80. Available from: https://doi.org/10.1080/1354570042000315163

58. Månsdotter A, Ekman B, Feldman I, Hagberg L, Hurtig AK, Lindholm L. We Propose a Novel Measure for Social Welfare and Public Health: Capability-Adjusted Life-Years, CALYs. Appl Health Econ Health Policy [Internet]. 2017 Aug 1 [cited 2018 Sep 11];15(4):437–40. Available from: https://doi.org/10.1007/s40258-017-0323-0

59. Lorgelly P, Lorimer K, Fenwick E, Briggs A. The Capability Approach: Developing an instrument for evaluating public health interventions [Internet]. University of Glasgow; 2008 Aug [cited 2020 Jun 8] p. 59. Available from: https://www.gcph.co.uk/assets/0000/0426/Capabilities_-_executive_summary__August_2008.pdf

60. Schlander M, Garattini S, Holm S, Kolominsky-Rabas P, Nord E, Persson U, et al. Incremental cost per quality-adjusted life year gained? The need for alternative methods to evaluate medical interventions for ultra-rare disorders. J Compar Effect Res [Internet]. 2014 Jul [cited 2023 Oct 28];3(4):399–422. Available from: https://becarispublishing.com/doi/10.2217/cer.14.34

61. Walters K, Frost R, Kharicha K, Avgerinou C, Gardner B, Ricciardi F, et al. Home-based health promotion for older people with mild frailty: The HomeHealth intervention development and feasibility RCT. Health Technol Assess. 2017 Dec;21(73):1–128.

62. Månsdotter A, Ekman B, Feldman I, Hagberg L, Hurtig AK, Lindholm L. Towards capability-adjusted life-years (CALYs) in public health and social welfare: Results from a pilot study on ranking capabilities (in preparation). 2017.

63. Finance SM of. Får vi det bättre? om mått på livskvalitet: betänkande [Do we get it better? about quality of life measures: report]. Stockholm: Fritze; 2015. 151 p. (Statens offentliga utredningar).

64. de Meyrick J. The Delphi method and health research. Health Education [Internet]. 2003 Jan 1 [cited 2020 Oct 27];103(1):7–16. Available from: https://doi.org/10.1108/09654280310459112

65. Williams A. Intergenerational equity: An exploration of the 'fair innings' argument. Health Econ [Internet]. 1997 Mar 1 [cited 2018 Oct 9];6(2):117–32. Available from: https://pubmed.ncbi.nlm.nih.gov/9158965/

66. Mannheimer LN, Lehto J, Östlin P. Window of opportunity for intersectoral health policy in Sweden — open, half-open or half-shut? Health Promotion International [Internet]. 2007 [cited 2023 Dec 5];22(4):307–15. Available from: https://www.jstor.org/stable/45152942

67. Synnevåg ES, Amdam R, Fosse E. Intersectoral Planning for Public Health: Dilemmas and Challenges. Int J Health Policy Manag [Internet]. 2018 Jul 1 [cited 2023 Dec 5];7(11):982–92. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6326631/

68. Stiglitz J, Sen A, Fitoussi J. Report of the Commission on the Measurement of Economic Performance and Social Progress (CMEPSP). 2009 Jan 1;

69. Gosepath S. Equality. In: Zalta EN, editor. The Stanford Encyclopedia of Philosophy [Internet]. Summer 2021. Metaphysics Research Lab, Stanford University; 2021 [cited 2023 Oct 28]. Available from: https://plato.stanford.edu/archives/sum2021/entries/equality/

70. Forslund A. Om sänkta arbetsgivaravgifter för ungdomar [About reduced social security contributions for the young] [Internet]. 2023 [cited 2023 Nov 18]. Available from: https://www.ifau.se/Press/Forskningssammanfattningar/om-sankta-arbetsgivaravgifter-for-ungdomar/

71. Saez E, Schoefer B, Seim D. Hysteresis from employer subsidies. Journal of Public Economics [Internet]. 2021 Aug 1 [cited 2023 Aug 16];200:104459. Available from: https://www.sciencedirect.com/science/article/pii/S0047272721000955

72. Socialavgiftslag. Socialavgiftslag [Act on Social Fees]. 2007:284 2007.

73. Forslund A, Liljeberg L. Unga som varken arbetar eller studerar [Youth not in employment, education, or training] [Internet]. IFAU; 2021 [cited 2023 Nov 18]. Report No.: 2021:12. Available from: https://www.ifau.se/globalassets/pdf/se/2021/r-2021-12-unga-som-varken-arbetar-eller-studerar.pdf

74. Furlong A. Not a very NEET solution: Representing problematic labour market transitions among early school-leavers. Work, Employment and Society [Internet]. 2006 Sep 1 [cited 2023 Nov 18];20(3):553–69. Available from: https://doi.org/10.1177/0950017006067001

75. Youth and Civil Society SA for. Swedish Agency for Youth and Civil Society [Internet]. 2023 [cited 2023 Dec 7]. Available from: https://www.mucf.se/

76. Sweden S. Price Base Amount [Internet]. Statistiska Centralbyrån; 2023 [cited 2023 Dec 7]. Available from: https://www.scb.se/en/finding-statistics/statistics-by-subject-area/prices-and-consumption/consumer-price-index/consumer-price-index-cpi/pong/tables-and-graphs/price-basic-amount/price-base-amount/

77. MUCF. Statistik om unga som varken arbetar eller studerar [Statistics on youth that neither work nor study] [Internet]. 2022 [cited 2023 Nov 18]. Available from: https://www.mucf.se/verktyg/statistik-om-unga-som-varken-arbetar-eller-studerar

78. Ruesga-Benito SM, González-Laxe F, Picatoste X. Sustainable Development, Poverty, and Risk of Exclusion for Young People in the European Union: The Case of NEETs. Sustainability [Internet]. 2018 Dec [cited 2023 Dec 7];10(12, 12):4708. Available from: https://www.mdpi.com/2071-1050/10/12/4708

79. Salvà-Mut F, Tugores-Ques M, Quintana-Murci E. NEETs in Spain: An analysis in a context of economic crisis. International Journal of Lifelong Education [Internet]. 2018 Mar 4 [cited 2023 Dec 7];37(2):168–83. Available from: https://doi.org/10.1080/02601370.2017.1382016

80. Rahmani H, Groot W. Risk Factors of Being a Youth Not in Education, Employment or Training (NEET): A Scoping Review. International Journal of Educational Research [Internet]. 2023 Jan 1 [cited 2023 Nov 18];120:102198. Available from: https://www.sciencedirect.com/science/article/pii/S0883035523000629

81. MacDonald R. Youth transitions, unemployment and underemployment: Plus ça change, plus c'est la même chose? Journal of Sociology [Internet]. 2011 Dec 1 [cited 2023 Nov 18];47(4):427–44. Available from: https://doi.org/10.1177/1440783311420794

82. Helbling LA, Sacchi S. Scarring effects of early unemployment among young workers with vocational credentials in Switzerland. Empirical Research in Vocational Education and Training [Internet]. 2014 Dec 16 [cited 2023 Nov 18];6(1):12. Available from: https://doi.org/10.1186/s40461-014-0012-2

83. Ralston K, Everington D, Feng Z, Dibben C. Economic Inactivity, Not in Employment, Education or Training (NEET) and Scarring: The Importance of NEET as a Marker of Long-Term Disadvantage. Work, Employment and Society [Internet]. 2022 Feb 1 [cited 2023 Nov 18];36(1):59–79. Available from: https://doi.org/10.1177/0950017020973882

84. Thompson R, Russell L, Simmons R. Space, place and social exclusion: An ethnographic study of young people outside education and employment. Journal of Youth Studies [Internet]. 2014 Jan 2 [cited 2023 Nov 18];17(1):63–78. Available from: https://doi.org/10.1080/13676261.2013.793793

85. Robertson PJ. The casualties of transition: The health impact of NEET status and some approaches to managing it. British Journal of Guidance & Counselling [Internet]. 2019 May 4 [cited 2023 Dec 7];47(3):390–402. Available from: https://doi.org/10.1080/03069885.2018.1455168

86. Strandh M, Winefield A, Nilsson K, Hammarström A. Unemployment and mental health scarring during the life course. European Journal of Public Health [Internet]. 2014 Jun 1 [cited 2023 Aug 27];24(3):440–5. Available from: https://doi.org/10.1093/eurpub/cku005

87. Andersson FW, Gullberg Brännstrom S, Mörtvik R. Long-term scarring effect of neither working nor studying. International Journal of Manpower [Internet]. 2018 Jan 1 [cited 2023 Dec 12];39(2):190–204. Available from: https://doi.org/10.1108/IJM-12-2015-0226

88. Al-Janabi H, Flynn TN, Coast J. Development of a self-report measure of capability wellbeing for adults: The ICECAP-A. Qual Life Res [Internet]. 2012 Feb [cited 2019 Mar 15];21(1):167–76. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3254872/

89. Canaway A, Al-Janabi H, Kinghorn P, Bailey C, Coast J. Development of a measure (ICECAP-Close Person Measure) through qualitative methods to capture the benefits of end-of-life care to those close to the dying for use in economic evaluation. Palliat Med [Internet]. 2017 Jan 1 [cited 2023 Sep 12];31(1):53–62. Available from: https://doi.org/10.1177/0269216316650616

90. Coast J, Huynh E, Kinghorn P, Flynn T. Complex Valuation: Applying Ideas from the Complex Intervention Framework to Valuation of a New Measure for End-of-Life Care. Pharmacoeconomics [Internet]. 2016 May 1 [cited 2020 May 12];34(5):499–508. Available from: https://doi.org/10.1007/s40273-015-0365-9

91. Coast J, Flynn TN, Natarajan L, Sproston K, Lewis J, Louviere JJ, et al. Valuing the ICECAP capability index for older people. Social Science & Medicine [Internet]. 2008 Sep 1 [cited 2021 Mar 4];67(5):874–82. Available from: https://www.sciencedirect.com/science/article/pii/S0277953608002542

92. Flynn TN, Huynh E, Peters TJ, Al-Janabi H, Clemens S, Moody A, et al. Scoring the Icecap-a capability instrument. Estimation of a UK general population tariff. Health Econ. 2015 Mar;24(3):258–69.

93. Grewal I, Lewis J, Flynn T, Brown J, Bond J, Coast J. Developing attributes for a generic quality of life measure for older people: Preferences or capabilities? Soc Sci Med [Internet]. 2006 Apr 1 [cited 2019 Mar 13];62(8):1891–901. Available from: http://www.sciencedirect.com/science/article/pii/S0277953605004454

94. Husbands S, Mitchell PM, Floredin I, Peters TJ, Kinghorn P, Byford S, et al. The Children and Young People Quality of Life Study: A protocol for the qualitative development of attributes for capability wellbeing measures for use in health economic evaluation with children and young people. Wellcome Open Res [Internet]. 2022 Mar 30 [cited 2023 Sep 12];7:117. Available from: https://wellcomeopenresearch.org/articles/7-117/v1

95. Huynh E, Coast J, Rose J, Kinghorn P, Flynn T. Values for the ICECAP-Supportive Care Measure (ICECAP-SCM) for use in economic evaluation at end of life. Soc Sci Med [Internet]. 2017 Sep 1 [cited 2020 May 12];189:114–28. Available from: http://www.sciencedirect.com/science/article/pii/S0277953617304458

96. Anand P, Hunter G, Carter I, Dowding K, Guala F, Van Hees M. The Development of Capability Indicators. J Human Dev Capabil [Internet]. 2009 Mar 1 [cited 2019 Mar 15];10(1):125–52. Available from: https://doi.org/10.1080/14649880802675366

97. Lorgelly PK, Lorimer K, Fenwick EAL, Briggs AH, Anand P. Operationalising the capability approach as an outcome measure in public health: The development of the OCAP-18. Soc Sci Med [Internet]. 2015 Oct 1 [cited 2019 Mar 15];142:68–81. Available from: http://www.sciencedirect.com/science/article/pii/S027795361530054X

98. Simon J, Anand P, Gray A, Rugkåsa J, Yeeles K, Burns T. Operationalising the capability approach for outcome measurement in mental health research. Soc Sci Med [Internet]. 2013 Dec 1 [cited 2019 Mar 15];98:187–96. Available from: http://www.sciencedirect.com/science/article/pii/S027795361300525X

99. Netten A, Burge P, Malley J, Potoglou D, Towers A, Brazier J, et al. Outcomes of social care for adults: Developing a preference weighted measure. Health Technol Assess. 2012 Apr 2;16(16).

100. Anand P, van Hees M. Capabilities and achievements: An empirical study. The Journal of Socio-Economics [Internet]. 2006 Apr 1 [cited 2022 Oct 10];35(2):268–84. Available from: https://www.sciencedirect.com/science/article/pii/S1053535705001083

101. Hofmann K, Schori D, Abel T. Self-Reported Capabilities Among Young Male Adults in Switzerland: Translation and Psychometric Evaluation of a German, French and Italian Version of a Closed Survey Instrument. Soc Indic Res [Internet]. 2013 Nov 1 [cited 2022 Oct 10];114(2):723–38. Available from: https://doi.org/10.1007/s11205-012-0170-1

102. Helter TM, Coast J, Łaszewska A, Stamm T, Simon J. Capability instruments in economic evaluations of health-related interventions: A comparative review of the literature. Qual Life Res. 2019 Dec 24;

103. Ubels J, Hernandez-Villafuerte K, Schlander M. The Value of Freedom: A Review of the Current Developments and Conceptual Issues in the Measurement of Capability. J Human Dev Capabil [Internet]. 2022 Mar 19 [cited 2022 Mar 21];0(0):1–27. Available from: https://doi.org/10.1080/19452829.2022.2053506

104. Brazier J, Peasgood T, Mukuria C, Marten O, Kreimeier S, Luo N, et al. The EQ-HWB: Overview of the Development of a Measure of Health and Wellbeing and Key Results. Value in Health [Internet]. 2022 Apr [cited 2022 Oct 14];25(4):482–91. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1098301522000833

105. Veldwijk J, Marceta SM, Swait JD, Lipman SA, de Bekker-Grob EW. Taking the Shortcut: Simplifying Heuristics in Discrete Choice Experiments. Patient [Internet]. 2023 Jul 1 [cited 2023 Dec 1];16(4):301–15. Available from: https://doi.org/10.1007/s40271-023-00625-y

106. Daniels N. Just Health: Meeting Health Needs Fairly. Cambridge: Cambridge University Press; 2007.

107. United Nations U. Transforming Our World, the 2030 Agenda for Sustainable Development. A/RES/70/1 2015.

108. Health and Social Affairs M of. God och jämlik hälsa - en utvecklad folkhälsopolitik [God and equal health - a developped public health policy]. Prop. 2017/18:249 2017.

109. Health and Social Affairs M of. Hälso- och sjukvårdslag [Health and Medical Services Act] [Internet]. Apr 1, 2017. Available from: https://lagen.nu/2017:30

110. Kungörelse. Kungörelse 1974:152 Regereingsformen [Instrument of governement] [Internet]. Kapitel 2 1974. Available from: https://lagen.nu/1974:152

111. Hardin R, Cullity G. The Free Rider Problem. In: Zalta EN, editor. The Stanford Encyclopedia of Philosophy [Internet]. Winter 2020. Metaphysics Research Lab, Stanford University; 2020 [cited 2020 Nov 13]. Available from: https://plato.stanford.edu/archives/win2020/entries/free-rider/

112. VanderWeele TJ. On the promotion of human flourishing. PNAS [Internet]. 2017 Aug 1 [cited 2020 Jun 14];114(31):8148–56. Available from: https://www.pnas.org/content/114/31/8148

113. Meili KW, Mulhern B, Ssegonja R, Hjelte J, Norström F, Feldman I, et al. CALY-SWE value set: An integrated approach for a valuation study based on an online-administered TTO and DCE survey [Internet]. Zenodo; 2023 Apr [cited 2023 Jul 4]. Available from: https://zenodo.org/record/7870705

114. Finch AP, Meregaglia M, Ciani O, Roudijk B, Jommi C. An EQ-5D-5L value set for Italy using videoconferencing interviews and feasibility of a new mode of administration. Social Science & Medicine [Internet]. 2022 Jan 1 [cited 2022 Mar 17];292:114519. Available from: https://www.sciencedirect.com/science/article/pii/S0277953621008510

115. Sun S, Chuang LH, Sahlén KG, Lindholm L, Norström F. Estimating a social value set for EQ-5D-5L in Sweden. Health and Quality of Life Outcomes [Internet]. 2022 Dec 23 [cited 2023 Feb 24];20(1):167. Available from: https://doi.org/10.1186/s12955-022-02083-w

116. Ramos-Goñi JM, Oppe M, Slaap B, Busschbach JJV, Stolk E. Quality Control Process for EQ-5D-5L Valuation Studies. Value in Health [Internet]. 2017 Mar [cited 2022 Mar 17];20(3):466–73. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1098301516340669

117. Meili K, Lindholm L. In Pursuit of Weights For CALY : Exploring Methods for Measuring And Calculating Capability Adjusted Life Year Weights [Internet] [M.Sc. Thesis]. [Umeå]: Umeå University; 2017 [cited 2017 Dec 18]. Available from: http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-142729

118. Meili KW, Månsdotter A, Lindholm L. Time Trade-Off Iteration for CALY Weights: A Pilot Study (unpublished). 2018.

119. Group TP. PHP: Hypertext Preprocessor [Internet]. 2022 [cited 2022 Sep 11]. Available from: http://php.net/

120. Foundation MD. MariaDB [Internet]. 2023 [cited 2023 Oct 28]. Available from: https://mariadb.org/

121. Hernández-Nieto C, Sánchez X, Salinas P. A Mobile First Approach for the Development of a Sustainability Game. Procedia Computer Science [Internet]. 2015 [cited 2023 Dec 3];75:182–5. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1877050915036972

122. Montgomery DC. Design and Analysis of Experiments [Internet]. Wiley; 2012. 752 p. Available from: https://books.google.com?id=KSRZpwAACAAJ

123. Morgan-Wall T, Khoury G. Optimal Design Generation and Power Evaluation R: The skpr Package. J Stat Soft [Internet]. 2021 [cited 2021 Sep 27];99(1). Available from: http://www.jstatsoft.org/v99/i01/

124. Street DJ, Burgess L, Louviere JJ. Quick and easy choice sets: Constructing optimal and nearly optimal stated choice experiments. Int J Res Mark [Internet]. 2005 Dec [cited 2018 Sep 11];22(4):459–70. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0167811605000510

125. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian Data Analysis [Internet]. 0th ed. Chapman and Hall/CRC; 2013 [cited 2023 Dec 23]. Available from: https://www.taylorfrancis.com/books/9781439898208

126. McElreath R. Statistical Rethinking: A Bayesian Course with Examples in R and Stan. 1 edition. Boca Raton: Chapman and Hall/CRC; 2015. 487 p.

127. Team SD. Stan [Internet]. 2022. Available from: https://mc-stan.org

128. Gabry J, Češnovar R, Bales B, Morris M, Popov M, Lawrence M, et al. Cmdstanr [Internet]. 2022 [cited 2022 Aug 11]. Available from: https://mc-stan.org/cmdstanr

129. Brussels: Belgian Health Care Knowledge Centre (KCE) HSR (HSR). An EQ -5D-5L value set for Belgium – How to value health-related quality of life? Brussels; 2021 p. 110.

130. Bijlenga D, Birnie E, Bonsel GJ. Feasibility, Reliability, and Validity of Three Health-State Valuation Methods Using Multiple-Outcome Vignettes on Moderate-Risk Pregnancy at Term. Value in Health [Internet]. 2009 Jul [cited 2022 Sep 17];12(5):821–7. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1098301510607450

131. Bartley M. Health inequality: An introduction to concepts, theories and methods. Second edition. Cambridge, UK ; Malden, MA, USA: Polity; 2017. 244 p.

132. Organization PAH. Methodological Summaries: Measuring Inequity in Health. Epidemiological Bulletin [Internet]. 1999 Mar [cited 2023 Dec 1];20(1). Available from: https://iris.paho.org/handle/10665.2/56950

133. Sweden S. Population by age and sex. Year 1860 - 2022. [Internet]. Statistikdatabasen; 2022 [cited 2023 Dec 10]. Available from: https://www.statistikdatabasen.scb.se/pxweb/en/ssd/START__BE__BE0101__BE0101A/BefolkningR1860N/

134. Sweden S. Population aged 15-74 (LFS) by sex, age and labour status. Year 2001 - 2022 [Internet]. Statistikdatabasen; 2020 [cited 2023 Nov 18]. Available from: http://www.statistikdatabasen.scb.se/pxweb/en/ssd/START__AM__AM0401__AM0401A/NAKUBefolkning2Ar/

135. Sweden S. Population aged 15-74 (LFS) by sex, age and labour status, previous definitions. Year 1970 - 2020 [Internet]. Statistikdatabasen; 2020 [cited 2023 Nov 18]. Available from: https://www.statistikdatabasen.scb.se/pxweb/en/ssd/START__AM__AM0403__AM0403A/NAKUBefolkning2ArTD/

136. Sweden S. Income distribution (fractiles) for persons registered in the national population register during the whole year by region, sex and age. Year 2010 - 2021 [Internet]. Statistikdatabasen; 2021 [cited 2023 Nov 18]. Available from: http://www.statistikdatabasen.scb.se/pxweb/en/ssd/START__HE__HE0110__HE0110A/SamNetFrakt/

137. Sweden S. Average salary and salary dispersion by sector, occupation (SSYK 2012) and sex . Year 2014 - 2022 [Internet]. Statistikdatabasen; 2022 [cited 2023 Nov 18]. Available from: http://www.statistikdatabasen.scb.se/pxweb/en/ssd/START__AM__AM0110__AM0110A/LoneSpridSektorYrk4A/

138. OECD. Conversion rates - Purchasing power parities (PPP) - OECD Data [Internet]. theOECD; 2023 [cited 2023 Dec 13]. Available from: http://data.oecd.org/conversion/purchasing-power-parities-ppp.htm

139. Sweden S. Local taxes [Internet]. 2023 [cited 2023 Dec 14]. Available from: https://www.scb.se/en/finding-statistics/statistics-by-subject-area/public-finances/local-government-finances/local-taxes/

140. Baroni L. Lägre kostnader för senior arbetskraft [Lower costs for senior employees]. 2020. Report No.: 2018:10.

141. PricewaterhouseCoopers. Vad kostar en anställd? [Internet]. Vad kostar en anställd? [What does an employee cost?]; 2016 [cited 2023 Dec 12]. Available from: https://blogg.pwc.se/foretagarbloggen/sa-mycket-kostar-en-anstalld

142. and Welfare NB of health. Nationella riktlinjer - Metodbeskrivning [National guidelines - Methodods description] [Internet]. National Board of health and Welfare; 2015. Available from: https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/dokument-webb/nationella-riktlinjer/nationella-riktlinjer-metodbeskrivning.pdf

143. Sweden S. Statistics Sweden [Internet]. 2023 [cited 2018 Oct 14]. Available from: https://www.scb.se

144. Cint. Cint [Internet]. 2020 [cited 2020 Nov 6]. Available from: https://www.cint.com

145. LimeSurvey. LimeSurvey: An Open Source survey tool [Internet]. Hamburg, Germany: Limesurvey GmbH; 2022 [cited 2020 Nov 6]. Available from: http://www.limesurvey.org

146. Authority SER. Etikprövningsmyndigheten [Swedish Ethical Review Authority] [Internet]. Etikprövningsmyndigheten; 2023 [cited 2023 Oct 28]. Available from: https://etikprovningsmyndigheten.se/en/

147. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2023 [cited 2020 Nov 6]. Available from: https://www.R-project.org/

148. OECD. How's Life? 2013: Measuring Well-being [Internet]. OECD; 2013 [cited 2020 Jul 27]. (How's Life?). Available from: https://www.oecd-ilibrary.org/economics/how-s-life-2013_9789264201392-en

149. Johansson S, Erikson R, Jonsson JO, Tåhlin M. LNU68 - Swedish level of living survey 1968. Stockholm University; 1968.

150. Månsdotter A, Ekman B, Meili KW, Feldman I, Hagberg L, Hurtig AK, et al. Towards capability-adjusted life years in public health and social welfare: Results from a Swedish survey on ranking capabilities. PLoS One. 2020;15(12):e0242699.

151. Shiroiwa T, Fukuda T, Ikeda S, Shimozuma K. QALY and Productivity Loss: Empirical Evidence for "Double Counting." Value in Health [Internet]. 2013 Jun 1 [cited 2020 Jan 20];16(4):581–7. Available from: http://www.sciencedirect.com/science/article/pii/S1098301513000648

152. Sedlacek T, Havel F by V. Economics of Good and Evil: The Quest for Economic Meaning from Gilgamesh to Wall Street. Oxford, New York: Oxford University Press; 2011. 376 p.

153. Aronsson M, Husberg M, Kalkan A, Eckard N, Alwin J. Differences between hypothetical and experience-based value sets for EQ-5D used in Sweden: Implications for decision makers. Scand J Public Health [Internet]. 2015 Dec [cited 2023 Oct 27];43(8):848–54. Available from: http://journals.sagepub.com/doi/10.1177/1403494815596910

154. Burström K, Teni FS, Gerdtham U-G, Leidl R, Helgesson G, Rolfson O, et al. Experience-Based Swedish TTO and VAS Value Sets for EQ-5D-5L Health States. PharmacoEconomics [Internet]. 2020 Aug [cited 2022 May 5];38(8):839–56. Available from: https://link.springer.com/10.1007/s40273-020-00905-7

155. Schneider P. The QALY is ableist: On the unethical implications of health states worse than dead. Qual Life Res [Internet]. 2022 May 1 [cited 2023 Oct 27];31(5):1545–52. Available from: https://doi.org/10.1007/s11136-021-03052-4

156. Gandhi M, Rand K, Luo N. Valuation of Health States Considered to Be Worse Than Death—An Analysis of Composite Time Trade-Off Data From 5 EQ-5D-5L Valuation Studies. Value in Health [Internet]. 2019 Mar 1 [cited 2023 Jul 5];22(3):370–6. Available from: https://www.sciencedirect.com/science/article/pii/S1098301518361953

157. Jakubczyk M. Re-revisiting the Utilities of Health States Worse than Dead: The Problem Remains. Med Decis Making [Internet]. 2023 Oct 1 [cited 2023 Nov 29];43(7-8):875–85. Available from: https://doi.org/10.1177/0272989X231201147

158. Tilling C, Devlin N, Tsuchiya A, Buckingham K. Protocols for Time Tradeoff Valuations of Health States Worse than Dead: A Literature Review. Med Decis Making [Internet]. 2010 Sep [cited 2023 Nov 29];30(5):610–9. Available from: http://journals.sagepub.com/doi/10.1177/0272989X09357475

159. Lindholm L, Rosén M. On the measurement of the nation's equity adjusted health. Health Econ [Internet]. 1998 Nov 1 [cited 2018 Oct 9];7(7):621–8. Available from: https://pubmed.ncbi.nlm.nih.gov/9845255/

160. Cadham CJ, Prosser LA. Eliciting Trade-Offs Between Equity and Efficiency: A Methodological Scoping Review. Value in Health [Internet]. 2023 Jun 1 [cited 2023 Nov 29];26(6):943–52. Available from: https://www.valueinhealthjournal.com/article/S1098-3015(23)00055-4/fulltext

161. Cookson R, Mirelman AJ, Griffin S, Asaria M, Dawkins B, Norheim OF, et al. Using Cost-Effectiveness Analysis to Address Health Equity Concerns. Value Health [Internet]. 2017 Feb [cited 2019 May 16];20(2):206–12. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5340318/

162. Radhakrishnan A, Muir JM, Freitag A, Stillman IO, Sarri G. Efficiency-Equity Trade-offs in Cost-effectiveness Analyses; An Umbrella Review of Previous Systematic Literature Reviews. In 2023.

163. Elster J. Sour grapes – utilitarianism and the genesis of wants. In: Sen A, Williams B, editors. Utilitarianism and Beyond [Internet]. Cambridge: Cambridge University Press; 1982 [cited 2023 Dec 5]. p. 219–38. Available from: https://www.cambridge.org/core/books/utilitarianism-and-beyond/sour-grapes-utilitarianism-and-the-genesis-of-wants/A0E4BFE0600CE2C6DD5E61B335969AB7

164. Robeyns I. Wellbeing, Freedom and Social Justice: The Capability Approach Re-Exam ined (Robeyns) [Internet]. Utrecht University; 2021 [cited 2023 Dec 5]. Available from: https://socialsci.libretexts.org/@go/page/103063

165. Mitchell P. Adaptive Preferences, Adapted Preferences. Mind [Internet]. 2018 Oct 1 [cited 2023 Dec 5];127(508):1003–25. Available from: https://academic.oup.com/mind/article/127/508/1003/5051681

166. Månsdotter A. Health, economics, and feminism : On judging fairness and reform. 2006 [cited 2023 Nov 28]; Available from: https://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-731

167. Loewenstein G. Because It Is There: The Challenge of Mountaineering... for Utility Theory. Kyklos [Internet]. 1999 [cited 2023 Nov 29];52(3):315–43. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-6435.1999.tb00221.x

168. Netten A, Ryan M, Smith P, Skatun D, Healey A, Knapp M, et al. The Development of a Measure of Social Care Outcome for Older People. 2002 Jan 1;

169. Kinghorn P, Afentou N. Eliciting a monetary threshold for a year of sufficient capability to inform resource allocation decisions in public health and social care. Social Science & Medicine [Internet]. 2021 Jun 1 [cited 2021 Aug 25];279:113977. Available from: https://www.sciencedirect.com/science/article/pii/S0277953621003099

170. Kinghorn P. Using deliberative methods to establish a sufficient state of capability well-being for use in decision-making in the contexts of public health and social care. Social Science & Medicine [Internet]. 2019 Nov 1 [cited 2023 Dec 5];240:112546. Available from: https://www.sciencedirect.com/science/article/pii/S0277953619305404

171. Mitchell PM, Roberts TE, Barton PM, Coast J. Assessing sufficient capability: A new approach to economic evaluation. Soc Sci Med [Internet]. 2015 Aug 1 [cited 2019 Mar 13];139:71–9. Available from: http://www.sciencedirect.com/science/article/pii/S0277953615300137

172. van Hout BA, Shaw JW. Mapping EQ-5D-3L to EQ-5D-5L. Value Health. 2021 Sep;24(9):1285–93.

173. Kot SM, Paradowski PR. The atlas of inequality aversion: Theory and empirical evidence on 55 countries from the Luxembourg Income Study database. Equilibrium Quarterly Journal of Economics and Economic Policy [Internet]. 2022 Jun 25 [cited 2023 Dec 16];17(2, 2):261–316. Available from: https://journals.economic-research.pl/eq/article/view/1972

174. Esping-Andersen G. The three worlds of welfare capitalism. Princeton, N.J: Princeton University Press; 1990. 248 p.

175. Kuivalainen S, Nelson K. The Nordic welfare model in a European perspective.

176. Kuisma M. Oscillating meanings of the Nordic model: Ideas and the welfare state in Finland and Sweden. Critical Policy Studies [Internet]. 2017 Oct 2 [cited 2023 Dec 16];11(4):433–54. Available from: https://doi.org/10.1080/19460171.2016.1200475

177. OECD. Social Expenditure Database (SOCX) [Internet]. 2023 [cited 2023 Dec 16]. Available from: https://www.oecd.org/social/expenditure.htm

178. Riksdagsförvaltningen. Socialtjänstlag (2001:453) Svensk författningssamling 2001:2001:453 t.o.m. SFS 2018:1894 - Riksdagen [Internet]. 2018 [cited 2019 Mar 25]. Available from: https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/socialtjanstlag-2001453_sfs-2001-453

179. Téhard B, Detournay B, Borget I, Roze S, Pouvourville GD. Value of a QALY for France: A New Approach to Propose Acceptable Reference Values. Value in Health [Internet]. 2020 Aug 1 [cited 2023 Dec 16];23(8):985–93. Available from: https://www.valueinhealthjournal.com/article/S1098-3015(20)30188-1/fulltext

180. Palmer S, Torgerson DJ. Economics notes: Definitions of efficiency. BMJ [Internet]. 1999 Apr 24 [cited 2024 Jan 4];318(7191):1136–6. Available from: https://www.bmj.com/lookup/doi/10.1136/bmj.318.7191.1136

181. Meili KW, Jonsson H, Lindholm L, Månsdotter A. Perceived changes in capability during the COVID-19 pandemic: A Swedish cross-sectional study from June 2020. Scand J Public Health [Internet]. 2021 Jul 2 [cited 2021 Sep 22];14034948211023633. Available from: https://doi.org/10.1177/14034948211023633

182. Marmot M. Fair society, healthy lives : The Marmot Review : Strategic review of health inequalities in England post-2010 [Internet]. 2010 [cited 2023 Dec 5]. Available from: https://www.parliament.uk/globalassets/documents/fair-society-healthy-lives-full-report.pdf

183. CRediT. CRediT [Internet]. CRediT; 2023 [cited 2024 Jan 4]. Available from: https://credit.niso.org/

184. 'CALY-SWE project'. CALY Sweden [Internet]. CALY Sweden; 2022 [cited 2022 Oct 14]. Available from: https://caly.se/

Epilogue

All plans go wrong – not having a plan is wrong

Sitting in my 92' Opel Astra and listening to Elton John's Crocodile rock, never would I have imagined that I would end up doing a PhD. The year was 2015 and I was in high spirit – just having learned that I got into the Public Health Master's Programme with a Specialization in Health Economics at the Department of Epidemiology and Global Health at Umeå University.

However, I became fascinated by the department paradigm that contrasted so much with the one in my previous studies: Suddenly the people and not profit was in focus. Especially intriguing was the course of *Equity and Health*, and I realized in full consciousness what I was instinctively internalizing before: Since childhood my motto was "Saving is a virtue". But monetary richness was never the main purpose behind. Rather, the avoidance of waste to facilitate the things that truly matter.

I finished the master's and moved back to Switzerland. As the opportunity of working on the CALY-SWE project arose did I not entirely understand the idea. When I grasped the width of it though, I got fascinated and was convinced of the project's potential for society and was willing to go to great lengths to contribute – moving to Sweden away from family and friends. Working in academia comes with downsides and there are times of regret. Still, this work constitutes all in all a very rewarding and honouring experience. It revealed to me on my own example the importance of the freedom to pursue one's one aspirations and values.

Appendix 1: Additional results

Name (Cmp.)	N (Cmp.)	Mean weight (95% CI)	Cmp. weight	Weight Δ (95% CI)
Population sample	3639	0.837 (0.831, 0.842)		
Disadvantage				
Ever NEET (Population)	797 (2842)	0.762 (0.748, 0.778)	0.837	-0.075 (-0.087, -0.061)
Unemployed (Population)	167 (3472)	0.668 (0.639, 0.699)	0.837	-0.169 (-0.197, -0.137)
Poor self-rated health (Population)	102 (3537)	0.465 (0.433, 0.501)	0.837	-0.371 (-0.406, -0.335)
Long-term sick leave (Population)	87 (3552)	0.573 (0.528, 0.616)	0.837	-0.264 (-0.305, -0.223)
Payment difficulties (Population)	576 (3063)	0.676 (0.659, 0.692)	0.837	-0.16 (-0.175, -0.146)
Homeless (Population)	30(3609)	0.702 (0.633, 0.764)	0.837	-0.135 (-0.206, -0.074)
Plurality				
Women (Men)	1833 (1791)	0.829 (0.82, 0.837)	0.846	-0.017 (-0.029, -0.005)

Birthplace	147	0.833	0.837	-0.004
Global South (Other birthplace)	(3482)	(0.807, 0.858)		(-0.031, 0.02)
Lower education (Other education)	313 (3318)	0.778(0.756, 0.8)	0.842	-0.064 (-0.088, -0.039)
Not Swedish-born parents (Swedish-born parents)	488 (3128)	0.814 (0.798, 0.831)	0.84	-0.027 (-0.044, -0.009)
Rural (Urban)	285 (3246)	0.791 (0.767, 0.815)	0.841	-0.05 (-0.074, -0.025)
Informal work (Other occupations)	48 (3565)	0.799(0.744, 0.844)	0.837	-0.037 (-0.092, 0.008)
Age 18–29 (Population)	611 (3028)	0.842 (0.828, 0.856)	0.837	0.006 (-0.007, 0.018)
Age 30–44 (Population)	943 (2696)	0.827 (0.816, 0.838)	0.837	-0.009 (-0.019, 0)
Age 45–64 (Population)	1140 (2499)	0.817 (0.805, 0.829)	0.837	-0.02 (-0.029, -0.011)
Age 65–105 (Population)	945 (2694)	0.866 (0.857, 0.877)	0.837	0.03 (0.021, 0.038)

Table 7.2: Weights per group. Adapted from S3. CI (confidence interval); comparator (Cmp.); not in employment, education, or training (NEET). Δ stands for 'difference'.

Appendix 2: Timeline CALY-SWE and PhD project activities

- 2016: Publication of the CALY-SWE proposal*
- 2017: Master's thesis on piloting methods for a weight valuation study [117]
- 2018: Working paper on CALY-SWE weight anchoring
- 2018: Forte Grant
- 2019: Pilot study on measuring capability using a web panel
- 2019: Artwork for the CALY-SWE logo
- 2020: Publication of the ranking study
- 2020: The www.caly.se website went live
- 2021: Study on perceived capabilities during the COVID-19 [181]
- 2021: Data collection for a qualitative study to inform the valuation study
- 2022: Study on the development of phrasing and capability attribute selection [1]
- 2022: Data collection distribution study and value set study
- 2023: CALY-SWE video
- 2023: Presentation at the International Health Economics Association (IHEA) conference in Cape Town, South Africa
- 2023: Publication of value set
- 2024: PhD defence

For simplification, only the final date is listed, although some items on the list spanned multiple years.

*Without my involvement

Appendix 3: Contributions

This appendix outlines approximately what I contributed in the individual studies, based on the [183] taxonomy.

General: Data curation and management

S1: Contributed to conceptualization, quantitative data collection and analysis for frequency words and inequality aversion, administrative facilitation of the Delphi process, manuscript writing, manuscript revision, and data visualization.

S2: Contributed to conceptualization, pilot studies, experimental design, sample size simulations, web development, qualitative study design and data collection, quantitative data collection, methodology and formal analysis, manuscript writing, manuscript revision, and data visualization.

S3: Contributed to conceptualization, quantitative data collection, methodology and formal analysis, manuscript writing, manuscript revision, and data visualization

S4: Contributed to conceptualization, quantitative data collection, figure digitalization, methodology and formal analysis, manuscript writing, manuscript revision, and data visualization

Appendix 4: Related resources

- CALY-SWE website: `https://www.caly.se` . Features project-related resources and the CALY-SWE questionnaire [184].
- Study on ranking capabilities among Swedish Health care professionals. Data collection in 2016 [150].
- Study on perceived capabilities during Covid-19 [181].

The next appendixcontains the CALY-SWE questionnaire.

Hälsa: ”Jag har ett bra allmänt hälsotillstånd (psykiskt och fysiskt) som nästan alltid (säg minst 95% av alla dagar) tillåter mig att arbeta eller ägna mig åt det jag vill.”

- ☐ Instämmer helt
- ☐ Instämmer delvis
- ☐ Instämmer inte

Nära relationer: ”Jag har tillgång till nära relationer (familj, vänner eller bekanta) som ger mig råd och stöd när jag behöver.”

- ☐ Instämmer helt
- ☐ Instämmer delvis
- ☐ Instämmer inte

Ekonomi och boende: ”Jag har en ekonomi (lön, annan inkomst eller besparingar) som alltid tillåter mig att ha en fast bostad och för det mesta (minst 8 gånger av 10) tillåter mig att köpa det jag tycker mig behöva.”

- ☐ Instämmer helt
- ☐ Instämmer delvis
- ☐ Instämmer inte

Sysselsättning: ”Jag har ett arbete eller annan sysselsättning (studier, praktik, hemarbete, vård av anhörig, etc.) som jag för det mesta är nöjd med. De senaste fem åren har jag minst 75% av tiden varit nöjd med det jag ägnat mig åt.”

- ☐ Instämmer helt
- ☐ Instämmer delvis
- ☐ Instämmer inte

CC BY-SA 4.0 – www.caly.se

Säkerhet: ”Jag har en trygghet som gör att risk för våld eller annan kriminalitet inte påverkar mig i mitt arbete eller i min fritid.”

- ☐ Instämmer helt
- ☐ Instämmer delvis
- ☐ Instämmer inte

Politiska och medborgerliga rättigheter: "Jag litar på att våra rättigheter som medborgare, och vårt aktiva deltagande i det demokratiska systemet, gör det möjligt att påverka våra gemensamma livsvillkor."

- ☐ Instämmer helt
- ☐ Instämmer delvis
- ☐ Instämmer inte

CC BY-SA 4.0 – www.caly.se

List of Figures

1.1 Spending of the Swedish municipalities and regions in 2022. Items closely relevant for social welfare are shown shaded. Less-relevant items are shaded white with a black border. 8

2.1 Lars Lindholm at the Delphi panel day in Stockholm, 23. of January, in Stockholm. 22
2.2 Distribution of attribute-level answers per frequency for each version (A, B, or C). *Indicates a significant difference in the proportions between the versions, using a t-test for proportions with a significance level of 0.05. 26
2.3 Inequality aversion answers. 26

3.1 Overview of the weight study. Adapted from [113]. 33
3.2 Left: survey for the qualitative interviews. Middle: survey for stage 1. Right: survey for stages 2 and 3. 33
3.3 Data quality indicators by the included percentile according to the combined inconsistency score (CIS). 40
3.4 Distribution of the time trade-off (TTO) answers. “Other states” refers to all other states in the design besides 111111 and the learning state. . 41
3.5 Comparison of the anchor scale, the discrete choice experiment (DCE) logit model, and the discapability model. The bars represent the magnitude of coefficients, normalized so that the coefficient sum per model without a constant corresponds to 100%. 42

4.1 Top left: overall answer level proportions. Right: answer level proportions by attribute. Bottom left: overall weight distribution. Capability-adjusted life year (CALY). 47
4.2 Weight component differences for the disadvantage groups. Adapted from S3. Not in employment, education, or training (NEET). 50
4.3 Weight component differences for the plurality groups. Adapted from S3. 51
4.4 Weight component differences for the age groups. Adapted from S3. . 52

4.5 Capability-adjusted life year (CALY) consequences on the population scale. Adapted from S3. Not in employment, education, or training (NEET) 53

5.1 Unemployment (top), influence value (middle), and not in employment, education, or training (NEET, bottom) proportions. The factual (solid lines) and counterfactual (dashed lines) scenarios are shown. The grey vertical lines represent the duration of the payroll tax-cut intervention. Adapted from S4. 58

5.2 Cost-acceptability curves for the cautious and optimistic scenarios from the fiscal perspective. Capability-adjusted life years (CALY). Adapted From S4. 60

7.1 Comparison of quality-adjusted life year (QALY) and capability-adjusted life years Sweden (CALY-SWE) anchoring on the 0 to 1 scale. 70

List of Tables

1.1 Overview of some prominent capability outcome measures. Adapted from S1. ICEpop CAPability (ICECAP) – measure for Adults (A) – measure for Older people (O) – Supportive Care Measure (SCM) – Close Person Measure (CPM) – Children and young people (CYP). Oxford capability (OxCAP). Capabilty-adjusted life years Sweden (CALY-SWE). 17

2.1 Delphi panel participants. Columns one to four contain the organization name, the SDG(s) related to the organization, if the organization representants participated in the start-up day, and if the representatives participated in the Delphi process, in order. Adapted from S1. SDG numbers: 1. No poverty, 3. Good health and wellbeing, 4. Quality education, 5. Gender equality. 8. Decent work and economic growth, 16. Peace, justice, and stronger institutions. Sustainable Development Goals (SDG). The Swedish Federation for Lesbian, Gay, Bisexual, Transgender, Queer, and Intersex Rights (RFSL). 24

2.2 Swedish phrasing and informal English translation. Adapted from the supplementary information to S2. 27

3.1 Timeline value set study. Time trade-off (TTO). Discrete choice experiment (DCE). 32

3.2 Results of the main model. Mean square error (MSE), Mean absolute error (MAE), and accuracy refer to the results of the cross validation. The 27,099 observations consisted of 20,364 discrete choice experiment (DCE) answers and 6,735 time trade-off (TTO) answers. 43

4.1 Disadvantage and plurality group categorizations. 46

5.1 Cost-effectiveness results. Costs in 2019 Dollars. *in millions. Capability-adjusted life years (CALY). Not in employment, education, or training (NEET). Adapted from S4. 59

6.1 Overview of the data sources used in the thesis studies. For the sample size column, the target sample size is displayed in brackets. *Data sources that were only used for methodological development, but do not contribute to the results of the studies directly. Third-party data sources such as Statistics Sweden [143] are not listed. Capability-adjusted life years Sweden (CALY-SWE). Discrete choice experiment (DCE). Time trade-off (TTO). 63

7.1 Overview of some prominent capability outcome measures, attributes, and selection thereof, and valuation method. Adapted from S1. Refer to table c1t1 for key references. Best-worst scaling (BWS). Time trade-off (TTO). Discrete choice experiments (DCE). ICEpop CAPability (ICECAP) – measure for Adults (A) – measure for Older people (O) – Supportive Care Measure (SCM) – Close Person Measure (CPM) – Children and young people (CYP). Oxford capability (OxCAP). Capability-adjusted life years Sweden (CALY-SWE). 75

7.2 Weights per group. Adapted from S3. CI (confidence interval); comparator (Cmp.); not in employment, education, or training (NEET). Δ stands for 'difference'. 105

List of Infoboxes

1.3.1 Questionnaire, Instrument, measure 4
1.3.2 Attributes and answer levels 4
1.3.3 State, or config 4
1.3.4 Weight, or quality score 4
1.3.5 Tarif, value set 4
1.3.6 Value set elicitation, valuing 4
1.3.7 Latent variable 5
1.5.1 Welfare, wellbeing, and QoL 10
1.7.1 CALY-SWE naming 13
1.7.2 Delphi panel 13
1.7.3 Equity efficiency trade-off 14
2.2.1 Fair-minded 21
3.2.1 Experimental design 35
3.2.2 Credible interval 36
3.2.3 Level sum score 37
4.2.1 Inequality, inequity, differences 45

Acknowledgement

Writing this thesis was a privilege, and many persons facilitated it directly or indirectly. In some ways, each and every encounter shaped my experiences and may have influenced this work – mostlyto the better. My gratitude extends first and foremost to the *Swedish population* that ultimately financed the work, through the Swedish Research Council for Health, Working Life and Welfare, *Forte*.
A deciding role played also the *participants* and *Delphi representatives* without whose participation the research would have been all but impossible.

On a personal level, the largest gratitude though is due for my supervisors:
Anna Månsdotter, for always challenging me in inspiring ways and being true role model for research and other aspects of life
Jan Hjelte, kindkly never failing deliver substantial review comments, rightfully reminding the health economists of social work matters, and not missing a meeting even while hunting
Lars Lindholm, brilliantly displaying how to pursue the essential, an indestructible optimist that brings out the best in others

For my wider journey I am infinitely thankful to my *mother* and *father* that made me who I am. Never will I forget their generosity, the many peaks I hiked up with my dad, nor the cheerful times with my mum.
Wholeheartedly I want to thank my wife *Helen* and our sons *Johan* and *Théoden* that – with love – supported me in deed, word and inspiration.
Also to my sisters *Rebekka* and *Seraina* I owe a great deal. They did not shy away from journeying far to see us, exemplifying that freedom of choice should not be taken for granted.
And to other relatives: My *Godmother* and *Godfather*, *Aunt*, *Grandmother*, *Grandfather*, *Cousins*, *Tatay*, *Alfa*, *Chloe*, and *Ate Melinda*. Gratefully indepted am i to my uncle *Christoph Züllig* that proofread the german summary.

A special thank goes to the many teachers that gifted me with a formidable education: *Frau Gans*, *Frau Dreyer*, *Frau Minder*, *Martin "Tigge" Sommer* – Was ich mache, mache ich gut, *Beat Eicher*, *Martin Lehner*, *Pierre Aeschlimann*, *Thomas Zwygart*, *Sabine Riard*, *Melanie Gerber*, *Alfred Schilt*, and *Matthias Walpen*.

Those working at the *Department of Epidemiology and Global Health* of course deserve a special mention. Many are not only excellent, inspiring researchers but also dear colleagues, providing the ideal working environment :

Faustine, my office neighbour that never fails to keep my French alive.
Mazen, my PhD godfather that showed me the ropes
Sara Forsberg, *Ulrika Harju*, *Ulrika Järvholm*, *Angelica Johansson*, *Karin Johanson*, *Eva Selin*, *Lena Mustonnen*, *Susanne Walther*, always cheerfully helping with administrative matters *Göran Lönnberg* that never fails to resolve any IT -related issue

I had the privilege to be taught by outstanding UmuEpiGH as scholars:
Isabela Goicolea, eloquently staying true to the cause
Per Gustafsson, vividly explaining the most complex things with passion
Anna-Karin Hurtig, a diligently humane leader, easily navigating complex health systems
Anneli Ivarsson, striving for excellence but never forgetting the human side of things
Jon Kinsman, teaching with the right message
Maria Lindkvist, a devoted and inspiring teacher, always putting the students first
Kristina Lindvall, thoroughly demonstrating the merits of qualitative research
Curt Löfgren, an inspiring teacher and economist
Fredrik Norström, always educationally offering an alternative view

Maria Nilsson, calmly determined to better the climate for health
Nawi Ng, a an inspiringly staunch, meticulous epidemiologist
Raman Preet, a true vocal champion for equity
Anni-Maria Pulkki-Brännström, never getting lost in the most complicated issues
Karl-Erik Renhorn, flawless organizer and thorough teacher
Klas-Göran Sahlen, never failing to see the essence and move things forward
Miguel San Sebastian, genuinely inspiring to not just criticize, but also to take action
Sun Sun, devoted researcher with untiring passion
Linda Richer Sunderg, a superb artist promoting sound implementation
Barbara Schumann, offering a grander perspective on health research
Lars Weinehall, demonstrating how to ask the right questions
Masoud Vazghemi, a cheerfull sharp thinker and excellent teacher

And others that deserve a special mention:
Peter Byas, for inspiring a random PhD students for social media
Brian and Jennifer, for being genuinely kind end encouraging
Håkan Jonsson, a thorough statistician with infectious enthusiasm
Henrik Holmberg, an excellent teacher that always sees the bigger picture

André, *Hedi* for offering valuable comments, and all other reviewers and opponents for the planning seminar and midterm seminar.
Adam, *Alieu*, *Anh*, *Anna*, *Anne Albert*, *Arian*, *Camilla*, *Carla*, *Catrine*, *Chanvo*, *Chama*, *Chavula*, *David*, *Edwinah*, *Edy*, *Ellia*, *Ester*, *Fethi*, *Guan*, *Hanna*, *Husseifa*, *Ida*, *Jennie*, *Jing*, *Jasper*, *Julia*, *Klara*, *Kapuya*, *Förskola Kragen*, *Linda*, *Mai*, *Magnus*, *Mike*, *Osvaldo*, *Panduleni*, *Paul*, *Puthy*, *Petter*, *Rhakal*, *Rodrigo*, *Wolfgong*, *Yercin* also deserve a mention for many cheerful encounters and their support.

Brendan Mulhern, *Fredrk Norström*, *Inna Feldman*, *Richard Ssegonja*, *Frida Jonsson*, *Kerstin Edin*, for excellent advice, collaboration and co-authorship
Sebastian Östlund and *Kalle* Grill for their contributions
Further I owe thanks to many at the ISPM: *Nicola Low* and *Michel Counotte* for scientific inspiration and trusting me to support their work. *Julien Riou* for showing me the Bayesian way . *Christian Althaus* for remaining vigilant

I also want to thank my friends:
Eric (also for critically reviewing the thesis), *Samuel*, *Daniel*, *Felizia* and *Dafinë* - carrying Seeländer spirit into the world
Jonas, *Sandra*, *Marc*, *Rene* and *Fredrik* - Friendship does not know borders
The *Filipino community* in Umeå - Spreading friendliness and kindness that close to the polar circle.
Gunnel and *Kjell-Arne*, our valued neighbours, and *Amir*, who hosted me so many times.
Anna-Mårta, *Bengt*, *Ingrid*; the most agreeable persons to converse and practice Swedish with
A large thank you also goes to the association Reformierte Studentenhäuser Zürich, especially *Friederike* and *Daniel* , and to the many i had the pleasure of meeting there; an institution that greatly facilitated my studies.

Of great worth were the numerous open source software pieces used in this thesis, among them *R* , *Rstudio*, LaTeX , the *Linux*, *archlinux* , *KDE* , and *stan*.
Importantly, attribution is due for all the brilliant, distinguished scientists whose work this thesis builds on – including *Amartya Sen*
Also to *Paul Tol* for permitting for his colour-blind and monochrome friendly vibrant contrast colour palette.

Previous PhD theses from the Department of Epidemiology and Global health

[163] **Melissa Brower Scribani**. *Insights on weight maintenance and impacts of obesity for two rural populations in the United States and Sweden*. 2023.

[162] **Johan Hambraeus**. *Interventional pain management focused on zygapophysial joint pain : a health-economic evaluation*. 2023.

[161] **Yercin Mamani-Ortiz**. *Cardiovascular disease prevention in Cochabamba, Bolivia : the importance of preventable risk factor distribution and inequalities for policy implementation*. 2023.

[160] **Mai Vu Thi Quynh**. *A feasibility and applicability study of a health-related quality of life measurement in Vietnam*. 2023.

[159] **Paul Joseph Amani**. *Does health insurance contribute to improving responsiveness of the health system? : the case of elderly in rural Tanzania*. 2022.

[158] **Mazen Baroudi**. *Leaving the door ajar : young migrants' sexual and reproductive health in Sweden*. 2022.

[157] **Mikael Emsing**. *Conflict management & mental health among Swedish police officers & recruits*. 2022.

[156] **Septi Kurnia Lestari**. *Active and healthy ageing in Europe : significance of social relationships*. 2022.

[155] **Prasad Liyanage**. *The Influence of Climate and Public Health Interventions on Aedes Vectors and Dengue in Sri Lanka*. 2022.

[154] **Nadja Trygg Fagrell**. *Knowledge diversification in public health through intersectionality*. 2022.

[153] **Anna Bengtsson**. *Pictorial presentation of subclinical atherosclerosis : a measure to reduce the risk for cardiovascular disease*. 2021.

[152] **Anne Gotfredsen**. *Carving out collective spaces : Exploring the complexities of gender and everyday stressors within rural youth leisure*. 2021.

[151] **Aditya L. Ramadona**. *Spatiotemporal prediction of arbovirus outbreak risk : the role of weather and population mobility*. 2021.

[150] **Julia Schröders**. *Diversity, dynamics and deficits : the role of social networks for the health of aging populations in Indonesia*. 2021.

[149] **Rakhal Gaitonde**. *Divergences, dissonances and disconnects : implementation of community-based accountability in India's national rural health mission*. 2020.

[148] **Junia Joffer**. *Health for future : self-rated health and social status among adolescents*. 2020.

[147] **Sulistyawati Sulistyawati**. *Dengue Prevention and Control in Indonesia : a case study in Yogyakarta City*. 2020.

[146] **Kamila Al-Alawi**. *Team-based approach in the management of diabetes at primary health care level in Muscat, Oman : challenges and opportunities*. 2019.

[145] **Robert Jonzon**. *Health assessments of asylum seekers within the Swedish healthcare system : a study of the interface between control and care, and how structure and procedures may influence access and coverage*. 2019.

[144] **Susanne Ragnarsson**. *Att klara skolan när huvudet dunkar och kroppen värker : en studie om återkommande smärta och skolprestation bland skolbarn*. 2019.

[143] **Eid Rodríguez** and **Daniel V**. *The rough journey to access health care : the case of leishmaniasis in the Bolivian rainforest*. 2019.

[142] **Pamela J. Tinc**. *Raising the (roll)bar : exploring barriers and facilitators to research translation in US public health*. 2019.

[141] **Nitin Gangane**. *Breast cancer in rural India : knowledge, attitudes, practices; delays to care and quality of life*. 2018.

[140] **Regis Hitimana**. *Health economic evaluation for evidence-informed decisions in low-resource settings : the case of Antenatal care policy in Rwanda*. 2018.

[139] **Frida Jonsson**. *The presence of the past : a life course approach to the social determinants of health and health inequalities in northern Sweden*. 2018.

[138] **Alireza Khatami**. *It is on my skin, on my soul, and on my life : development of a disease-specific quality of life instrument for adult patients with acute cutaneous leishmaniasis in Iran*. 2018.

[137] **Ida Linander**. *"It was like I had to fit into a category" : people with trans experiences navigating access to trans-specific healthcare and health*. 2018.

[136] **Jing Liu-Helmersson**. *Climate change, dengue and aedes mosquitoes : Past trends and future scenarios*. 2018.

[135] **Amaia Maquibar Landa**. *An insight into institutional responses to intimate partner violence against women in Spain*. 2018.

[134] **Per Nordin**. *Control or elimination : terms for public health interventions against tungiasis and schistosomiasis haematobium*. 2018.

[133] **Cahya Utamie Pujilestari**. *Abdominal obesity among older population in Indonesia : socioeconomic and gender inequality, pattern and impacts on disability and death*. 2018.

[132] **Nathanael Sirili**. *Health workforce development post-1990s health sector reforms : the case of medical doctors in Tanzania*. 2018.

[131] **Moses Tetui**. *Participatory approaches to strengthening district health managers' capacity : Ugandan and global experiences*. 2018.

[130] **Anna Westerlund**. *The role of implementation science in healthcare improvement efforts : investigating three complex interventions*. 2018.

[129] **Anna Brydsten**. *Yesterday once more? Unemployment and health inequalities across the life course in northern Sweden*. 2017.

[128] **Juan Antonio Córdoba Doña**. *Withstanding austerity : economic crisis and health inequalities in Spain*. 2017.

[127] **Laila Daerga**. *Att leva i två världar : hälsoaspekter bland renskötande samer*. 2017.

[126] **Jonas Hansson**. *Mind the blues : Swedish police officers' mental health and forced deportation of unaccompanied refugee children*. 2017.

[125] **Therese Kardakis**. *Strengthening lifestyle interventions in primary health care : the challenge of change and implementation of guidelines in clinical practice*. 2017.

[124] **Kateryna Karhina**. *Social capital and well-being in the transitional setting of Ukraine*. 2017.

[123] **Kanyiva Muindi**. *Air pollution in Nairobi slums : sources, levels and lay perceptions*. 2017.

[122] **Maquins Odhiambo Sewe**. *Towards climate based early warning and response systems for malaria*. 2017.

[121] **Johanna Sundqvist**. *Forced repatriation of unaccompanied asylum-seeking refugee children : towards an interagency model*. 2017.

[120] **Phan Minh Trang**. *Weather and extreme heat in association to mental disorders : The case of Hanoi, Vietnam*. 2017.

[119] **Masoud Vaezghasemi**. *Nutrition transition and the double burden of malnutrition in Indonesia : a mixed method approach exploring social and contextual determinants of malnutrition*. 2017.

[118] **Osama Ahmed Hassan Ahmed**. *Rift Valley fever : challenges and new insights for prevention and control using the "One Health" approach*. 2016.

[117] **Hagos Godefay Debeb**. *Maternal mortality then, now, and tomorrow : The experience of tigray region, northern ethiopia*. 2016.

[116] **Vijendra Ingole**. *Too hot! : an epidemiological investigation of weather-related mortality in rural india*. 2016.

[115] **Vu Duy Kien**. *Inequalities in non-communicable diseases in urban Hanoi, Vietnam : health care utilization, expenditure and responsiveness of commune health stations*. 2016.

[114] **Kaaren Mathias**. *Shadows and light : examining community mental health competence in North India*. 2016.

[113] **Anne Neumann**. *Prevention of type 2 diabetes : modeling the cost-effectiveness of diabetes prevention*. 2016.

[112] **Mikkel B. Quam**. *Imported infections' importance : global change driving Dengue dynamics*. 2016.

[111] **Eva Randell**. *Adolescent boys' health : managing emotions, masculinities and subjective social status*. 2016.

[110] **Linda Richter Sundberg**. *Mind the Gap : exploring evidence-based policymaking for improved preventive and mental health services in the Swedish health system*. 2016.

[109] **Ryan G Wagner**. *The Burden of Epilepsy : using population-based data to define the burden and model a cost-effective intervention for the treatment of epilepsy in rural South Africa*. 2016.

[108] **Berit Edvardsson**. *"Det är inte mig det är fel på, det är huset" : en studie av prognosfaktorer och bemötande med fokus på sjuka hus-syndromet*. 2015.

[107] **Thaddaeus Wandera Egondi**. *Making visible the invisible : Health risks from environmental exposures among socially deprived populations of Nairobi, Kenya*. 2015.

[106] **Ameneh Setareh Forouzan**. *Assessing responsiveness in the mental health care system : the case of Tehran*. 2015.

[105] **Tesfay Gebregzabher Gebrehiwet**. *No woman should die while giving life : Does the Health Extension Program improveaccess to maternal health services in Tigray, Ethiopia?*, 2015.

[104] **Paul Mee**. *Who died, where, when and why? : an investigation of HIV-related mortality in rural South Africa*. 2015.

[103] **Joseph Mumba Zulu**. *Integration of national community-based health worker programmes in health systems : Lessons learned from Zambia and other low and middle income countries*. 2015.

[102] **Fredinah Namatovu**. *The multifactorial etiology of celiac disease explored by combining several national registers*. 2015.

[101] **Bharat Randive**. *Study of conditional cash transfer programme Janani Suraksha Yojana for promotion of institutional births : Studies from selected provinces of India*. 2015.

[100] **Ailiana Santosa**. *Where are the world's disease patterns heading? : The challenges of epidemiological transition*. 2015.

[99] **Jennifer Crowe**. *Heat exposure and health outcomes in Costa Rican sugarcane harvesters*. 2014.

[98] **Alison Hernandez**. *Enabling the performance of nurses in rural Guatemala : the role of relationships*. 2014.

[97] **Elisabet Höög**. *Navigera i ständig förändring : facilitering av utvecklingsarbete inom vård och omsorg*. 2014.

[96] **Curt Löfgren**. *Catastrophic health expenditure in Vietnam : studies of problems and solutions*. 2014.

[95] **Paola Andrea Mosquera Méndez**. *Evaluation of a primary health care strategy implemented in a market-oriented health system : the case of Bogota, Colombia*. 2014.

[94] **Tej Ram Jat**. *Maternal health and health care in Madhya Pradesh state of India : an exploration using a human rights lens*. 2014.

[93] **Fatwa Sari Tetra Dewi**. *Working with community : exploring community empowerment to support non-communicable disease prevention in a middle-incom country*. 2013.

[92] **Kristina Edvardsson**. *Health promotion in pregnancy and early parenthood : the challenge of innovation, implementation and change within the Salut Programme*. 2013.

[91] **Elli Nur Hayati**. *Domestic violence against women in rural Indonesia : searching for multilevel prevention*. 2013.

[90] **Yien Ling Hii**. *Climate and dengue fever : early warning based on temperature and rainfall*. 2013.

[89] **Siddhivinayak Hirve**. *"In general, how do you feel today?" Self-rated health in the context of aging in India*. 2013.

[88] **Anand Krishnan**. *Gender inequity in child survival : travails of the girl child in rural north India*. 2013.

[87] **Kristina Lindvall**. *Being able to be stable : exploring primary weight maintenance as a public health strategy for obesity prevention*. 2013.

[86] **Faustine Nkulu Kalengayi**. *A world on the move : challenges and opportunities for hiv/aids and tuberculosis care and prevention among vulnerable migrant populations in Sweden*. 2013.

[85] **Katrina Nordyke**. *Mass screening for celiac disease : a public health intervention from the participant perspective*. 2013.

[84] **Cynthia Anticona Huaynate**. *Lead exposure in indigenous children of the Peruvian Amazon : seeking the hidden source,venturing into participatory research*. 2012.

[83] **Felix Kisanga**. *Child sexual abuse in urban Tanzania : possibilities and barriers for prevention*. 2012.

[82] **Hailemariam Lemma**. *Improving efficiency, access to and quality of the rural health extension programme in Tigray, Ethiopia : the case of malaria diagnosis and treatment*. 2012.

[81] **Anna Myléus**. *Towards explaining the Swedish epidemic of celiac disease : an epidemiological approach*. 2012.

[80] **Quang Ngoc Nguyen**. *Understanding and managing cardiovascular disease risk factors in Vietnam : integrating clinical and public health perspectives*. 2012.

[79] **Fredrik Norström**. *The burden of celiac disease and the value of having it diagnosed*. 2012.

[78] **Anna Rosén**. *Mass screening for celiac disease in 12-year-olds : Finding them and then what?*, 2012.

[77] **Ana Lorena Ruano**. *The role of social participation in municipal-level health systems : the case of Palencia, Guatemala*. 2012.

[76] **Pham Thai Son**. *Hypertension in Vietnam : from community-based studies to a national targeted programme*. 2012.

[75] **Gasto Frumence**. *The role of social capital in HIV prevention: experiences from the Kagera region of Tanzania*. 2011.

[74] **Rose Mjawa Laisser**. *Prevention of intimate partner violence : community and healthcare workers´ perceptions in urban Tanzania*. 2011.

[73] **Gunnar Lundqvist**. *Tobaksvanor i medelåldern : riskfaktormönster, rökstoppsattityder och erfarenheter av att sluta röka*. 2011.

[72] **Stephen Maluka**. *Strengthening fairness, transparency and accountability in health care priority setting at district level in Tanzania : opportunities, challenges and the way forward*. 2011.

[71] **Zainonisa Petersen**. *Smoking cessation during pregnancy : a person-centred approach among disadvantaged women in South Africa*. 2011.

[70] **Virgilio Mariano Salazar Torres**. *Intimate partner violence in Nicaragua : studies on ending abuse, child growth, and contraception*. 2011.

[69] **Ann Sörlin**. *Health and the elusive gender equality : Can the impact of gender equality on health be measured?*, 2011.

[68] **Le Van Hoi**. *Health for community dwelling older people : trends, inequalities, needs and care in rural Vietnam*. 2011.

[67] **L'e;onie Dapi N.** *Socioeconomic and sex differences in adolescents' dietary intake, anthropometry and physical activity in Cameroon, Africa*. 2010.

[66] **Malin Eriksson**. *Social capital, health and community action : implications for health promotion*. 2010.

[65] **Helene Johansson**. *En mer hälsofrämjande hälso- och sjukvård : hinder och möjligheter utifrån professionernas perspektiv*. 2010.

[64] **Ari Probandari**. *Revisiting the choice : to involve hospitals in the partnership for tuberculosis control in Indonesia*. 2010.

[63] **Susanne Waldau**. *Creating organisational capacity for priority setting in health care : using a bottom-up approach to implement a top-down policy decision*. 2010.

[62] **Maria Wiklund**. *Close to the edge : discursive, gendered and embodied stress in modern youth*. 2010.

[61] **Mark A Collinson**. *Striving against adversity. : the dynamics of migration, health and poverty in rural South Africa*. 2009.

[60] **Anders Emmelin**. *Counted - and then? : trends in child mortality within an Ethiopian demographic surveillance site*. 2009.

[59] **Isabel Goicolea**. *Adolescent pregnancies in the Amazon basin of Ecuador : a rights and gender approach to girls' sexual and reproductive health*. 2009.

[58] **Fyson Kasenga**. *Making it happen : prevention of mother to child transmission of HIV in rural Malawi*. 2009.

[57] **Hussein L Kidanto**. *Improving quality of perinatal care through clinical audit : a study from a tertiary hospital in Dar es Salaam, Tanzania*. 2009.

[56] **Maria Nilsson**. *Promoting health in adolescents : preventing the use of tobacco*. 2009.

[55] **Klas-Göran Sahlén**. *An ounce of prevention is worth a pound of cure : preventive home visits among healthy seniors*. 2009.

[54] **Julie Sorensen**. *Social marketing for injury prevention : changing risk perceptions and safety-related behaviors among New York farmers*. 2009.

[53] **Kjerstin Dahlblom**. *Home alone : sibling caretakers in león, nicaragua*. 2008.

[52] **Mesganaw Fantahun**. *Mortality and survival from childhood to old age in rural Ethiopia*. 2008.

[51] **Edward F Fottrell**. *Dying to count : mortality surveillance methods in resource-poor settings*. 2008.

[50] **Fikru Tesfaye**. *Epidemiology of cardiovascular disease risk factors in Ethiopia : the rural-ruban gradient*. 2008.

[49] **Stephen M Tollman**. *Closing the gap : applying health and socio-demographic surveillance to complex health transitions in South and sub-Saharan Africa*. 2008.

[48] **Lars Hagberg**. *Cost-effectiveness of the promotion of physical activity in health care*. 2007.

[47] **Lars Jerdén**. *Health-promoting health services : personal health documents and empowerment*. 2007.

[46] **Kerstin E Edin**. *Perspectives on intimate partner violence, focusing on the period of pregnancy*. 2006.

[45] **Dao Lan Huong**. *Mortality in transitional vietnam*. 2006.

[44] **Kathleen Kahn**. *Dying to make a fresh start : mortality and health transition in a new South Africa*. 2006.

[43] **Anna Månsdotter**. *Health, economics, and feminism : on judging fairness and reform*. 2006.

[42] **Hoang Van Minh**. *Epidemiology of cardiovascular disease in rural Vietnam*. 2006.

[41] **Anne Nafziger**. *A population perspective on obesity prevention : lessons learned from Sweden and the U.S.* 2006.

[40] **Nawi Ng**. *Chronic disease risk factors in a transitional country : the case of rural Indonesia*. 2006.

[39] **Margareta Norberg**. *Identifying risk of type 2 diabetes : epidemiologic perspectives from biomarkers to lifestyle*. 2006.

[38] **Nils Oscarson**. *Health economic evaluation methods for decision-making in preventive dentistry*. 2006.

[37] **Sven Hassler**. *The health condition in the Sami population of Sweden, 1961-2002 : Causes of death and incidences of cancer and cardiovascular diseases*. 2005.

[36] **Thanh Nguyen Xuan**. *The injury poverty trap in rural Vietnam : causes, consequences and possible solutions*. 2005.

[35] **Eliette Valladares Cardoza**. *Partner violence during pregnancy, psychosocial factors and child outcomes in Nicaragua*. 2005.

[34] **Jeremiah Chikovore**. *Gender power dynamics in sexual and reproductive health. A qualitative study in Chiredzi District, Zimbabwe*. 2004.

[33] **Maria Emmelin**. *Self-rated health in public health evaluation*. 2004.

[32] **Hoang Minh Hang**. *Epidemiology of unintentional injuries in rural Vietnam*. 2004.

[31] **Theresia Ninuk Sri Hartini**. *Food habits, dietary intake and nutritional status during economic crisis among pregnant women in Central Java, Indonesia*. 2004.

[30] **Torbjörn Lind**. *Iron and zinc in infancy : results from experimental trials in Sweden and Indonesiaa*. 2004.

[29] **Gunilla Ringbäck Weitoft**. *Lone parenting, socioeconomic conditions and severe ill-health : longitudinal register-based studies*. 2003.

[28] **SM Ziyaddin Hyder**. *Anemia and iron deficiency in women : impact of iron supplementation during pregnancy in rural Bangladesh*. 2002.

[27] **Adi Utarini**. *Evaluation of the user-provider interface in malaria control programme : The case of Jepara District, Central Java Province, Indonesia*. 2002.

[26] **Lena Grönblom-Lundström**. *Rehabilitation in light of different theories of health : Outcome for patients with low-back complaints - a theoretical discussion*. 2001.

[25] **Anneli Ivarsson**. *On the multifactorial etiology of celiac disease*. 2001.

[24] **Detty Siti Nurdiati**. *Nutrition and reproductive health in Central Java, Indonesia. An epidemiological approach*. 2001.

[23] **Ann Öhman**. *Profession on the move : Changing conditions and gendered development in physiotherapy*. 2001.

[22] **Tobias Andersson**. *Survival of mothers and their offspring in 19th century Sweden and contemporary rural Ethiopia*. 2000.

[21] **Yemane Berhane**. *Women's health and reproductive outcome in rural Ethiopia*. 2000.

[20] **Mary Carroll Ellsberg**. *Candies in hell : research and action on domestic violence against women in Nicaragua*. 2000.

[19] **Eva Johansson**. *Emerging perspectives on tuberculosis and gender in Vietnam*. 2000.

[18] **Lennarth Nyström**. *Assessment of population screening. The case of mammography*. 2000.

[17] **Ingrid Mogren**. *Reproductive factors' impact on the health of mother and offspring : an epidemiological study*. 1999.

[16] **Rodolfo Peña**. *Infant mortality in transitional Nicaragua*. 1999.

[15] **Elmer Zelaya Blandon**. *Adolescent pregnancies in Nicaragua. The importance of education*. 1999.

[14] **Maryam Mohamed Ibrahim**. *Child health in Somalia. An epidemiological assessment i rural communities during a pre-war period*. 1998.

[13] **Margareta Barnekow Bergkvist**. *Physical capacity, physical activity and health : a population based fitness study of adolescents with an 18 year follow-up*. 1997.

[12] **Bertil Forsberg**. *Urban air quality and indicators of respiratory problems*. 1997.

[11] **Lars Weinehall**. *Partnership for health : on the role of primary health care in a community intervention programme*. 1997.

[10] **Lars Lindholm**. *Health economic evaluation of community-based cardiovascular disease prevention : some theoretical aspects and empirical results*. 1996.

[9] **Marianne Wulff**. *Reproductive hazards in an industrial setting : an epidemiological assessment*. 1996.

[8] **Erik Bergström**. *Cardiovascular risk indicators in adolescents : the Umeå youth study*. 1995.

[7] **Japhet Killewo**. *Epidemiology towards the control of HIV infection in Tanzania with special reference to the Kagera region*. 1994.

[6] **Lulu Muhe**. *Child health and acute respiratory infections in Ethiopia : epidemiology for prevention and control*. 1994.

[5] **Berndt Stenberg**. *Office illness : the worker, the work and the workplace*. 1994.

[4] **Inger Brännström**. *Community participation and social patterning in cardiovascular disease intervention*. 1993.

[3] **Desta Shamebo**. *Epidemiology for public health research and action in a developing society : the Butajira Rural Health Project in Ethiopia*. 1993.

[2] **Anita Sandström**. *Epidemiology at a smeltery : changing patterns concerning lung cancer occurrence, work environment, smoking and risk perceptions over six decades*. 1992.

[1] **Måns Rosén**. *Epidemiology in planning for health : with special reference to regional epidemiology and the use of health registers*. 1987.

Department of Epidemiology and Global Health
Umeå University, SE-901 87 Umeå, Sweden
www.umu.se/english

Art by Linda Richter Sundberg

ISBN 978-91-8070-257-7 (print)
ISBN 978-91-8070-258-4 (pdf)
ISSN 0346-6612

UMEÅ UNIVERSITY