



Social Protection

Innovative Investment
in Long-Term Care

FEASIBILITY FRAMEWORK TOOL FOR SOCIAL INVESTMENTS IN LONG- TERM CARE

Adam Richards (Social Value UK), Ismo Linnosmaa, Hanna Jokimäki and Taru Haula (National Institute for Health and Welfare) Eds.

www.sprint-project.eu



The SPRINT Project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 94985.

Executive Summary

The definition of SI in LTC presented in Lopes (2017) calls for sustainable and equitable care provision that enhances quality of life and the ability to participate in society and the economy. To achieve this, there is the need to make decisions about the resourcing of long-term care (LTC) which take account of the multiple costs and benefits for different stakeholders. Therefore, extending previous reports that identified important conceptual and environmental issues which informed the dimensions of impact identified Richards (2018), this report has identified how to define and measure relevant outcomes of Social Investment (SI) in LTC.

The quantitative measurement of labour market, substitution, and health and wellbeing outcomes is essential to assist ex-ante decision-making. Appreciating the significance of different outcomes provides an understanding of which are material and require inclusion within governance and decision-making of SI in LTC.

A key element of understanding the significance of outcomes is appreciating their relative significance. Given the requirement to value outcomes in the same language as the costs of their production, this report therefore examined different approaches of monetisation. Although not presented as a traditional academic publication, a rapid literature review on the measurement of appropriate outcomes, as well as options to monetise was conducted. This facilitated the creation of a theoretical context for this report.

In addition, a systematic review of published empirical evidence was undertaken. This revealed substantial gaps in current rigorous evidence and identified where greater research is required. Notwithstanding challenges of measurement, monetising of outcomes, particularly intangible effects on people is even less developed. Labour market and substitution effects are more traditionally monetised. However, in addition to these, health and wellbeing concerns were also examined and valued. Although it was not possible to create a complete assessment of the outcomes of SI in LTC, illustrative examples of how to measure and value such outcomes was produced.

Value is a subjective construct, and the nature of health and wellbeing means there is unlikely to ever be a single value used consistently to analyse such outcomes. Therefore, no single financial proxy was identified for health and wellbeing outcomes. Instead, a range of options were identified that aim to contribute to the debate and demonstrate the possibility to do so.

This report identified that there are significant gaps in current evidence, in particular for intangible effects. Nevertheless, illustrative examples focused on informal care were created that highlight the potential to measure and value outcomes of LTC. The presented results are not to be taken by decision-makers as the evidence to inform their choices. However, using existing data this report demonstrates the potential to monetise different outcomes of LTC. In doing so, it has achieved its aim and also helped to identify where greater evidence is required.

The discussions presented in this report mirror those also taking place within impact investing, with a growing focus on key impact questions that need to be asked to account for the effects of decisions. Therefore, the potential presented in this report, and SPRINT overall is not limited to LTC policy makers, with results also being potentially of interest to other social actors such as impact investors.

To progress the debate and practice of SI in LTC, a framework tool has also been developed as part of this report. The tool is framed by the key impact questions. These need to be implemented in line with the principles of economic analyses generally, and social return on investment (SROI) that require value to be identified by those actors experiencing outcomes. This information provides those making decisions about the allocation of resources with an improved ability to understand the effects of their decisions, taking account of different stakeholders and their experiences.

Although it was not possible to create a complete appraisal of the costs and benefits of SI in LTC, this report does provide the practical advances required in line with the grant agreement. It demonstrates how to measure outcomes, taking account of important issues such as the extent of changes and causality, whilst also demonstrating the theoretical potential to identify the relative importance of outcomes. It has also identified the means by which this information can help to inform important decisions about the allocation of resources to support the aims of SI in LTC.

Key messages

- There is a lack of systematic evidence on the outcomes of long-term care.
- The potential to measure and value different effects of long-term care from the perspective of the different stakeholders is feasible when framed by the principles of Social Return on Investment and key impact questions.
- There is a need for decision-making to appreciate evidence informed by different stakeholder-perspectives.

Table of Contents

Executive Summary	2
Key messages	3
Acronyms and Abbreviations	6
1 Introduction	7
2 Aims and Objectives	7
3 Methods	8
4 Theoretical Context	10
4.1 Measuring Intangible Social Impacts	10
4.2 Specific Measures of the Intangible Outcomes of Social Care	13
4.3 Approaches to Monetising	14
4.3.1 <i>Cost-based Approaches</i>	15
4.3.2 <i>Stated preference techniques</i>	15
4.3.3 <i>Revealed preferences</i>	16
4.3.4 <i>Wellbeing Valuation approach</i>	17
4.4 The Outcomes of Social Investment in Long Term-Care	18
5 Monetising Outcomes of Long Term-Care – Illustrative Examples on Informal Care	19
5.1 Valuing Labour Market Effects	19
5.2 Valuing Substitution Effects	21
5.3 Valuing Health and Wellbeing	21
5.3.1 <i>Approaches to value health and wellbeing: Stated preference</i>	22
5.3.2 <i>Approaches to value health and wellbeing: The Wellbeing Valuation approach</i>	22
5.4 Monetising the Health and Wellbeing Effects of LTC Initiatives for Informal Care	24
6 Conclusions	31
7 References	40
8 Annex 1: Overview of the Results of Results for the Theoretical Context	45
9 Annex 2: Categories Employed in Web of Science Searches for Theoretical Context	49
10 Annex 3: Overview of the Health Results for the Empirical Review	50
11 Annex 4: Overview of Different Health and Wellbeing Measures	55
12 Annex 5: Overview of Approaches to Monetise Outcomes	57
13 Annex 6: Further Information on the Wellbeing Valuation Methodology	63
14 Annex 7: Diminishing Marginal Utility of SWEMWBS Scores	64
15 Annex 8: Labour Market Effects Variables and Evidence	65
16 Annex 9: Monetized Results of Labour Market Effects of Increased Provision of Informal Care	66
17 Annex 10: Substitution Effects Variables and Evidence	72
18 Annex 11: Monetized Results of Substitution Effects Based Upon the Provision of Informal Care Based Upon Sample of 420 Care Recipients	73

19	Annex 12: Willingness to Accept Values for Having Unmet Needs.....	74
20	Annex 13: Life-satisfaction Values of Levels For All ASCOT Domains	75
21	Annex 14: Health and Wellbeing Reference Value Options to Calculate the Worth of LTC Initiatives.....	76
22	Annex 15: Exchange Rates of Different Measures of Health and Wellbeing.....	77
	Acknowledgments.....	78

Acronyms and Abbreviations

ASCOT	Adult Survey Care Outcomes Toolkit
CBA	Cost-benefit analysis
DALY	Disability-adjusted life year
ELSA	English Longitudinal Survey of Ageing
EQ-5D	EuroQual – 5 Dimensions
ICECAP - O	Investigating Choice Experiments Capability Measures for Older People/ Adults
LTC	Long-term care
QALY	Quality-adjusted life year
SCRQoL	Social care-related quality of life
SHARE	Survey of Health, Ageing and Retirement in Europe
SI	Social investment
SROI	Social Return on Investment
SVI	Social Value International
SVUK	Social Value UK
SWEMWBS	Short Warwick-Edinburgh Mental Well-Being Scale
WTA	Willingness to accept
WTP	Willingness to pay

1 Introduction

This report extends earlier SPRINT research by Barbieri and Ghibelli (2018), Greve *et al.* (2018), Richards (2018) and Poškutė (2018), which identified a range of relevant dimensions of impact of social investment (SI) in long-term care (LTC), as well as appropriate principles to appropriately measure and manage the resultant impacts.

Although there are currently significant gaps in evidence, through illustrative examples, this report demonstrates the potential to measure and value relevant outcomes of SI in LTC. The dimensions of impact were assessed using the key impact questions also identified in the reports mentioned above to frame how this was operationalised. This defined the issues in greater detail and identified where greater research is required.

The means by which outcomes can be valued was examined in particular detail. Different options to monetise tangible and intangible outcomes were outlined, with attention paid to the potential to apply to LTC outcomes for multiple stakeholders. This facilitated the creation of illustrative examples of outcomes associated with informal care and the support of those caring for loved ones. Concerns about work-life balance and the effects on the health and wellbeing of informal carers is regularly discussed, and previous reports have indicated the pressures on such care. Informal care was therefore identified as a useful focus for the illustrative examples provided within this report.

Although not able to provide an examination of all relevant outcomes of SI in LTC, the illustrative examples do demonstrate the potential to measure and value outcomes in line with the principles of economic analysis and Social Return on Investment (SROI) as per the grant agreement.

In addition, and based upon the impact questions, a framework tool for SI in LTC was developed (http://sprint-project.eu/impact_tool; the user guide is attached to the end of this report). The tool provides an approach by which relevant evidence can be collated and analysed to support ex-ante decision-making about the allocation of resources. It also helps to further identify important gaps in evidence. An animation¹ has also been developed that summarises the key issues regarding the pressures facing LTC, the principles of SROI, and the challenges that those making decisions about resource-allocation face.

2 Aims and Objectives

The research question that frames this report is: *How can outcomes of long-term care be measured and monetised?*

¹ <https://www.youtube.com/channel/UCZK7iRVtPSbUCMlu67YEvig/featured>.

The following objectives underpin the work:

- To assess the means to measure tangible and intangible effects of long-term care;
- To illustrate various options available to monetise outcomes of long-term care;
- To provide illustrative examples of measured and valued outcomes of long-term care;
- To outline a framework tool that can collate evidence on outcomes of long-term care to assist decision-making.

3 Methods

The methodology for this report is two-fold. Firstly, a rapid review of the theoretical context was conducted. This provided a basis of the issues to be addressed regarding the measurement of LTC outcomes and valuation approaches. Secondly, a systematic review of available empirical evidence was undertaken. This examined the extent of available evidence, as well as supporting the creation of illustrative examples that measured and valued relevant LTC outcomes associated with informal care.

Search terms for the theoretical context were created in collaboration with partners of SPRINT in accordance with the co-designed structure for this report. This strengthened the search criteria, rather than relying on a single partner. As a result, a series of search terms were identified that importantly included any potentially important variations of terminology. For example, when searching for issues related to the measurement of quality of life, variations of both elements of the search – such as ‘measurement’, and ‘quantifying’, as well as variations of the phrase ‘quality of life’ were included. This approach increased the likelihood of retrieving as many relevant articles as possible. All searches were conducted on the Web of Science search engine, with initial results further refined with criteria including restrictions to peer reviewed material, the date of publication, and categories of interest (such as economics and gerontology).

A full list of search terms used in the rapid-review and the results at the different refinements is provided in Annex 1, while Annex 2 outlines the categories included in the search. Initially, 12,401 articles were identified, although after refinements, 1,147 articles were retrieved. These were subsequently screened by title and abstract, with those articles excluded typically being focused on a specific illness or disorder. This reduced the results to 178, which was further condensed to 123 with a second screening exercise that assessed those where relevance was uncertain based on the review of title and abstract alone. Therefore, in total 123 articles were identified for review. All information was stored on an Excel document, with an overall tab displaying the results of each stage of refinement, and subsequent tabs highlighting the results for each of the 9 searches conducted (this number of searches was required given the restrictions to the quantity of text permitted in each search). Each article was reviewed against a consistent template that created an

efficient means of compiling and comparing notes against key dimensions including identification of evidence to support the thesis, and any limitations identified.

Searches were also conducted for empirical evidence to understand the current level of evidence-availability and support the illustrative examples presented within this report. Again, the process of creating appropriate search criteria was conducted by SPRINT partners. This element was restricted to evidence relevant to the specific context of informal care provided for older people.

The search criteria employed was based upon the dimensions of impact and relevant outcomes identified in Greve *et al.* (2018). The systematic empirical review was conducted following the PRISMA approach (<http://www.prisma-statement.org/>). The literature search was conducted with an information specialist. The search words and combinations used were “aged” or “elderly” or “senior” or “older adults” or “older people” or “old old” or “very old” AND “informal care” AND “effect*” or “impact*” or “causal*” (the term ‘outcome’ was not included as this resulted in an unmanageable number of results in the millions) AND “employment” or “labor market” or “work hours” or “working” or “wages” or “formal service*” or “health” or “well-being” or “quality of life” or “life satisfaction”. Peer-reviewed literature was searched from the Ebsco, Proquest and Web of Science databases. Unlike the literature searches for the theoretical context element that was adequately addressed by a single search facility, SPRINT partners wanted to ensure a broader approach to ensure relevant evidence was identified. There were no restrictions on the date of publication. There is always the potential that search terms do not capture all relevant articles. For example, there was the potential to also include terms such as unpaid carers in addition to informal carers. However, given that some informal carers receive financial compensation, it was felt that the latter term was appropriate to capture the relevant evidence.

The eligibility of the retrieved records was reviewed independently by two researchers based on inclusion and exclusion criteria. If consensus was not reached, a third reviewer was requested. An article was considered eligible if it studied the causal effects of informal care on any of the following outcomes: a) health (either of the informal carer or the care recipient/beneficiary), b) wellbeing (either of the carer or the care recipient/beneficiary), c) labour market and d) substitution between different care forms, and was published in English. In addition, eligibility was restricted to those articles that had applied a sufficiently rigorous methodology. As such, the three options for studying causal effects of ‘randomised controlled trials’, OR ‘instrumental variables’, OR ‘matching’ were also included.

Only peer reviewed academic articles were included, with other forms of publications such as abstracts, book chapters and editorials excluded. However, in order to maximise the results, no restrictions in relation to the nation in which the study took place were included.

A summary table was used containing information of the included articles: authors, publishing year and country, the objective, data and methods, and the results of the study (Annex 3 presents the health results as an example). The findings were summarized thematically. From the 516 articles identified prior to the inclusion of the causal effects criteria, only 27 articles were identified. After removing the duplicates and screening the abstracts, 20 articles were assessed as

full-texts, of which 4 concerned the effects on health, 6 on labour market, 9 on substitution, and 1 for wellbeing.

The use of valuation techniques for substitution and labour-market effects is a well-recognised approach to evaluation and as such the example articles were purposively selected to provide a demonstration of valuation approach. However, the wellbeing effects article did not provide evidence of the extent of an outcome, rather only indicating to the increased potential of a change (i.e. the spill-over effects of ill-health from care recipient to informal carer). Having demonstrated a lack of robust evidence, to produce illustrative examples it was necessary to identify alternative sources. This was completed using partner input that identified a forthcoming international review (Brimblecombe *et al.* 2018) on support for informal carers. Although the evidence identified was not all as robust as the previous search had revealed, it served to provide the opportunity to show the potential to monetise intangible outcomes in line with the task requirements and framed by the SROI principles and impact questions outlined in Richards (2018)

In addition, all calculations within the illustrative examples were conducted by two researchers. This was done to ensure accuracy in the approaches undertaken and increase confidence in the results.

4 Theoretical Context

4.1 Measuring Intangible Social Impacts

Terms such as wellbeing, life satisfaction, and quality of life are often viewed synonymously. However, the OECD (2015) highlights an individual’s wellbeing as the principal concern:

Wellbeing		
Quality of life		Material conditions
Health status	Civic engagement and governance	Income and wealth
Work-life balance	Environmental quality	Jobs and earnings
Education and skills	Personal security	Housing
Social connections	Subjective well-being	

Sourced from OECD (2018)

Similarly, the concept of ‘well-being for all’ guides the definition of social cohesion from the European Council (European Foundation for the Improvement of Living and Working Conditions (EFILWC), 2010).

Some of the multiple issues of wellbeing are more obviously related to LTC than others, with many interconnected and components of more than one of the dimensions of impact highlighted in earlier SPRINT research. For example, although social impacts relate primarily to effects upon individuals, there are also macro-level social impacts that are captured within the dimensions of labour market and substitution effects. Equally, even those components that may appear less relevant will still likely impact on a person’s judgement of their wellbeing. As will be discussed, different measures of wellbeing have the potential to reflect these issues.

In line with the identified dimensions of impact, this report therefore focuses on health and subjective wellbeing as relevant constituents of quality of life. Although measurement of health-status is a reasonably well-developed area, concerns of subjective wellbeing are less universally understood.

Given the importance of policy choices that affect the lives of people, there are general conditions that need to be satisfied for health and wellbeing evidence to meaningfully contribute to policy debate. Aware that social information needs to be presented in a timely fashion to influence policy choices (Commission of the European Communities 2009), Dolan *et al.* (2011) identify criteria for theoretical rigour, policy relevance, and empirical robustness as essential. Accordingly, three options to account for wellbeing are proposed– those of preference satisfaction, objective lists, and mental states/subjective wellbeing (*ibid*).

Preference satisfaction utilises GDP as a proxy measure of wellbeing. This is based on the premise that the more income available, the greater the ability to satisfy one’s preferences, thereby improving wellbeing (Hicks *et al.* 2013). This macro-level measure in isolation has however been dismissed for failing to represent the broader considerations of wellbeing (e.g. Stiglitz *et al.* 2009). Although average income levels can provide an indication of national economies, they are limited in their ability to consider implications for different segments of the population.

In addition, evidence suggests both complexity and limitations in the link between income and wellbeing (see for example, Kahneman and Deaton 2010). In contrast, objective lists provide people with pre-determined options of social indicators that can complement those of economic statistics. Although there is no universal agreement as to the constituents of such pre-determined lists, national policy will often target certain indicators such as access to education and healthcare services (Dolan *et al.* 2011). These demonstrate the need for social elements to complement economic performance. However, they have been criticised for providing a range of options that is too long for meaningful analysis and it is left to analysts to make decisions of what is material (Self 2015). Similarly, as with increases to income, evidence suggests that objective indications of standards of living are only weakly correlated to individuals’ subjective wellbeing beyond a certain level (EFILWC 2010).

Subjective wellbeing is the third option. Although no universal agreement on the elements of subjective wellbeing, it is reasonable to classify the constituents into two broad categories of

hedonic and eudemonic happiness. Hedonic happiness considers positive and negative affect measures, relating to experienced components, such as moods, feelings, and emotions (OECD 2015) with questions often focusing on a defined and recent time- period (Kahneman and Deaton 2010).

Yet, hedonic happiness alone can be criticised as assessing a momentary state of happiness, and the limitations of policy dependent variables such as income and employment to influence it are understood (Layard 2016). Eudaimonia alternatively focuses on the idea that achieving one's potential is the true goal of life. What constitutes a good life is itself open to interpretation, yet there is general agreement that a person's sense of purpose, meaning, and worthwhileness are important (OECD 2015). This approach has a greater focus on the psychological components of life and can therefore be associated with measures of life-satisfaction. Regularly asked as a scaled-means to appreciate the current appraisal of one's life, it is conceptualised as a cognitive evaluation of life as a whole (Murtin *et al.* 2015). Further to this, and somewhat in contrast to questions of short-term happiness, the relatively enduring underlying state of satisfaction with life is argued to provide useful information (EFILWC, 2010). This is owing to its capacity to address both cognitive and affective dimensions, with the former focused on one's sense of satisfaction, and the latter with an assessment of life overall (European Social Survey 2015a).

Life-satisfaction is the most widely used means of assessing subjective wellbeing (EFILWC 2010), employed by approaches such as the Gallup World Poll. Surveys in 132 countries use Cantrill's Ladder ask people to identify a current rating of their life-satisfaction on a ladder from one to ten. These results inform initiatives including the World Happiness Report. Similarly, the European Social Survey (2015a, 2015b), European Quality of Life Survey, and SHARE employ the same scaled approach to measure self-reported life-satisfaction, albeit without Cantrill's Ladder.

Owing to the non-prescriptive nature of life-satisfaction questions, Layard (2016) argues this approach to be the most democratic means of assessing individual subjective wellbeing. There is also evidence to support 'ordinal interpersonal comparability', whereby different people reporting similar levels of life-satisfaction can be considered to be experiencing similar levels of wellbeing (Cojocar and Diagne 2015).

However, use of a single indicator of subjective wellbeing is not without its critics. The influential work of Amartya Sen has focused attention on both the functioning and capabilities of people as essential to wellness (Nussbaum and Sen 1993). The former represents the fundamental elements necessary to achieve positive outcomes, yet without the ability to enact them through the latter considerations, people are restricted from fulfilling their potential. A further criticism of life-satisfaction measures relates to the potential for people to adapt to their life circumstances. Social comparison theory (e.g. Voicu 2015) argues that evidence reflects how people adapt to present circumstances (Deaton 2008). However, the consistent pattern for populations from less wealthy nations to report lower average life-satisfaction scores suggests that people are able to distinguish their desired utility from their current situation.

Kapteyn *et al.* (2014) also identify that retrospective judgement of life-satisfaction can be cognitively challenging, only being constructed when requested. Similarly, the order of questions

has potential to affect people's responses (Lee *et al.* 2016). Diener *et al.* (2013) outline the potential for temporarily accessible information to influence respondents if they base their responses on heuristics. It is therefore advised that questions of life-satisfaction are separated from those that ask about the health-status, and any comparison of life-satisfaction results must be based on consistent ordering of questions. Although not without risks, Kroll and Delhey (2013) outline that cumulative evidence demonstrates that such bias, along with measurement concerns of reliability and validity, are potentially overstated.

However, a further issue when considering the potential for decision-makers to make use of the results to improve decision-making is the lack of weighted domains/components. This restricts their ability to identify the importance of different outcomes and where improvements are possible.

4.2 Specific Measures of the Intangible Outcomes of Social Care

While generic measures of wellbeing can provide useful insight, the use of bespoke approaches that are created for particular situations have the potential to yield more effective results (Kaambwa *et al.* 2015). Although aware that no single issue is exclusively responsible for the wellbeing of people (Niedzwiedz *et al.* 2014), for many of those involved in LTC, its effect will be significant.

One well-established and specific approach is the EQ-5D instrument that addresses five issues to measure and indirectly value quality-adjusted health-related quality of life (van Leeuwen *et al.* 2015a). Underpinning the traditional health measures of Quality-Adjusted Life Years (QALYs) and Disability-Adjusted Life Year (DALYs) responses indicate the utility value of health states. When combined with the number of years living in that state, results correspond to a state between zero and one, where the former represents a state equal to death, and the latter being that of full health.

EQ-5D measures the health of individuals, although it has been criticised for potentially underestimating key impacts of initiatives (Huter *et al.* 2016). As such, alternative preference-based measures that consider a broader conception of quality of life (here, the language of quality of life is essentially synonymous with that of subjective wellbeing) have been developed (van Leeuwen *et al.* 2015b). As with EQ-5D, measures such as Adult Social Care Outcomes Toolkit (ASCOT) and ICEpop CAPability measure for Older people (ICECAP-O) ask respondents to identify their current state against a range of issues, employing a four and five-point scale respectively. In addition to providing an overall perspective on wellbeing, the measures include various domains that also provide decision-makers with the increased potential to identify where improvements are possible. Both ASCOT and ICECAP-O were developed in the UK and are conceptually based upon the influential work of Sen, with ASCOT addressing both functional and capability concerns, whilst ICECAP-O focuses exclusively on capability issues. Both approaches are available for both care recipients and those providing informal care, with appropriate variations included (Annex 4

provides an overview of each instrument). Furthermore, while there are long-standing criticisms of subjective measures, both ASCOT and ICECAP-O are at least as reliable as health-focused measures (van Leeuwen *et al.* 2015a). Consistent advice outlines that although feasible to use both EQ-5D and one of the broader measures, it is not recommended that the latter options are used interchangeably (*ibid.*).

4.3 Approaches to Monetising

The process of valuation is something that people do all of the time – making decisions between options, taking into account the potential costs and benefits of different choices. It is also a process that decision-makers in organisational-contexts undergo regularly, whereby experience and personal beliefs can influence perceptions of how valuable different issues are. Although people and organisations value things regularly, there is often a lack of transparency by which these value-judgements are made.

Layard (2016) outlined the need for a common currency to understand concerns of wellbeing. Although not explicitly demanding the monetisation of changes to wellbeing, if decision-making is to have the ability to compare the costs and benefits associated with SI in LTC, as outlined in Richards (2018) there needs to be a common mechanism of accounting for both outcomes with market values, and those without.

Based on principles of welfare economics, valuing intangible outcomes is inherently subjective, and although not immune from criticism, recognition of its importance is growing. This can be seen in the growth of interest in impact investing and developments such as the Dow Jones Sustainability Index incorporating the monetisation of intangible outcomes. Developments in impact investing (e.g. The Impact Management Project 2018) demonstrate extremely strong similarities to the impact questions outlined in Richards (2018) and the Principles of Social Value (Social Value International 2015) that underpin them.

The fundamental purpose of monetising outcomes is twofold. By providing a consistent unit of measurement, outcomes can be compared to one another, as well as to the costs of their production. Although there is increasing awareness of the potential to monetise intangible outcomes, there is no universally accepted method.

Richards (2018) outlines that the Principles of Social Value and the general principles of economic analyses aim to understand value from the perspective of those social actors affected. Therefore, it is feasible that the same outcome has potential to be valued differently based upon the individual or institutional actor experiencing them. Greve (2018) also highlights that monetisation can be challenging.

4.3.1 Cost-based Approaches

Issues such as the time people commit to providing informal care can be valued using cost-based approaches such as opportunity costs relating to the foregone earnings of those of working-age individuals providing care. Many informal carers are however beyond working-age. This is not to say that their time-contribution should not be valued. Leisure time is still valuable, and although decision-makers may decide that an alternative hourly-rate to those working age informal carers is appropriate, it remains important to value the time-contribution of all informal carers if decision-making is to take account of all material issues

In addition to the value to informal carers, there is also cost-based value for institutional actors that requires consideration. When viewed from the perspective of LTC providers/investors the time contributed by informal carers has a corresponding value. Replacement costs are conventionally used to represent the value of avoided expenses required to replace voluntary care with formal services. However, there is the potential to consider further valuable outcomes. Foregone tax income, and potentially increased welfare expenditure will influence decision-making. In addition is the need to consider the impacts on productivity. Even for those informal carers that do not leave employment, there is the risk of presenteeism, whereby job performance reduces, with potential implications for employers and economies. A suitable means to value this outcome is the friction-cost approach where each hour of work is valued until another employee takes over the role. This approach does however have the potential to underestimate the value of the transition process, where a new employee may be less productive for time.

These valuation approaches have the advantage of being reasonably accessible and standardised. However, whilst those representing the value to institutional actors may provide useful indications of institutional value, those for individuals can be argued not to fully represent the value of the impacts on people's lives (Garrido-Garcia *et al.* 2015). As such, is the need to consider alternative means of monetising non-market outcomes (Annex 5 provides an overview of different approaches, not all of which are relevant to LTC).

4.3.2 Stated preference techniques

Stated preference techniques can be based on either contingent valuations or choice modelling methods (Fujiwara and Campbell 2011). The former technique asks respondents to value welfare implications (van den Berg *et al.* 2013), contingent on specific characteristics of that being valued. For example, asking people the maximum they would be willing to pay (WTP) for a positive outcome, or the minimum they would be willing to accept (WTA) to experience negative outcomes. Choice modelling methods on the other hand, asks respondents to rank, choose, or rate alternative options of a good/service, based on varying attributes. Inclusion of the costs for the different options allows respondents to identify their preference, aware of the costs and benefits. For instance, people could be asked to rank/rate varying health or social care initiatives and the resultant outcomes that would be experienced.

Although done regularly, people do not have fixed preferences that are readily accessed. Rather, values are constructed when asked (Oliver 2013). As such, is the potential risk that responses are

unduly influenced by heuristics. Priming effects of the order of questions is often cited (Voicu 2015), as is sensitivity to the scale of changes (Bobinac *et al.* 2013), as well as the potential for results to be affected by the anchoring of irrelevant cues (Fujiwara and Campbell 2011). Also, the ability to value complex issues can be cognitively challenging (Bayrak and Kriström 2015). Similarly, techniques can be subject to non-responses, as well as strategic and protest votes. Strategic responses intend to reduce the usefulness of results. For example, if respondents believe that indicating a WTP for a free service may induce a charge for services. Protest votes can also be employed for similar reasons, and to counter what some may feel are unethical requests to place a 'price' on sensitive issues (Garrido-Garcia *et al.* 2015).

There is also evidence of disparities between WTP and WTA values, with the former consistently lower. This 'valuation gap' (Bayrak and Kriström 2015), can potentially be explained by the 'endowment effect' that suggests that as preferences are context dependent, human nature usually means that losses are of greater significance than gains (*ibid.*). Given the potential constraints on stated preference approaches, it is suggested that measures are taken to reduce the potential limitations. These include careful consideration of the ordering of questions, the use of intervals, rather than asking for definitive points of value, and using validation questions to eliminate less reliable responses (Gao *et al.* 2016).

Although not without criticism, stated preference techniques have been employed by policy makers and researchers alike in health economics for some time. For example, the value of a QALY has regularly been examined from the perspective of those effected (e.g. Martin-Fernandez *et al.* 2014), and less frequently from a societal perception (e.g. Bobinac *et al.* 2013). Meta-studies report there is considerable variation in the WTP value of a QALY, with results suggesting higher results for life-extension, than for quality of life improvements (Ryen and Svensson 2015). In addition to valuing psychological states, stated preference techniques can also be applied to concerns with seemingly objective values. For instance, although accounting costs of hospital beds are regularly cited, Page *et al.* (2017) contingent valuation exercise highlighted the value of beds from the perspective of hospital chief executive officers.

Application of stated preference to aspects of LTC are also available, with studies examining the WTA value for providing informal care. Again, not without potential ethical considerations, results demonstrate general consistency between the two stakeholder groups of those currently providing informal care, and those that do not. Results do however demonstrate a higher WTA to provide care for a hypothetical person, than a loved-one (Garrido-Garcia *et al.* 2015).

4.3.3 Revealed preferences

Rather than asking people to state preferences, revealed preferences are identified through market production and consumption. They can employ traditional approaches to consider the time costs, such as travel expenditure, as well as examining market prices of a similar good or service that would provide broadly consistent changes as those experienced. Alternatively, hedonic pricing can be used to reflect how the price of a market good or service is affected by

different attributes (Fujiwara and Campbell 2011), such as the price of homes in favoured school-catchment areas.

An advantage of revealed preferences is that inferences are based on actual behaviours, and the time and cost requirements will likely be lower than conducting a stated preference approach. However, they assume perfectly competitive markets, and the full knowledge of all in the market place, as well as presuming alternatives to be a perfect substitute.

4.3.4 Wellbeing Valuation approach

Wellbeing Valuation method is a more recent development. Based on the assumption that life-satisfaction is a reliable proxy for individual welfare, marginal rates of substitution represent the amount of income required to provide the same change in utility as provided by a change in life circumstances (Fujiwara 2013). Therefore, rather than asking people directly for the value of changes, or using existing market prices, value is inferred through large-scale panel-data surveys.

Publicly-available results from the UK present values for a range of outcomes across segmented age-categories, and some geographical areas, based upon the average values for these sections of society (e.g. Fujiwara *et al.* 2014, Fujiwara and Vine 2015, HACT 2016). This can make the wellbeing approach more time and cost effective than alternatives such as WTP exercises, although has the potential risk that values do not represent those of the particular stakeholders of interest.

Essentially, this valuation approach is designed to understand causal relationships between outcomes, income, and life-satisfaction. Through a ‘simultaneous model methodology’ the marginal rate of substitution (MRS) is identified that an individual would require as a change to their income to provide the equivalent improvement, or to compensate for the detriment to their wellbeing as experienced from a change to their life circumstances (Fujiwara, ND). For example, if an individual were to report that a particular LTC intervention improved their life satisfaction by 10%, it is possible through statistical estimations to identify how much income they would require to experience the same increase in life satisfaction. Annex 6 provides more detail on this methodology.

Whilst this methodology has been employed to a range of situations, there remains a fairly limited set of publicly available data, although bespoke results can be obtained through the lead proponents of the technique. Although developed in the UK, the methodology has been applied in both Australia² and Hong Kong³.

Although offering a valuable contribution to the debate on the valuation of non-market outcomes, there are limitations. The current data is based upon large-scale panel surveys, and although results provide some segmentation, they still highlight effects on wellbeing for the average person (within each broad segment) in a society. Some of the outcomes represent changes that are

² See <https://asvb.com.au/>

³ See <https://hkisia.wordpress.com/2015/04/20/hki-sia-commissioned-groundbreaking-research-on-wellbeing-valuation/>

naturally binary, for instance someone is either in full-time employment or not. However, many others, such as changes to levels of anxiety or depression, are not situations where everyone will likely experience either the worst or best possible states. In these cases, it is important to understand the marginal value of changes.

Recent developments provide results that also address the diminishing rates of substitution. This development addresses Layard's (2016) requirement that weightings be applied to different levels of wellbeing, and Annex 7 demonstrates these results. Using the short version of the Warwick-Edinburgh Mental Wellbeing Scale (SWMBWBS), monetary values demonstrate the relative importance of changes between any two points on the Scale (Trotter and Rallings-Adams 2017). Using seven positively worded statements (the full WEMWBS version has 14 statements), each with five options to reflect an individual's experience of the statements (with corresponding scores of 1 to 5), SWEMWBS reflects the current mental health/wellbeing of people from a score of 7 to 35. At the time of writing, results illustrating diminishing value were only reported for the average UK citizen, although the intention to segment based on the consistent categories of age and geography are outlined.

Ideally, counterfactual evidence would be gained from random controlled experiments, although asking individuals to consider the statements at varied points in an intervention can also provide useful evidence (*ibid.*). If applying the latter approach, the authors also suggest application of an average deadweight figure of 27% from the UK Homes and Communities Agency's Additionality Guide (2014).

4.4 The Outcomes of Social Investment in Long Term-Care

Greve *et al.* (2018) identified the dimensions of impact that serve as the common criteria for different social investors to consider the relevant outcomes of LTC.

Framed by the broad dimensions of economic return, social impacts, and risk, relevant areas were identified that addressed concerns of health and wellbeing, labour market implications, and substitution effects. In addition to identifying what is relevant, it is also necessary to measure and value to appreciate the significance of outcomes. Only with both assessments can decision-makers be confident that they have information about material issues that can improve their decision-making.

Barbieri and Ghibelli (2018) demonstrated the multiple outcomes of SI in LTC and addressed in detail the substitution and labour market effects of informal care. Although potentially challenging, if such concerns can be measured effectively, they are more readily quantified and monetised. However, in accordance with Knapp's (1984) Production on Welfare (PoW) framework (and recent developments in Evans-Lacko and Knapp 2017), there is also an imperative to examine how LTC address care recipients' welfare shortfalls. The paradigm of SI in LTC also focuses attention on preventing, as well as repairing/minimising effects on individuals' welfare-which align with dimensions of impact by Greve *et al.* (2018). The PoW model can consider such concerns as

part of the 'normative endeavour' of social care (Knapp, 1984). Similarly, the PoW model was extended by Malley and Fernandez (2010) to recognise the stress on informal care providers as an additional stakeholder.

Broadly, these concerns can be conceptualised as relating to wellbeing, and present additional challenges for measurement. Therefore, the following sections of this report examine relevant evidence to provide illustrative examples to measure and value outcomes of LTC using informal care as the context.

5 Monetising Outcomes of Long Term-Care – Illustrative Examples on Informal Care

There is widespread recognition of the importance of informal care. Ghibelli *et al.* (2017) identified the scale of informal care, while Barbieri and Ghibelli (2018) also outlined financial and social effects of substituting formal care provision with informal care. This awareness contrasts the conventional position of viewing informal carers as a free resource that delivers cost savings for policy makers (van den Berg *et al.* 2014). This report extends the conceptual work and demonstrates the potential to monetise the effects of informal care. The examples presented are not exhaustive of the issues to consider, and do not relate to the same interventions so cannot be considered as outcomes to be aggregated. The intention is to demonstrate the potential to measure and value such changes generally. The analysis is structured by the impact questions that were identified in Richards (2018). Designed to identify the key requirements that can aid decision-making, they also form the basis for the subsequent recommendations presented in section 6.

5.1 Valuing Labour Market Effects

Notwithstanding the importance to also consider macro-labour-market effects such as the friction costs of lost productivity, labour market effects on individuals traditionally uses opportunity costs as suitable financial proxies that represent foregone income. The two examples below use existing national and pan-national data, as well as modelled assumptions to illustrate the potential to value these micro-level effects.

Bolin *et al.* (2008) report upon a reduction in employment probability of 3.7% for informal carers aged 50+, when the care they provide is increased by 10% across 8 European nations (3.2% for males and 2.8% for females). Therefore, based on available data it is possible to represent the value of these negative outcomes experienced by informal carers. The same authors also reveal a reduction of hours worked by 2.6% when informal care is increased by 10%.

Here we outline the process undertaken and display included estimations. We outline the potential to identify necessary data for the average individual and aggregated gross financial effect to those informal carers affected. Initially the discussion outlines estimations of the quantity of those affected, followed by issues of individual financial value.

Riedel and Kraus (2011) provide information on the proportion of the population aged 50+ that provide informal care (the figures used excluded care to a child), with the population size (aged 50 – 64 to represent those of working age) for relevant nations sourced from the International Labour Organization (ILO 2018a). However, there is a lack of robust evidence on the employment rates of informal carers. Therefore, informal care rates were multiplied by national employment rates for those aged 55-64 (Eurostat 2018). Although these average rates may not account for endogenous factors that may impact upon the rates for informal carers, they do provide data segmented by sex. Therefore, estimates of the relevant national populations was calculated by:

*Relevant proportion of population that provide informal care * proportion of those aged 50+ that provide care to family other than children * employment rate for those aged 55-64*

As an example, based on this approach the results for Denmark were;

$$1,111,000 * 5.3% * 88% * 67.8% = 35,132$$

Employment income levels are unavailable for informal carers specifically. Therefore, to avoid over-claiming, for each of the nations included in Bolin *et al.* (2008) research, both the average national employment incomes, and the lowest average salaries by occupation type were used (ILO 2018b). For example, in Denmark the lowest average salary is received by those people working in ‘services and sales work’. Although these figures serve as reasonable proxies for the value lost by leaving the employment market, they are gross figures and there would be likely be variable levels of income received by the individual through welfare benefits.

Similar estimations have been applied to appreciate the effect of reducing the number of hours worked. Calculations based on average annual income levels, incomes from the lowest occupation type, and minimum wages where applicable are reduced by 2.6% based on the evidence from (Bolin *et al.* 2008) on the reduction in hours worked. For minimum wage calculations, the OECD (nd) identifies the average number of hours worked through declared dependent employment, and these are reduced by the percentage of fewer hours worked. Although not necessarily representing the net loss in income, results serve as reasonable representations of the value.

Annex 8 provides a summary explanation of the variables included against impact question 3 to 9, and Annex 9 highlights estimates for those nations included in the paper from Bolin and colleagues (2008). Although evidence of average effects is presented for changes in the provision of informal care, there is no indication of how many carers are likely to experience the necessary conditions for the change to occur. Therefore, modelled estimations are provided that account for 0.1% and 1% of the relevant population to demonstrate different potential scales of impact.

5.2 Valuing Substitution Effects

Avoiding the need for (or potentially further) formal health and/or social care is conventionally valued using replacement costs associated with the avoided care. However, there is the potential to value using WTP from the perspective of relevant management (Page *et al.* 2017).

Wolff and Kasper (2004) report upon the effect of informal care on the delayed discharge of females aged 65+ from hospitalisation in Baltimore (using an age-stratified random sample of 420 people). Delayed discharge was identified as delays of at least two additional days compared to the average length of stay for the same diagnostic related group.

Although based on a small sample size, results demonstrated that 24.3% of the sample were hospitalised in year 1 (102.06 people), with an average of 29.6% being delayed (30.21 people). Providing a reasonable baseline estimation, significant results based upon multivariate models⁴ are highlighted for particular segments of care recipients (Annex 10). Identifying informal carers characterised as competent, overloaded, and captive in their roles highlights the relative impact of the different states.

The World Health Organisation (WHO) (2011) provide average unit costs for a hospital bed in the US as \$628.43 (2008 prices), with calculations in Annex 11 based on 2 days delayed discharge presented for the sample of 420 people. The results demonstrate that those informal carers that can be characterised as 'competent' create positive value through reducing delayed discharges, whilst those that are 'overloaded' or 'captive' there is negative value when compared the baseline average.

Wolff and Kasper (2004) also indicate that care recipients with informal carers that are competent and involved in their healthcare are also more likely to be hospitalised. Such evidence should certainly be further examined and included in any decision-making, but given the lack of additional information these have not been valued in this report.

5.3 Valuing Health and Wellbeing

To identify the worth of interventions on health and wellbeing outcomes, it is important to establish values that can act as suitable representations of the total worth of these outcomes from which the significance of the effects can be calculated. The requirement for relative values to be identified from the perspective of those stakeholders affected by activities is fundamental to SROI and economic analysis generally. Here we use existing evidence to examine the potential of both stated preference and Wellbeing Valuation approaches and present different options that can be

⁴ Adjusted for year, age, income, living arrangements, self-reported health status, task limitations, musculoskeletal, neurological and cardiopulmonary conditions, depression, sensory impairment and diabetes, but not for other caregiver characteristics

used to identify the value of the effects on people. Revealed preferences are not considered owing to the lack of opportunity to reflect values that stakeholders would identify.

5.3.1 Approaches to value health and wellbeing: Stated preference

Stated preference approaches are reasonably traditional means to valuing the health-status of people, although there is no universal value that is applied. Robinson *et al.* (2013) highlighted variety in the value of a QALY. Results for 9 European nations are presented based upon standard gamble and time trade-off approaches for two EQ-5D questions. The trimmed⁵ average across all nations was identified as \$18,247 to \$34,097 (converted to US\$ using PPP in 2010). Similarly, Gyrd-Hanse and Kjær (2012) report upon a Danish study identifying a range of 20,404 DKK (€2,720) to 722,763DKK (€96,366) (2005 prices), from aggregated and disaggregated approaches respectively⁶. For the purposes of this report, these latter values will be employed.

Netten *et al.* (2002) also identified stated preference valuations for the domains of Social Care Related Quality of Life (SCRQoL) (discussed in detail below). Results for weekly thresholds for financial amounts required to accept needs being unmet provide annual values between £60,008 to £69,108 (Annex 12). Identified as both credible and reliable (*ibid.*) results are higher than costs, suggesting a positive net result. The authors however urge caution, mindful that WTA values are usually higher than WTP, and given the nature of interventions, the latter is the preferred methodology. Burge *et al.* (2010) also identified WTA values using discrete choice experiments for the ASCOT domains. In this case £787 per week was identified as the average WTA value to substitute the best situation for control over daily life, with the research also identifying potentially important segments of service-users based on their location.

Whilst the above studies consider stated preference values from the perspective of those stakeholders affected, the National Institute for Clinical Excellence (NICE) provide a WTP threshold from an institutional perspective. The UK government employ a value of approximately £30,000 as the ceiling for treatments that provide a QALY. Although use of this value has been criticised for a lack of empirical basis (Donaldson *et al.* 2010), it is widely employed within the UK to inform decisions.

5.3.2 Approaches to value health and wellbeing: The Wellbeing Valuation approach

This section outlines two options to apply the Wellbeing Valuation methodology. The first uses the ASCOT framework and the second examines the developments associated with changes recorded by the SWEMWBS.

UK local authority social service-users annually rate their current state against seven of the SCRQoL domains (dignity is not assessed), with four possible responses representing best to worst scenarios.

⁵ Trimming is a standard convention, and in this study the highest 1% of responses were trimmed.

⁶ Aggregated approaches assess the ratio of the means, whereas disaggregated methods calculate the averages of the individual WTP per QALY.

Preference weights identified by social care service-users (based on a sample of home-care service-users) demonstrate the relative value of the eight domains of SCRQoL for each of the four states (Netten *et al.* 2012). The most significant domain, as identified by service-users, control over daily life has a range of zero to one. This indicates the lowest level as equivalent to being dead, and the best comparable to perfection. The other domains are not as significant, with neither highest nor lowest levels being comparable to states equivalent to perfection or death. This does not suggest that the other domains are not important to people, but it makes clear that both best and worst situations for the other domains do not have the same importance to people.

Control over daily life's preference weighting provide the potential to value this domain in relation to the wellbeing valuation approach. As outlined, until the most recent development in wellbeing valuations using SWEMWBS, all values were presented as binary concerns. This is identical to the situation of a zero to one weighting between levels one to four.

Within the valued outcomes available from the Wellbeing Valuation approach is one titled 'feel in control of life.' This is a very similar descriptor to the domain within ASCOT. Both measures employ a four-point scale, although whereas ASCOT provides four statements that ask respondents to describe how much control people have over their daily life, HACT (2016) has a statement that states; 'I feel that what happens to me is out of my control' with options of 'Often, Sometimes, Not often, Never' provided. Testing would be required to ensure consistency and that values could be reasonably transferred, but there is an intuitive similarity between the two that suggests potential opportunity.

The Wellbeing Valuation for 'feel in control of life' has a value of £16,427 per annum for someone aged 50+ in an unknown geographical location (*ibid.*). Given the ASCOT weighting of 1 assigned to level one, it is therefore possible to apply this value to the ASCOT domain for someone who reports their current status as being at level one (i.e. the best state). Weightings are available for each SCRQoL domain-level. Therefore, assuming a linear relationship between the domain-levels monetary values for each level are revealed by multiplying the corresponding weighting by the monetary value for level one. For example, level two of the control over daily life domain is weighted at 0.919. Therefore, 91.9% of the anchor value of £16,427 reveals a value of £15,096.

This approach can be applied to all SCRQoL domains, using the control over daily life as the anchor value from which all others are calculated. For example, level one in the domain of occupation has a weighting of 0.962, resulting in 96.2% of the anchor value being a useful proxy (£15,803). Annex 13 presents the results of this approach.

Someone who reports the best possible state for all domains, would have an aggregated value of £118,522. This represents the 'best' annual value of SCRQoL for a social care service-user in the UK. Alternatively, someone reporting the worst possible levels for all domains would have a corresponding value of £23,902.

There are some differences in the included domains within the informal carers SCRQoL. However, control over daily life is present, and this provides the consistent means to create an anchor value with the life-satisfaction valuation methodology. Although there is potential for different preference weights for informal carers to those of care recipients, Netten *et al.* (2012) identified

the weights of service-users agreed with the general population. Therefore, although further work could refine the preferences of informal carers, the same methodology can be applied. There is also the potential to employ values from different age-categories as appropriate from the Wellbeing Valuation data bank (i.e. under 25, 25-49, and 50+).

5.4 Monetising the Health and Wellbeing Effects of LTC Initiatives for Informal Care

Options that can be employed to identify the marginal value of LTC interventions on health and wellbeing were provided above. Annex 14 summarises the reference values employed to demonstrate this potential, with values standardised to 2017 and presented in US dollars. Purchasing power parity (PPP) was employed to convert the results to a standard currency (US\$) (OECD 2018). These results were then converted to 2017 values using the consumer price index to provide consistent results⁷. Not all studies have been utilised in this report. Those selected illustrate a range of possible values and are intended to demonstrate the theoretical potential to value LTC outcomes, not provide definitive results to be used by decision-makers.

Where interventions identify effects on health or wellbeing it is important to be aware of the scale for each measurement approach. For example, EQ-5D has a range of 0 to 1, whilst SWEMWBS scores between 7 to 35. The optimal result for these and other measures will therefore have a value equivalent to the totals identified above, from which effects of initiatives represent a proportion of the total possible value.

Table 1 presents an example of how to value informal care to those in receipt of care. The table also presents two different LTC initiatives to support informal carers. All three examples apply each of the examined approaches to monetisation. These are illustrative and hide the complexity that is behind much of the cited work that has produced the different values. These are framed by impact questions 2 to 9 as identified in Richards (2018) (impact question 1 considers the problem, and question 10 asks decision-makers to identify the changes that are most important and require management).

Barnay and Juin (2016) identify that the provision of informal care reduces the risk of depression by 42% for care recipients using the Mental-Health Inventory (MHI-5). Although such evidence is potentially useful at the macro-level, it does not provide evidence of the amount of change experienced per individual. Therefore, for illustrative purposes, other evidence presented in the paper that highlights a reduction of 1.8 points on the MHI-5 for each one-unit increase in informal care is assessed.

Both studies with evidence of the effects of LTC interventions use the SF-36 scale to measure mental health. The first study examined the provision of telephone support to informal carers of stroke survivors after rehabilitation compared to those not receiving the service (Grant *et al.*

⁷ Website used; <https://westegg.com/inflation/>

2002). The second study analysed effects of support to families of stroke victims (Mant *et al.* 2000). The studies assessed informal carers at 13 weeks and 6 months respectively – therefore, without better evidence, values are presented for 6 and 12 months for all studies. Both studies show negative effects, and results are presented in US\$ for 2017.

Table 1: Illustration of how to value of informal care LTC interventions on health / wellbeing – linear values

What is the solution?	Who changes?	What changes?	How is the change measured?	How much change occurred? (per individual)	How much change is down to the intervention?		How long does the change last?	What is the relative importance of the change (per individual)?			
					Counterfactual	Attribution		WTP UK government	WTP patient lower boundary	WTP patient upper boundary	Wellbeing valuation
Informal care provision – one-unit increase in provision	French care recipient aged 65+ reporting at least one ADL/IADL	Reduction in depression	Mental-Health Inventory scale	Average of 1.8-unit decrease	Ordinary least square and instrumental variables employed to control for endogeneity (included within how much change occurred column)	No data available	6 months	\$190	\$294	\$943	\$765
							12 months	\$380	\$587	\$1,887	\$1,531
Initial face to face interactions with nurse,	Family members providing informal	Mental health / wellbeing	SF-36 mental health variable	Average of 17.07 decrease	7.69 identified through random control trials	No data available	6 months	\$989	\$139	\$4,881	\$3,988

followed by telephone calls to provide problem-solving support	care to stroke survivors						12 months	\$1,978	\$277	\$9,831	\$7,976
Family support from family-support organiser - with activities centred on needs of family	Informal carers of stroke survivors – not including those survivors in residential care	Mental health / wellbeing	SF-36 - average of the 9 domains	4.56 out of 100	The effect of the initiative is the difference between experiment and control groups	No data available	6 months	\$480	\$67	\$2,387	\$1,937
							12 months	\$961	\$135	\$4,775	\$3,874

The results demonstrate the potential to value. However, they assume a linear relationship between the outcome and value. Recent developments that value responses to SWEMWBS on a diminishing marginal utility model provides potential for application to SI in LTC.

Mukuria *et al.* (2014) at the request of the UK Department of Health identified exchange rates between different measures of health and wellbeing. Psychometric analysis assessed relationships between different measures using summary statistics and correlations. In addition, factor analysis examined if they addressed more than a single dimension, whilst regression analysis (eta squared) identified the capacity of wellbeing measures to distinguish between people with differing health needs (*ibid*).

Complementing this work, Layard (2016) presents the marginal effects of a one-unit change in different measures to that of the life-satisfaction scale. When this is multiplied by a one standard deviation change in each explanatory variable computed effects are comparable. Annex 15 presents the exchange rates, identifying the effect of a change in measures on other options where the data is available. For example, each unit change in SF-36 results in a 1.07 change in SWEMWBS.

If the baseline score on SWEMWBS is known, the amount of change can identify the corresponding proportion of value created by an intervention. For example, someone starting from a baseline of 14 points with a change of 10.7 points would now have a score of 24.7. This is equivalent to 86% of the total value (Annex 7) (Fujiwara *et al.* 2017). To model this approach, 86% of the different reference values can be identified as suitable proxies. Table 2 applies this illustrative approach, and is again not a presentation of results that should be used for decision-making, but illustrates the theoretical potential for valuation of relevant outcomes.

As no exchange rates are known for the MHI-5 scale, the evidence from Barnay and Juin (2016) cannot currently be applied to the diminishing utility calculations. However, for the two studies on the effects of informal care interventions, results represent changes of 9.38 and 4.56 on the SF-36 scale. With no baseline information available, results are presented for three different options assuming one-year of value being experienced. Baselines of 7, 15 and 26 on SWEMWBS are employed, representing the lowest possible score, the lowest that is identified as having any value, and that which is the mid-point of scores that have assigned value (half way between 16 and 35) respectively.

Table 2: Illustration of how to value of informal care LTC interventions on health / wellbeing – diminishing values

What is the solution?	Who changes?	What changes? And How much change occurred? (per individual)	What is the relative importance of the change (per individual)?											
			WTP UK government			WTP patient lower boundary			WTP patient upper boundary			Wellbeing valuation		
Activity	Stakeholder segment	Outcome	Baseline 7	Baseline 15	Baseline 26	Baseline 7	Baseline 15	Baseline 26	Baseline 7	Baseline 15	Baseline 26	Baseline 7	Baseline 15	Baseline 26
Initial face to face interactions with nurse, followed by telephone calls to provide problem-solving support	Family members providing informal care to stroke survivors	Mental health / wellbeing 9.38 (17.07 – 7.69)	\$16,577	\$18,074	\$2,025	\$2,325	\$2,535	\$284	\$82,378	\$89,819	\$10,061	\$66,836	\$72,873	\$8,163

Family support from family-support organiser - with activities centred on needs of family	Informal carers of stroke survivors – not including those survivors in residential care	Mental health / wellbeing 4.56 out of 100	\$0	\$13,835	\$1,286	\$0	\$1,941	\$180	\$0	\$68,753	\$6,393	\$0	\$55,782	\$5,187
---	---	--	-----	----------	---------	-----	---------	-------	-----	----------	---------	-----	----------	---------

6 Conclusions

This report has extended previous work produced in SPRINT which identified the principled-approach necessary to account for the consequences of LTC interventions, as well as highlighting the relevant outcomes for multiple social investors. Using illustrative examples focused on informal care, this report has demonstrated the potential to measure and value these outcomes.

Measuring and valuing outcomes is intended to inform decision-makers as to the significance of different outcomes. This is essential for decisions to take account of material effects of investment options and compare the costs and benefits of different options. The principles of social return on investment (SROI) and economic analysis intend to guide how this information is collected, and both require that measurement and valuation is informed by those social actors affected by decisions.

The impact questions structure what data is required. At present there remain significant gaps in current evidence on the effects of LTC interventions, and Tables 3, 4 and 5 highlight where greater information is required. However, across the dimensions of impact identified in earlier SPRINT research, evidence was used in this report to provide illustrative examples of the potential to analyse the significance of stakeholder outcomes. Although still requiring further research, in both labour market and substitution effects, some rigorous data was identified that demonstrates the extent of outcomes experienced by those providing and receiving informal care. There is however less available evidence of health and wellbeing outcomes, and this has been highlighted as an area requiring significant development. In addition to collecting appropriate data from highly rigorous approaches such as those employing random control trials, there is potential value in considering alternative approaches such as real-world data (e.g. <https://rwe-navigator.eu/use-real-world-evidence/sources-of-real-world-data/>). Health interventions have been analysed using such approaches, where data is sourced from alternative sources such as existing patient registers. In a similar vein, this report has demonstrated that available evidence that may use values from different years and currencies has the potential to be applied to other nations via appropriate conversions. Although offering potentially more rapid assessments, it remains important to ensure that any such approaches maintain a focus on stakeholder-informed evidence in line with the requirements of economic analyses and the principles of SROI.

This report has used data to demonstrate the potential to measure and value relevant of health and wellbeing. There are a variety of indicators that can measure these concerns, and this makes it more difficult to compare alternatives. However, research and practice to provide consistent and meaningful data is developing. Studies that illustrate exchange rates between different measures of health and wellbeing improve opportunities to understand the relationship between them.

When wanting to identify where improvements are possible, it is also important to appreciate more than the average results of wellbeing assessments. Life-satisfaction measures are one of the most commonly used indicators, and although providing useful evidence of the overall effect of

LTC interventions, they do not provide insights as to where improvements can be made. However, some measures of wellbeing such as ASCOT provide more than a single indicator. Overall results are the consequence of different domains scores, and this information alone can inform where LTC interventions are having positive effects, and where they are less effective. Research to test the application of ASCOT into other countries that appreciates the relevant preference-weightings is also ongoing.

Measurement of outcomes is important, but as highlighted the purpose of valuing different outcomes of any activities is to provide a consistent unit by which comparisons can be made. This will improve decision-makers' ability to compare different outcomes that may also have differing extents of change. When using the language of money, comparison to the costs of production can also be made. For outcomes of SI in LTC relating to labour market and substitution effects, the appropriate means to monetise are relatively straight-forward. Opportunity and replacement costs can represent the value to different stakeholders and are often readily available through existing data sources. Further research is however required to also include macro-level labour effects such as the potential friction costs of lost productivity.

How to value intangible changes such as the health and wellbeing of informal carers is however less immediately obvious. Different options to value such changes were presented, and although the illustrative examples are not the results that decision-makers should currently use, they do show the potential to understand how important the different outcomes are to those affected. Appreciating the subjective nature of valuation, stated preference and Wellbeing Valuation approaches were outlined as allowing stakeholders to inform the values that are used by in decision-making.

A key issue is the lack of understanding of the unit value of changes to health and wellbeing. The value of changes is therefore often viewed as having a linear relationship. Although there are some advances in this area, there is still the need for further research to be more sensitive to the relative value of different amounts of changes experienced.

Although accuracy of valuations is important if wanting to ensure a satisfactory return on investment is produced, it is more important to appreciate the relative value of different outcomes in relation to one another. Rather than striving for what could be considered false-precision, information about how important different outcomes are in relation to one another can provide decision-makers with information that can assist their decisions about resource allocation. Essentially, the data needs to be fit-for-purpose, commensurate to the audience and decision the evidence seeks to support.

Developments such as those that have provided weightings to the different domains of health and wellbeing measures again provide decision-making processes with useful information. Care recipients in the UK have identified the relative importance of ASCOT domains, and this provides an understanding of which elements are of greatest significance. When this information is combined with monetary values and data on the changes experienced by stakeholders, decision-

making will be more informed as to which stakeholder segments are best served by different LTC interventions.

Value is a subjective construct, and for decisions to take account of outcomes, those making the decisions need to agree on the values that are used. It is essential that values are informed by those affected by decisions, but for these to influence decisions about the allocation of LTC resources, they also need to be credible. LTC is a complex system involving multiple, and often competing motivations of social actors. There are also considerable pressures on those making decisions to ensure equitable and sustainable LTC provision. Notwithstanding the need for greater evidence of the effects of SI in LTC, this report has demonstrated the potential for outcomes to be measured and valued. Although there is still the need for more evidence on the outcomes of SI in LTC, and a requirement for decision-makers to agree on the valuation approaches used, this report has illustrated how doing so can provide improved insight. The challenge to introduce and embed this approach across national LTC systems would undoubtedly take considerable time and resources. However, the potential to measure and value LTC interventions as pilot studies is something that could help create short-term benefits and influence ex-ante decision-making about the allocation of resources.

Table 3: Labour-market effects; recommendations

Impact question	Current evidence	Improvements required
3. Who experiences changes?	Some evidence is segmented to different stakeholders. Much is limited to broad stakeholder groups such as informal carers	Segmentation of informal carers possible by range and combination of characteristics including age, sex, occupation type, health-status (carer and informal care), living arrangements, intensity of care provided
		Greater segmentation of individual stakeholders. This is potentially affected by all subsequent impact questions - with improved data collection highlighting differences in experiences that leads to improved understanding of the initiatives that work best for different stakeholders
4: What changes are experienced?	Some effects on hours worked and probability of remaining in the employment market for informal carers is available	Segmentation of informal carers by changes experienced
		Greater understanding of the interrelationship between immediate labour-market effects impact upon later concerns such as pension entitlements.
		Improved understanding of short and longer-term macro-level effects on state and municipal departments
5: How can we measure the changes?	Changes measured using large-scale data from sources such as SHARE, or employing random control groups	Segmentation informed by the results of this question - see impact question 6
6: How much of each change has happened?	Average reduction in hours worked and probability of leaving the employment market is available for some nations	Segmentation by scale of changes experienced
	Some segmented data by age, sex, or living arrangements available	Baseline data required - appropriate employment rates and average number of hours worked before intervention/increase in care provision
7: How much of each change	Limited data accounting for counterfactual evidence	Segmentation by different levels of counterfactual and attribution

is caused by our activities?	Attribution levels unlikely to be examined	Consensus required for different stakeholder segments. Conduct appropriate studies to identify counterfactual evidence - control groups and/or develop/use statistical evidence - ask stakeholders with experience, or likely to experience changes
8: How long do we need to measure the changes for?	Evidence of duration of changes often unknown, or limited to last data collection point	Segmentation by different lengths of time that experiences last Longitudinal evidence required - conduct appropriate studies - control groups and/or develop/use statistical evidence - ask stakeholders with experience, or likely to experience changes
9: What is the relative importance of the different changes?	Opportunity costs largely available based on average national earnings - by sex / age / occupation	Segmentation by different opportunity costs Appropriate opportunity costs needed for the different occupations of informal carers

Table 4: Substitution effects; recommendations

Impact question	Current evidence	Improvements required
3. Who experiences changes?	Some evidence is segmented to different stakeholders. Much is limited to broad stakeholder groups such as informal carers	Segmentation necessary by range and possible combination of characteristics including age, sex, occupation type, health-status, living arrangements, intensity of care provided
		The necessary segmentation is potentially affected by all subsequent impact questions - with improved data collection highlighting differences in experiences. Improved data collection and analysis can lead to improved understanding of the initiatives that work best for different stakeholders
4: What changes are experienced?	Some evidence of the effects of the relationships between informal and formal care, and the effect of preventative interventions.	Continued development to understand the effects of a LTC intervention, and combination of interventions on alternative formal care
		Greater understanding of the effect of preventative measures. Ex-ante modelling required to estimate the effects as well as longitudinal studies of the effects of interventions
5: How can we measure the changes?	Changes measured using random control groups	Segmentation informed by the results of this question - see impact question 6
6: How much of each change has happened?	Results often presented in comparison to population average	Segmentation by comparison to those in similar circumstances such as being in receipt or providing informal care
		Baseline data required - appropriate employment rates and average number of hours worked before intervention/increase in care provision
7: How much of each change is caused by our activities?	Limited data accounting for counterfactual evidence	Segmentation by different levels of counterfactual and attribution

	Attribution levels unlikely to be examined	Consensus required for different stakeholder segments. Conduct appropriate studies to identify counterfactual evidence - control groups and/or develop/use statistical evidence - ask stakeholders with experience, or likely to experience changes
8: How long do we need to measure the changes for?	Evidence of duration of changes often unknown, or limited to last data collection point	Segmentation by different lengths of time that experiences last
		Longitudinal evidence required - conduct appropriate studies - control groups and/or develop/use statistical evidence - ask stakeholders with experience, or likely to experience changes
9: What is the relative importance of the different changes?	Replacement unit costs of the avoided care provision available in many nations	Consistent data collection of relevant costs segmented by national/regional/institutional setting

Table 5: Health and wellbeing effects; recommendations

Impact question	Current evidence	Improvements required
3. Who experiences changes?	Limited segmentation of care recipients or informal carers	Segmentation necessary by range and possible combination of characteristics including age, sex, occupation type, health-status, living arrangements, intensity of care provided
		The necessary segmentation is potentially affected by all subsequent impact questions - with improved data collection highlighting differences in experiences. Improved data collection and analysis can lead to improved understanding of the initiatives that work best for different stakeholders
4: What changes are experienced?	Limited robust evidence of the effects of specific interventions - evidence of wellbeing effects particularly limited	Segmentation needs to be informed by changes experienced
		Need to better understand the health and wellbeing effects of different interventions
		Initial qualitative data collection required, and systems to regularly ensure relevant outcomes being assessed
		Potential to apply existing specific LTC measures such as ASCOT, although these still require local testing
5: How can we measure the changes?	Some application of health and wellbeing measures - various options available with demonstrable validity and reliability	Segmentation informed by the results of this question - see impact question 6
		Decision-makers need to agree on measures to inform decision-making – these should consider both objective and subjective indicators and include the provision of domains that enable decision-makers to understand the areas where improvements are possible
	Increasing potential to calculate exchange rates from one measure to another to aid comparability	Where different measures are being used, there is the need to continue development of

		exchange rates, examining any important differences between stakeholder segments
6: How much of each change has happened?	Where data exists, generally presents average changes in health or wellbeing	Segmentation can be informed by the scale of changes experienced
		Baseline data required - appropriate health and wellbeing levels before interventions
		Frequent quantitative data collection and analysis systems required to identify level of changes experienced segmented appropriately
7: How much of each change is caused by our activities?	Limited data accountings for counterfactual evidence and attribution levels unlikely to be examined	Segmentation to be informed by different levels of counterfactual and attribution
		Consensus required for different stakeholder segments. Conduct appropriate studies to identify counterfactual evidence - control groups and/or develop/use statistical evidence - ask stakeholders with experience, or likely to experience changes
8: How long do we need to measure the changes for?	Evidence of duration of changes often unknown, or limited to last data collection point	Segmentation by different lengths of time that experiences last
		Longitudinal evidence required - conduct appropriate studies - control groups and/or develop/use statistical evidence - ask stakeholders with experience, or likely to experience changes
9: What is the relative importance of the different changes?	Options available to directly or indirectly value changes	Segmentation to be informed by different preferences
		Need to involve stakeholders to develop appropriate preferences for different outcomes for the stakeholder segments
		Consensus required from decision-makers that values are reasonable and can inform decision-making – viable options to employ stated preference or Wellbeing Valuation approaches

7 References

- Barbieri D, Ghibelli P (2018) *Formal Versus Informal Long-term Care: Economic and Social Impacts*, SPRINT Working Paper D4.4, SPRINT, Brussels.
- Barnay T, Juin S (2016) Does home care for dependent elderly people improve their mental health? *Journal of Health Economics*, 45, 149-160.
- Bayrak OK, Kriström B (2015) Is there a valuation gap? The case of interval valuation, *Centre for Environmental and Resource Economics Working Paper*, 13.
- Bobinac A, Van Exel NJA, Rutten FFH, Brouwer WBF (2013) Valuing QALY gains by applying a societal perspective, *Health Economics*, 22, 1272–1281.
- Bolin K, Lindgren B, Lundborg P (2008b) Your next of kin or your own career? Caring and working among the 50+ of Europe, *Journal of Health Economics*, 27, 3, 718–738.
- Brimblecombe N, Fernandez JL, Knapp M, Rehill A, Wittenberg R (2018) Review of the international evidence on support for unpaid carers, *Journal of Long-Term Care*, forthcoming.
- Burge P, Netten A, Gallo F (2010) Estimating the value of social care, *Journal of Health Economics*, 29, 883 – 894.
- Cojocarua A, Diagne MF (2015) How reliable and consistent are subjective measures of welfare in Europe and Central Asia? Evidence from the second Life in Transition Survey, *Economics of Transition*, 23, 1, 75-103.
- Commission of the European Communities (2009) Communication from the Commission to the Council and the European Parliament: GDP and beyond; Measuring progress in a changing world.
- Deaton A (2008) Income, health and wellbeing around the world: evidence from the Gallup Poll, *Journal of Economic Perspectives*, 22, 2, 53-72.
- Diener E, Inglehart R, Tay L (2013) Theory and validity of life satisfaction scale, *Social Indicators Research*, 112, 497–527.
- Dolan P, Layard R, Metcalfe R (2011) Measuring subjective wellbeing for public policy: recommendations on measures, *Centre for Economic Performance, Special Paper No. 23. Report to the ONS*.
- Donaldson C, Baker R, Mason H, Jones-Lee M, Lancar E, Wildman J, Bateman I, Loomes G, Robinson A, Sugden J, Pinto Prades JL, Ryab M, Shackley P, Smith R (2011) The social value of a QALY: raising the bar or barring the raise? *BMC Health Services Research*, 11 , 8.
- European Foundation for the Improvement of Living and Working Conditions (2010) *Subjective Well-being in Europe: Second European Quality of Life Survey*, EFILWC, Dublin.
- European Social Survey (2015a) *Measuring and Reporting on Europeans' Wellbeing: Findings from the European Social Survey*, ESS ERIC, London.

- European Social Survey (2015b) *European's Personal and Social Wellbeing; Topline Results from Round 6 of the European Social Survey*, City University London, London.
- Eurostat (2018) *Employment rates by sex, age and NUTS 2 regions (%)*, Eurostat.
- Evans-Lacko S, Knapp M (2017) Health care – evaluating the overall system, in Greve B (ed.) *Evaluations of Social Policy*, Elgar, Cheltenham, pp. 279-299.
- Fujiwara D (ND) *Measuring the Social Impact of Community Investment: The Methodology Paper*, HACT.
- Fujiwara D (2013) *A General Method for Valuing Non-Market Goods Using Wellbeing Data: Three-Stage Wellbeing Valuation*, CEP Discussion Paper No. 1233, Centre for Economic Performance, London.
- Fujiwara D, Campbell R (2011) *Valuation Techniques for Social Cost-Benefit Analysis: Stated Preference, Revealed Preference and Subjective Well-Being Approaches*, HM Treasury.
- Fujiwara D, Keohane K, Clayton V, Hotopp U (2017) *Mental Health and Life Satisfaction: The Relationship between the Warwick Edinburgh Mental Wellbeing Scale and Life Satisfaction; A pilot study*.
- Fujiwara D, Kurdna L, Dolan P (2014) *Quantifying and Valuing the Wellbeing Impacts of Culture and Sports*, Department for Culture, Media and Sport, London.
- Fujiwara D, Vine J (2015) *The Wellbeing Value of Tackling Homelessness*, Simetrica and HACT.
- Gao Z, House LA, Xie J (2016) Online survey data quality and its implication for willingness-to-pay: a cross-country comparison, *Canadian Journal of Agricultural Economics*, 64, 199–221.
- Garrido-Garcia S, Sanchez-Martinez FI, Abellan-Perpinan JM, van Excel J (2015) Monetary valuation of informal care based on carers' and noncarers preferences, *Value in Health*, 18, 832-840.
- Ghibelli P, Barbieri D, Fernandez JL, Knapp M (2017) *The Role of Public and Private Actors in Delivering and Resourcing Long-term Care Services*, SPRINT Working Paper D2.3, SPRINT, Brussels.
- Grant JS, Elliott TR, Weaver M, Bartolucci AA, Giger JN (2002) Telephone intervention with family caregivers of stroke survivors after rehabilitation, *Stroke*, 33, 8, 2060–2065.
- Greve B (2017) *The Long-Term Care Resourcing Landscape*, SPRINT Working Paper D2.4, SPRINT, Brussels.
- Greve B, Lopes A, Richards A (2018) *Social Investment Criteria in the Field of Long-term Care*, SPRINT Working Paper D4.1, SPRINT, Brussels.
- Gyrd-Hanse D, Kjær T (2012) Disentangling WTP per QALY data: different analytical approaches, different answers, *Health Economics*, 222–237.
- HACT (2016) *Social Value Bank*, HACT.
- Hicks S, Tinkler L, Allin P (2013) Measuring subjective well-being and its potential role in policy: perspectives from the UK Office for National Statistics, *Social Indicators Research*, 114, 73–86.

- Homes and Communities Agency (2014) *Additionality Guide*, 4th edition, Homes and Communities Agency, London.
- Huter K, Kocot E, Skarbak-Kissimova K, Dubas-Jakóbczyk K, Rothgang H (2016) Economic evaluation of health promotion for older people-methodological problems and challenges, *BMC Health Services Research*, 16, 5, 328, 429-479.
- International Labour Organization (2018a) *Working-age Population by Sex and Age (Thousands)*, ILO.
- International Labour Organization (2018b) *Mean Nominal Monthly Earnings of Employees by Sex and Occupation – Harmonized Series*, ILO.
- Kaambwa B, Gill L, McCaffrey N, Lancsar E, Cameron ID, Gray L, Ratcliffe J (2015) An empirical comparison of the OPQoL-Brief, EQ-5D-3L and ASCOT in a community dwelling population of older people, *Health and Quality of Life Outcomes*, 13, 164.
- Kahneman D, Deaton A (2010) High income improves evaluation of life but not emotional well-being, *Psychological and Cognitive Sciences*, 107, 38.
- Kapteyn A, Lee L, Tassot C, Vonkova H, Zamarro G (2014) Dimensions of subjective well-being, *Social Indicators Research*, 123, 625–660.
- Knapp M (1984) *Economics of Social Care*, MacMillian, London.
- Knapp M, Glanz A, Richards A (2017) *Glossary of Key Terms*, SPRINT Working Paper D2.2, SPRINT, Brussels.
- Kroll C, Delhey J (2013) A happy nation? Opportunities and challenges of using subjective indicators in policymaking, *Social Indicators Research*, 114, 13-28.
- Layard R (2016) *Measuring Wellbeing and Cost-effectiveness Analysis Using Subjective Wellbeing*, Discussion Paper 1, What Works Wellbeing, London.
- Lee S, McClain C, Han S (2016) Question order sensitivity of subjective well-being measures: focus on life satisfaction, self-rated health, and subjective life expectancy in survey instruments, *Qualitative Life Research*, 25, 2497-2510.
- Lopes A (2017) *Conceptual Report on Long-term Care*, SPRINT Working Paper D2.1, SPRINT, Brussels.
- Malley J, Fernandez JL (2010) Measuring quality in social care services: theory and practice, *Annals of Public and Cooperative Economics*, 81, 4, 559-582.
- Mant J, Carter J, Wade DT, Winner S (2000) Family support for stroke: a randomised controlled trial', *Lancet*, 2, 356, 9232, 808-813.
- Martin-Fernandez J, Polentinos-Castro E, del Cura-Gonzalez MI, Ariza-Cardiel G, Abraira V, Gil-LaCruz AI, Garcia-Perez S (2014) Willingness to pay for a quality-adjusted life year: an evaluation of attitudes towards risk and preferences, *BMC Health Services Research*, 14, 287.

- Mukuria C Rowen D, Peasgood T, Brazier J (2014) *An Empirical Comparison of Well-being Measures Used in the UK: Research Interim Report*, Policy Research Unit in Economic Evaluation of Health and Care Interventions, Sheffield and York.
- Murtin F, Boarini R, Cordoba J, Ripoll M (2015) *Beyond GDP: Is there a law of one shadow price?*, *OECD Statistics Working Papers*, 2015/05, OECD, Paris.
- Natural Capital Coalition (2016) *Natural Capital Protocol*, Natural Capital Coalition, London.
- Netten A, Burge P, Malley J, Potoglou D, Towers AM, Brazier J, Flynn T, Forder J, Wall B (2012) *Outcomes of social care for adults: developing a preference-weighted measure*, *Health Technology Assessment*, 16, 16.
- Netten A, Ryan M, Smith P, Skatun D, Healey A, Knapp M, Wykes T (2002) *The Development of a Measure of Social Care Outcome for Older People*, PSSRU Discussion Paper 1690/2, Personal Social Services Research Unit, Kent.
- Niedzwiedz CL, Katikireddi SV, Pell JP, Mitchell R (2014) *The association between life course socioeconomic position and life satisfaction indifferent welfare states: European comparative study of individuals in early old age*, *Age and Ageing*, 43, 431-436.
- Nussbaum M, Sen A (1993) *Capability and well-being*, in Nussbaum M, Sen A (eds.) *The Quality of Life*, Clarendon Press, Oxford, pp. 30-53.
- OECD (nd) *Average Annual Hours Actually Worked Per Worker*, OECD, Paris.
- OECD (2018) *Purchasing Power Parities (PPP)*, OECD, Paris.
- OECD (2015) *How's Life? 2015 Measuring Well-Being*, OECD, Paris.
- Oliver A (2013) *Testing procedural invariance in the context of health*, *Health Economics*, 22, 272–288.
- Page K, Barnett AG, Graves N (2017) *What is a hospital bed day worth? A contingent valuation study of hospital Chief Executive Officers*, *BMC Health Services Research*, 17, 137.
- Poškutě V (2018) *Determination of Success Factors for Investing in Long-term Care*, SPRINT Working Paper D4.3, SPRINT, Brussels.
- Richards A (2018) *Social Return on Investment Framework Application Guide to Long-term Care*, SPRINT Working Paper D4.2, SPRINT, Brussels.
- Richards A, Linnosmaa I, Jokimäki H, Haula T (2018) *Feasibility Framework Tool for Social Investment*, SPRINT Working Paper D5.1, SPRINT, Brussels.
- Richards A, Nicholls J (2017) *A Discussion Document on the Valuation of Social Outcomes*, Social Value International, Liverpool.
- Riedel M, Kraus M (2011) *Informal care provision in Europe: Regulation and Profile of Providers*, ENEPRI Research Report No. 96, European Network of Economic Policy Research Institutes.

- Robinson A, Gyrd-Hansen D, Bacon P, Baker R, Pennington R, Donaldson C (2013) Estimating a WTP-based value of a QALY: The ‘chained’ approach’, *Social Science and Medicine*, 92, 92–104.
- Ryen L, Svensson M (2015) The willingness to pay for a Quality Adjusted Life Year: a review of the empirical literature, *Health Economics*, 24, 1289–1301.
- Self A (2015) Quality of life measurement and application to policy: experiences from the UK Office for National Statistics, *Social Indicators Research*, 130, 147-160.
- Social Value International (2015) *The Seven Principles of Social Value*, Social Value International, Liverpool.
- Stiglitz JE, Sen A, Fitoussi JP (2009) *Report by the Commission on the Measurement of Economic Performance and Social Progress*.
- The Impact Management Project (2018) *The Impact Management Project*.
- Trotter L, Rallings Adams MK (2017) *Valuing improvements in Mental Health; Applying the Wellbeing Valuation Method to WEMWBS*, HACT, London.
- van den Berg B, Gafni A, Portrait F (2013) *Attributing a Monetary Value to Patients’ Time: A Contingent Valuation Approach*, Centre for Health Economics, University of York.
- van den Berg B, Fiebig DG, Hall J (2014) Well-being losses due to care-giving, *Journal of Health Economics*, 35, 123–131.
- van Leeuwen KM, Bosmans JE, Jansen APD, Hoogendijk EO, van Tulder M, van der Horst HE, Ostelo RO (2015a) Measurement properties of the EQ-5D-3L, ICECAP-O, and ASCOT in frail older adults, *Value in Health*, 18, 35-43.
- van Leeuwen KM, Jansen APD, Muntingam ME, Westerman MJ, van Tulder MW, van der Horst H.E. (2015b) Exploration of the content validity and feasibility of the EQ-5D-3L, ICECAP-O and ASCOT in older adults, *Health Services Research*, 15.
- Voicu B (2015) Priming effects in measuring life satisfaction, *Social Indicators Research*, 124, 993–101.
- Wolff JL, Kasper JD (2004) Informal caregiver characteristics and subsequent hospitalization outcomes among recipients of care, *Aging Clinical and Experimental Research*, 16,4, 307-313.
- World Health Organization (2011) WHO-CHOICE unit cost estimates for service delivery – Estimation file.

8 Annex 1: Overview of the Results of Results for the Theoretical Context

		Base search term	Refinement #1: Article	Refinement #2: English	Refinement #3: Year of publication 2013 to 2017	Refinement #4: Categories	Screening by title and abstract
1	You searched for: TOPIC: ("SROI") OR TOPIC: ("Social Return on Investment") OR TOPIC: ("Social outcome measurement") OR TOPIC: ("Non-financial measures") OR TOPIC: ("Non-financial measurement") OR TOPIC: ("Management of social outcomes") OR TOPIC: ("Monetization of outcomes") OR TOPIC: ("Monetisation of outcomes") OR TOPIC: ("Monetization of social outcomes") OR TOPIC: ("Monetisation of social outcomes") OR TOPIC: ("Valuation of social outcomes") OR TOPIC: ("SI in LTC") OR TOPIC: ("Social investment in LTC") OR TOPIC: ("SI in long term care") OR TOPIC: ("Social Investment in long term care") OR TOPIC: ("CBA in healthcare") OR TOPIC: ("Cost benefit analysis in healthcare")	181	114	111	56	39	7
2	TOPIC: ("Wellbeing measurement") OR TOPIC: ("Well-being measurement") OR TOPIC: ("Wellbeing measures") OR TOPIC: ("Well-being measures") OR TOPIC: ("Measurement of wellbeing") OR TOPIC: ("Measurement of well-being") OR TOPIC: ("Measuring wellbeing") OR TOPIC: ("Measuring well-being") OR TOPIC: ("Valuation of wellbeing") OR TOPIC: ("Valuation of well-being") OR TOPIC: ("Wellbeing of older people") OR TOPIC: ("Well-being of older people") OR TOPIC: ("Wellbeing in LTC") OR TOPIC: ("Well-being in LTC") OR TOPIC: ("Wellbeing in long term care") OR TOPIC: ("Well-being in long term care") OR TOPIC: ("Quantifying wellbeing") OR TOPIC: ("Quantifying well-being") OR TOPIC: ("Managing wellbeing") OR TOPIC: ("Managing well-being") OR TOPIC: ("Wellbeing from social investment") OR TOPIC: ("Well-being from social investment") OR TOPIC: ("Wellbeing of older people") OR TOPIC: ("Well-being of older people")	774	660	634	267	143	32

3	TOPIC: ("Quality of life measurement") OR TOPIC: ("QoL measurement") OR TOPIC: ("Quality of life measures") OR TOPIC: ("QoL measures") OR TOPIC: ("Measurement of Quality of life") OR TOPIC: ("Measurement of QoL") OR TOPIC: ("Measuring Quality of life") OR TOPIC: ("Measuring QoL") OR TOPIC: ("Valuation of Quality of life") OR TOPIC: ("Valuation of QoL") OR TOPIC: ("Indicators of Quality of life") OR TOPIC: ("Indicators of QoL") OR TOPIC: ("Quality of life in LTC") OR TOPIC: ("QoL in LTC") OR TOPIC: ("Quality of life in long term care") OR TOPIC: ("QoL in long term care") OR TOPIC: ("Quantifying Quality of life") OR TOPIC: ("Quantifying QoL") OR TOPIC: ("Managing Quality of life") OR TOPIC: ("Managing QoL") OR TOPIC: ("Quality of life from social investment") OR TOPIC: ("QoL from social investment") OR TOPIC: ("Quality of life of older people") OR TOPIC: ("QoL of older people")	5,332	4,076	3,917	1,133	310	40
4	TOPIC: ("Life satisfaction measurement") OR TOPIC: ("Life-satisfaction measurement") OR TOPIC: ("Life satisfaction measures") OR TOPIC: ("Life-satisfaction measures") OR TOPIC: ("Measurement of Life satisfaction") OR TOPIC: ("Measurement of Life-satisfaction") OR TOPIC: ("Measuring Life satisfaction") OR TOPIC: ("Measuring Life-satisfaction") OR TOPIC: ("Valuation of Life satisfaction") OR TOPIC: ("Valuation of Life-satisfaction") OR TOPIC: ("Indicators of Life satisfaction") OR TOPIC: ("Indicators of Life-satisfaction") OR TOPIC: ("Life satisfaction in LTC") OR TOPIC: ("Life-satisfaction in LTC") OR TOPIC: ("Life satisfaction in long term care") OR TOPIC: ("Life-satisfaction in long term care") OR TOPIC: ("Quantifying Life satisfaction") OR TOPIC: ("Quantifying Life-satisfaction") OR TOPIC: ("Managing Life satisfaction") OR TOPIC: ("Managing Life-satisfaction") OR TOPIC: ("Life satisfaction from social investment") OR TOPIC: ("Life-satisfaction from social investment") OR TOPIC: ("Life satisfaction of older people") OR TOPIC: ("Life-satisfaction of older people")	86	75	72	33	16	6
5	TOPIC: ("Happiness measurement") OR TOPIC: ("Happiness measures") OR TOPIC: ("Measurement of Happiness") OR TOPIC: ("Measuring Happiness") OR TOPIC: ("Valuation of Happiness") OR TOPIC: ("Indicators of Happiness") OR TOPIC: ("Happiness in LTC") OR TOPIC: ("Happiness in long term care") OR TOPIC: ("Quantifying Happiness") OR TOPIC: ("Managing Happiness") OR TOPIC: ("Happiness from social investment") OR TOPIC: ("Happiness of older people")	88	67	66	28	16	5

6	TOPIC: ("Valuing outcomes") OR TOPIC: ("Value of outcomes") OR TOPIC: ("Valuation of outcomes") OR TOPIC: ("Monetising outcomes") OR TOPIC: ("Monetizing outcomes") OR TOPIC: ("Weighting of outcomes") OR TOPIC: ("Preferences of outcomes") OR TOPIC: ("Valuing non-financial outcomes") OR TOPIC: ("Value of non-financial outcomes") OR TOPIC: ("Valuation of non-financial outcomes") OR TOPIC: ("Monetising non-financial outcomes") OR TOPIC: ("Monetisation of non-financial outcomes") OR TOPIC: ("Monetizing non-financial outcomes") OR TOPIC: ("Monetization of non-financial outcomes") OR TOPIC: ("Weighting non-financial outcomes") OR TOPIC: ("Preferences of non-financial outcomes") OR TOPIC: ("Valuing impacts") OR TOPIC: ("Value of impacts") OR TOPIC: ("Valuation of impacts") OR TOPIC: ("Monetising impacts") OR TOPIC: ("Monetisation of impacts") OR TOPIC: ("Monetizing impacts") OR TOPIC: ("Monetization of impacts")	94	80	79	27	7	1
7	TOPIC: ("Contingent valuation techniques") OR TOPIC: ("Stated preference techniques") OR TOPIC: ("Revealed preference techniques") OR TOPIC: ("Stated preference valuations") OR TOPIC: ("Revealed preference valuations") OR TOPIC: ("Wellbeing valuations") OR TOPIC: ("Well-being valuations") OR TOPIC: ("Subjective well-being valuations") OR TOPIC: ("Subjective wellbeing valuations") OR TOPIC: ("Utility weighting") OR TOPIC: ("Preference weighting") OR TOPIC: ("Stakeholder valuations") OR TOPIC: ("Purpose of valuation") OR TOPIC: ("Purpose of valuing") OR TOPIC: ("Purpose of monetisation") OR TOPIC: ("Wellbeing values") OR TOPIC: ("Well-being values") OR TOPIC: ("Subjective well-being values") OR TOPIC: ("Subjective wellbeing values") OR TOPIC: ("Utility-weighting") OR TOPIC: ("Preference-weighting") OR TOPIC: ("Stakeholder values") OR TOPIC: ("Utility preferences") OR TOPIC: ("Utility-preferences") OR TOPIC: ("Approach to valuation")	494	410	405	131	46	4
8	TOPIC: ("How to value outcomes") OR TOPIC: ("Contingent choice experiments") OR TOPIC: ("Choice modelling experiments") OR TOPIC: ("Hedonic-pricing method") OR TOPIC: ("Contingent valuation") OR TOPIC: ("How to weight outcomes") OR TOPIC: ("Contingent choice-experiments") OR TOPIC: ("Choice-modelling experiments") OR TOPIC: ("Hedonic pricing experiments") OR TOPIC: ("How to weight preferences") OR TOPIC: ("Contingent-valuation")	5,339	4,676	4,597	1,385	569	28

9	TOPIC: ("Why value outcomes") OR TOPIC: ("Why do we value outcomes") OR TOPIC: ("Purpose of valuation") OR TOPIC: ("Purpose of valuing") OR TOPIC: ("Purpose of monetisation") OR TOPIC: ("Motivations for valuation") OR TOPIC: ("Motivations for valuing outcomes") OR TOPIC: ("Purpose of valuing outcomes") OR TOPIC: ("Benefits of valuation") OR TOPIC: ("Advantages of valuation") OR TOPIC: ("Disadvantages of valuation") OR TOPIC: ("Opportunities of valuation") OR TOPIC: ("Strengths of valuation") OR TOPIC: ("Weaknesses of valuation") OR TOPIC: ("Strengths and weaknesses of valuation")	13	11	7	2	1	0
---	--	----	----	---	---	---	---

9 Annex 2: Categories Employed in Web of Science Searches for Theoretical Context

Web of Science Category terms

Behavioral Sciences

Economics

Family Studies

Geriatrics & Gerontology

Gerontology

Health Care Sciences & Services

Health Policy & Services

Management

Nursing

Primary Health Care

Rehabilitation

Social Issues

Social Sciences, Biomedical

Social Sciences, Interdisciplinary

Social Work

Sociology

Statistics & Probability

10 Annex 3: Overview of the Health Results for the Empirical Review

Authors, publishing year and country	Title	Objective	Data and methods	Findings
Barnay T & Juin S, 2016, France	Does home care for dependent elderly people improve their mental health?	To estimate the effects of both formal home care (provided by professional workers) and informal care (provided by the family and other relatives) on the mental health of dependent elderly individuals living at home in France.	<p><i>Data:</i> Ordinary Households section of the French Disability and Health Survey (2008) for non-institutionalized people, sample of dependent elderly persons aged 65 and over</p> <p><i>Main dependent variables:</i> depression (0/1), and Mental-Health Inventory (MHI-5) built from five questions of the Short Form Health Survey (SF-36) each scoring from 1 to 5 and total score transformed to a 0-100 scale (100 being the best possible health)</p> <p><i>Methods:</i> OLS, Instrumental variables models using 2SLS</p> <p><i>Instrumental variables:</i></p> <p><i>For informal care:</i> the proportion of daughters, having at least one child who has no child, having at least one child who has no partner, and having</p>	<p>The null hypothesis of exogeneity of formal and informal care was rejected in both models ($p=0.049$ for depression and $p=0.004$ for MHI-5), indicating that IV models were preferred to OLS.</p> <p><i>Receiving informal care</i> reduced the risk of depression by 42% ($p<0.1$) and a one-unit increase in <i>formal care</i> hours improved the MHI-5 by 1.8 points on the 0-100 scale ($p<0.05$).</p>

			<p>at least one child who lives nearby</p> <p><i>For formal care:</i> the proportion of individuals aged 75 and over who received the Personal Autonomy Allowance (PAA) at the departmental level in 2008</p> <p>Selected those with the highest first stage F-statistics: the proportion of daughters and the proportion of PAA recipients for the depression model; having at least one child who has no partner and the proportion of PAA recipients for the MHI-5 model</p>	
<p>Coe NB & Van Houtven CH, 2009, USA</p>	<p>Caring for mom and neglecting yourself? The health effects of caring for an elderly parent</p>	<p>To study whether adult (non-co-residing) childrens` continued and initial informal caregiving to a mother has effects on their mental and physical health</p>	<p><i>Data:</i> Health and Retirement Survey (HRS, 7 waves from 1992-2004) for a nationally representative sample of the near elderly in the US (ages 50 to 64 entered the sample initially)</p> <p><i>Main dependent variable:</i> CES-D8 index, 8-point scale (scores of 4 or 5 and above consistent with probable clinical depression)</p> <p><i>Methods:</i> Instrumental variable techniques</p>	<p>The Hausman test for endogeneity revealed that health and caregiving were endogenous for continued care, but endogeneity did not exist for selection into caregiving.</p> <p><i>Continued caregiving</i> increased the reported number of depressive symptoms measured by CES-D8 by 0.6 on an 8-point scale (a 47% increase from the mean score) compared to caregivers who stopped caregiving due to the death of the mother. The increase was 0.6 for married women and</p>

			<p><i>Instrumental variables:</i></p> <p><i>Continued caregiving:</i> death of the mother</p> <p><i>Initial caregiving:</i> number of children in a family, percent of children who are girls, eldest child is a daughter, total number of kids among all siblings, total number of grand kids, total number of siblings working in the wave prior to caregiving</p>	<p>0.65 for married men (both $p < 0.05$).</p> <p>Continued caregiving decreased the likelihood of reporting excellent or very good health by 10 % for married women ($p < 0.1$) but increased it almost 17 % for married men ($p < 0.05$), compared to those who stopped caregiving due to the death of the mother.</p> <p>No significant health effect was found for caregiving single women, but single caregiving men reported 40% more likely to have a heart condition than those who stopped providing care to their mothers ($p < 0.05$).</p> <p><i>Initial caregiving</i> increased the reported number of depressive symptoms by 0.17 on an 8-point scale for married women (a 15% increase from the mean score, $p < 0.05$). There were no other significant health effects among married women. No significant health effect was found for married men.</p>
<p>Do YK, Norton EC, Stearns SC & Van Houtven CH, 2015, South Korea</p>	<p>Informal care and caregiver's health</p>	<p>To study the effect of informal caregiving on health and health care use of women who are</p>	<p><i>Data:</i> Korean Longitudinal Study of Aging (KLoSA, 3 waves from 2006-2010) for non-</p>	<p>Statistical evidence of endogeneity was found in 5 of the 12 estimated models (4/6 models for the</p>

		caregivers	<p>institutionalized South-Korean adults aged 45 or older, sample of women with any living parents-in-law and women with any living parents</p> <p><i>Main dependent variables:</i> Dummy variable for pain affecting daily activities, Dummy for reporting fair-to-poor self-rated overall health, Dummies for any outpatient care use and for out-of-pocket spending for outpatient care, Dummies for any regular prescription drug use and for out-of-pocket spending for regular prescription drug use</p> <p><i>Methods:</i> Instrumental variable techniques</p> <p><i>Instrumental variables:</i> any (Activity of Daily Living (ADL) limitations of the mother-in-law and of the father-in-law for the sample of daughters-in-law; any ADL limitations of the mother and father for daughters</p>	<p>daughter sample and 1/6 model for the daughter-in-law sample)</p> <p>Informal caregiving increased</p> <p>-the probability of pains affecting daily activities by 13.1 %-points (p<0.01) for the daughter-in-law sample and by 41.3 %-points (p<0.01) for the daughter sample</p> <p>- the probability of reporting fair to poor self-rated health by 15.7 %-points (p<0.05) for the daughter-in-law sample and by 56.1 %-points (p<0.01) for the daughter sample</p> <p>- the probability of any outpatient use by 29.8 %-points (p<0.05) and the probability of any regular prescription drug use by 34 %-points (p<0.01) for the daughter sample</p>
Schmitz H & Westphal M, 2015, Germany	Short- and medium-term effects of informal care provision on female caregivers' health	To estimate the effect of informal care provision on female caregivers' health	<i>Data:</i> German Socio-Economic Panel (SOEP, years 2002-2010), a yearly repeated representative longitudinal	In the first year MCS decreased by 2.00 units (or 20% of a standard deviation, SD, p<0.01) for a woman caring at least 2h per day. Effect

			<p>survey of households and persons above the age of 18 living in Germany, sample restricted to non-professional female caregivers</p> <p><i>Main dependent variables:</i> a mental and a physical health score (MCS and PCS), based on information from the SF-12v2 questionnaire (a component of the SOEP), scales range 0-100</p> <p><i>Methods:</i> Propensity score methods combined with regression methods (regression adjusted matching approach)</p>	<p>reduced to 16% of a SD ($p < 0.01$) three years after treatment assignment and settled to below 12% five and seven years afterwards. There were no differences in the effect between 1 and 2h of care per day. A considerably stronger short-term effect was found for 3h of care per day, with a reduction of MCS by 31% of a SD ($p < 0.01$). Effect did not remain on that high level but dropped back to regions similar to those for 1 and 2h.</p> <p>In contrast, for PCS there was basically a zero effect throughout all periods.</p>
--	--	--	---	--

11 Annex 4: Overview of Different Health and Wellbeing Measures

General effect being measured	Measure	Description	Further information
Health	EQ-5D	An instrument that measures health-related quality of life across 5 dimensions including mobility, self-care, and anxiety/depression. Scores can be presented as a single index number and are valued by reference to national experiments that ask respondents to identify preference weights of different health states. This is then combined with techniques such as time-trade-off or standard gamble to calculate QALYs/DALYs.	https://euroqol.org/eq-5d-instruments/
	DALY (Disability adjusted life year)	A measure of the years of live lost to disease, illness or premature death. See QALY for more detail.	https://www.nice.org.uk/Glossary?letter=D
	QALY (quality-adjusted life year)	A measure of health that expresses results as the number of years of life adjusted to reflect the quality of life. Scores are anchored to preference weighted states of 0 – 1. The former represents a year of life lived in a state equivalent to death, whereas a score of 1 is representative of a year lived in perfect health.	https://www.nice.org.uk/glossary?letter=q
Wellbeing	ASCOT	A measure of social care related quality of life developed in the UK. 8 domains are used for care recipients and 7 for informal carers to measure their social care related quality of life (SCRQoL).	https://www.pssru.ac.uk/ascot/
	GHQ-12 (General Health Questionnaire)	The short version of the measure to understand a person’s current health-status and how that compares to their usual state.	https://www.gl-assessment.co.uk/products/general-health-questionnaire-ghq/

	ICECAP	A measure of wellbeing for health and social care application developed in the UK relating to Sen's capabilities approach.	https://www.birmingham.ac.uk/research/activity/mds/projects/HaPS/HE/ICECAP/index.aspx
	Life-satisfaction	Measures life-satisfaction on a 0 – 10 point scale with respondents asked to rate their current state. Variations such as Cantrill's Ladder employ the same approach and underpin the Gallup World Poll.	https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/bulletins/measuringnationalawellbeing/2015-09-23 http://www.gallup.com/analytics/213704/world-poll.aspx
	ONS-4	The 4 questions employed by the UK Office for National Statistics to measure the wellbeing of citizens. Each question has a scale of 0-10 from 'not at all' to 'completely'.	https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/methodologies/surveysusingthe4officefornationalstatisticpersonalwellbeingquestions
	SF-36	Measures health across 8 concepts including bodily pain and emotional wellbeing with 36 questions self-reported upon by patients.	https://www.rand.org/health/surveys_tools/mos/36-item-short-form/scoring.html
	SWEMWBS (Shortened Warwick-Edinburgh Mental Wellbeing Scale)	A shortened 7-item scale to measure mental wellbeing for general populations, programmes and policies. As with the original 14-item scale, responses are scored 1-5, providing a possible range of 14-35.	https://warwick.ac.uk/fac/med/research/platform/wemwbs/development/swemwbs/

12 Annex 5: Overview of Approaches to Monetise Outcomes

Category	Approach	Outline of the approach	Data required	Advantages	Disadvantages
Cost-based approaches	Replacement costs	The cost of replacing an informal service with formal services. Or, the cost of replacing one formal service with another.	The price of providing private or public health / social care services (e.g. replacing therapeutic services with GP visits).	+ Provides surrogate measures of value for regulatory services (which are difficult to value by other means). + A readily transparent and defensible method when based on market data.	- Can overestimate values. - Does not consider social preferences for services or behaviour in the absence of the services. - The replacement service probably only represents a proportion of the full range of services provided by the alternative service.
Revealed preference approaches	Market prices	The price of a good or service that best reflects what we want to value.	Market price of goods or services.	+ A readily transparent and defensible method since based on market data. + It reflects an individual's willingness to pay (WTP).	- Only applicable where a market exists for the goods or services and this data is readily available. - Risk of undervaluation as people will often value things more highly than the price paid.

Effect on production	Changes in the output of a marketed good or service to a measurable change in goods.	Data on changes in the output of a. Data on cause and effect relationship (e.g. how informal care affects employment)	+ If data is available, it is a relatively straightforward technique to apply.	- Necessary to recognize and understand the relationship between marketed goods or services and the output of the product.
Travel costs	The amount of time and money people spend visiting a particular service or event.	The amount of time and money that people spend visiting a particular service or event.. (e.g. the number of hours and cost of fuel to attend)	+ Based on actual behaviour (what people do) rather than a hypothetically stated WTP. + The results are relatively easy to interpret and explain.	- Approach is limited to direct use of recreational benefits. - Difficulties in apportioning costs when trips are to multiple places or are for more than one purpose. - Considering travel costs alone ignores the opportunity cost of time while travelling. - Risk of undervaluation.
Hedonic pricing	The difference in property prices or wage rates that can be ascribed to the different qualities of a property or position.	Usually data relating to differences in property prices or wage rates that can be ascribed to the different qualities (e.g. a landscape view of a property, access to better school results).	+ Readily transparent and defensible method since based on market data and WTP. + Property markets are generally very responsive so are good indicators of values.	- Approach is largely limited to benefits related to property. - The property market is affected by a number of factors, so the effect needs to be isolated or it may be overvalued.

Wellbeing approaches	Wellbeing Valuation (WV)	Wellbeing valuations assess the relationship between life circumstances (e.g. health status, volunteering, attending community groups) and levels of self-reported wellbeing, and what level of income change would provide the same level of change in wellbeing.	Large statistical datasets (e.g. the British Household Panel Survey).	+ Necessary datasets publicly available. + Additional datasets can be created.	- Data needed may not be available for either the outcome or for a specific stakeholder group in which case costs will be higher.
		Respondents asked directly for their	Large statistical datasets (e.g. the	+ Avoids the need for willingness to pay (WTP) scenarios which rely on	

Stated preference approaches	Hybrid stated preference / Wellbeing valuation	willingness to accept (WTA) compensation for a loss such that their level of wellbeing does not change.	British Household Panel Survey). Stated value that people place on the wellbeing associated with a good or service (e.g. access to a library service); demographic and biographical information on survey respondents obtained through survey questionnaires.	hypothetic entrance fees. + Produces values per visit similar to willingness to pay (WTP) valuations.	- Data needed for Wellbeing Valuation may not be available in which case costs will be higher.
		Infer values of outcomes by asking people directly what is their willingness to pay (WTP) for them, or their willingness to accept (WTA)	Stated value that people place on a good or service (e.g. existence of a species, increased confidence); demographic and biographical	+ Captures both use and non- use values. + Extremely flexible - it can be used to estimate the economic value of virtually anything. + Gives a much more accurate	- The results are subject to numerous different biases from respondents. - e.g. respondents may express a positive WTP to promote a “warm glow” effect, overestimating value. - e.g. if the cost is perceived as a tax,

	<p>Choice experiments (CE), or Multi-Choice Experiments (MCE) Including the Valuation game (as a form of CE)</p>	<p>Presents a series of alternative resource or use options, each defined by various attributes set at different levels (including price), and asks respondents to select which option (i.e. sets of attributes at different levels) they prefer (e.g. numbers of species present and percentage coral cover).</p> <p>Participants asked to place value on outcomes by comparing preferences, or by comparing goods or services which have known market values.</p>	<p>As for CV above, although CE contrasts several different scenarios. An appropriate set of “levels” are required for the different parameters (e.g. ranging from 0% health status to 100%).</p> <p>Relative values that people place on goods or services or preferences to outcomes.</p> <p>Demographic and biographical information..</p>	<ul style="list-style-type: none"> + Captures both use and non- use values. + Provides theoretically more accurate values for marginal changes (e.g. values per % increase in coral cover). + Gives a much more accurate outcome than benefit transfers. + Extremely flexible and useful for defining outcomes, recognizing subgroups of stakeholders. Order of magnitude valuation for service design. + Captures both use and non- use values. 	<p>The results are subject to numerous different biases from respondents.</p> <ul style="list-style-type: none"> - Can be mentally challenging for respondents to truly weigh up the alternative choices given to them in the time available. - The results are subject to numerous different biases from respondents. - Preferences need to align with market costs where more than one outcome is being valued for service design purposes.
--	--	---	---	---	--

	Auction game (as a form of CE)	Involves participants bidding to determine their maximum WTP for an outcome, good or service.	As for CV above	<ul style="list-style-type: none"> + Captures both use and non- use values. + Extremely flexible - it can be used to estimate the economic value of virtually anything. + Gives a much more accurate outcome than benefit transfers. 	- The results are subject to numerous different biases from respondents.
Benefit (Value) transfer	Benefit transfer	Involves transferring value estimates from existing economic valuation studies to the study site in question, making adjustments where appropriate.	<p>Valuations from similar studies elsewhere.</p> <p>Data on key variables from different studies (e.g. GDP per person).</p>	<ul style="list-style-type: none"> + Relatively low-cost when there is a similarity between that which is being valued. 	<ul style="list-style-type: none"> - The results may not be relevant to the stakeholder group for which the value is being calculated. - Existing valuation studies may be more robust and numerous for some services than for others.

13 Annex 6: Further Information on the Wellbeing Valuation Methodology

Fujiwara (ND) outlines the formula as;

$$MRS = \frac{\beta Q}{\beta M}$$

Where Q is the effect of outcomes, and M is the effect of a change of income.

The methodology is reported to be the most robust approach to wellbeing valuations (*ibid.*), employing a three-stage Wellbeing Valuation technique, and is not subject to the same risks of bias as a single equation model. Therefore, in accordance with recommendations from the UK Treasury's Green Book, an instrumental variable approach, based on exogenous changes to an individual's income as the result of a lottery win is used owing to the random nature of such wins. The results illustrated a causal effect of a log-point increase in income as increasing life-satisfaction by 1.013 index points per year.

Additionally, data from multiple large-scale surveys in the UK are used to conduct multivariate analysis to calculate effects of different outcomes on an individual's life-satisfaction. When first calculated, the median UK household income of £30,000 was employed as the base value (this figure has been updated), and average values for compensating surplus (essentially comparable to WTP for a positive outcome, and WTA a negative outcome) are calculated for individuals across different age and geographic locations (more information is available at <http://www.hact.org.uk/value-calculator>).

The formulas employed to calculate the income and outcome models are:

Income model; $LS_i = f(\ln(M_i))$

Outcome model; $LS_i = g(Q_i)$

where LS = life satisfaction, Q = the outcome in question and M = income.

14 Annex 7: Diminishing Marginal Utility of SWEMWBS Scores

Category	Overall SWEMWBS score	Full model value	Proportion of change in value compared to total	Category	Overall SWEMWBS score	Full model value	Proportion of change in value compared to total
1	7-14	£0	N/A	7	25-26	£24,225	4.8%
2	15-16	£9,639	36%	8	27-28	£24,877	2.4%
3	17-18	£12,255	9.8%	9	29-30	£25,480	2.3%
4	19-20	£17,561	19.8%	10	31-32	£25,856	1.4%
5	21-22	£21,049	13.0%	11	33-34	£26,175	1.2%
6	23-24	£22,944	7.1%	12	35	£26,793	2.3%

Sourced from Fujiwara *et al.* (2017)

15 Annex 8: Labour Market Effects Variables and Evidence

Impact question	Data utilised
3. Who experiences changes?	<p>ILO (2018) population sizes of different nations aged 50-64 (segmented by sex).</p> <p>Riedel and Kraus (2011) proportion of population aged 50+ that provide informal care (segmented by sex).</p> <p>Eurostat (2018) employment rates for 55 – 64-year olds (segmented by sex).</p> <p>OECD (ND) average number of hours worked in declared dependent employment.</p>
4: What changes are experienced?	<p>Data taken from Bolin <i>et al.</i> (2008):</p> <ol style="list-style-type: none"> 1. A reduced probability (average of 3.7%) of being in the labour-market with an increase of 10% in informal care provision (segmented to 3.2% for males and 2.8% for females). 2. Reduced hours worked (average of 2.6%) with an increase of 10% in informal care provision.
5: How can we measure the changes?	Data supplied based on assessment against control groups.
6: How much of each change has happened?	Modelled estimations provided based upon 0.1% and 1% of the relevant populations.
7: How much of each change is caused by our activities?	<p>Evidence of counterfactual provided by the study – using control groups.</p> <p>No evidence of attribution included.</p>
8: How long do we need to measure the changes for?	Estimated duration of one year included.
9: What is the relative importance of the different changes?	<p>ILO (2018) average monthly salaries – segmented also by sex.</p> <p>ILO (2018) lowest monthly salaries by occupation type – segmented also by sex.</p> <p>ILO (2018) monthly salaries based on national minimum wage.</p>

16 Annex 9: Monetized Results of Labour Market Effects of Increased Provision of Informal Care

Data to Inform Modelling of Labour Market Effects

Country	Quantity of informal carers			Employment rates			Income			
	Total	Males	Females	Employment rate (54-65-year olds)	Males (54 - 65-year olds)	Females (54 - 65-year olds)	Gross average income (2011)	Males average income	Females average income	Minimum wage
Austria	87,940	41,600	45,728	49.20%	57.60%	41.10%	\$ 38,496	\$ 47,172	\$ 29,040	N/A
Elementary occupations				\$ 28,548	\$ 31,764	\$ 25,368				
Denmark	51,817	26,862	25,397	67.80%	71.90%	63.60%	\$ 51,192	\$ 56,844	\$ 46,176	N/A
Services and sales work				\$ 36,624	\$ 37,884	\$ 36,048				
France	483,640	210,937	267,387	49.80%	51.60%	48.10%	\$ 41,364	\$ 45,240	\$ 37,080	\$ 21,900

Elementary occupations							\$ 26,352	\$ 27,984	\$ 24,852	
Germany	772,949	409,413	363,095	68.60%	73.70%	63.50%	\$ 47,628	\$ 53,244	\$ 41,304	\$ 22,212
Elementary occupations							27432	\$ 28,944	\$ 26,496	
Greece	123,232	40,879	81,465	36.30%	46.20%	27.20%	\$ 30,324	\$ 32,700	\$ 27,684	£ 13,428
Elementary occupations							\$ 20,100	\$ 21,432	\$ 18,468	
Italy	723,969	322,453	391,559	50.30%	61.70%	39.70%	\$ 40,128	\$ 43,488	\$ 35,868	N/A
Skilled agricultural, forestry and fishery workers							\$ 26,496	\$ 26,232	\$ 28,536	
Netherlands	110,436	58,949	51,512	63.50%	72.80%	54.20%	\$ 44,700	\$ 49,776	\$ 39,396	£ 22,932
Elementary occupations							\$ 24,132	\$ 24,492	\$ 23,736	

Spain				49.10%	55.70%	42.80%	\$ 34,044	\$ 38,160	\$ 29,556	£ 13,884
Elementary occupations	457,917	179,408	275,153				\$ 18,672	\$ 23,388	\$ 15,840	

Reducing Hours Worked

Presented Results are based on the national average salaries and lowest average salary by occupation-type

	Model a - 0.1% of informal carers					Model b - 1% of informal carers				
Country and occupation-type with lowest average salary	Gross average income (2011)	Males average income	Females average income	Aggregated males and females	Minimum wage (where appropriate)		Males average income	Females average income	Aggregated males and females	Minimum wage (where appropriate)
Austria	\$43,305	\$29,388	\$14,190	\$43,579	N/A		\$293,884	\$141,902	\$435,787	
Elementary occupations	\$32,115	\$19,789	\$12,396	\$19,789			\$197,892	\$123,959	\$197,892	
Denmark	\$46,760	\$ 28,545	\$19,392	\$47,937	N/A		\$285,447	\$193,921	\$479,368	

Services and sales work	\$33,453	\$19,024	\$15,139	\$34,163			\$190,238	\$151,388	\$341,625	
France	\$259,028	\$ 128,026	\$ 123,993	\$252,020	\$137,141		\$1,280,263	\$1,239,934	\$2,520,197	\$1,371,415
Elementary occupations	\$165,021	\$ 79,193	\$ 83,104	\$162,297			\$791,929	\$831,037	\$1,622,966	
Germany	\$656,614	\$ 417,708	\$247,605	\$665,313	\$306,222		\$4,177,081	\$2,476,047	\$6,653,128	\$3,062,215
Elementary occupations	\$378,186	\$ 227,071	\$158,835	\$385,906			\$2,270,705	\$1,588,353	\$3,859,059	
Greece	\$35,269	\$16,057	\$15,949	\$32,006	\$15,618		\$160,570	\$159,493	\$320,063	\$156,176
Elementary occupations	\$23,378	\$10,524	\$10,640	\$21,164			\$105,240	\$106,398	\$211,638	
Italy	\$379,935	\$ 224,955	\$ 144,967	\$ 369,921	N/A		\$2,249,546	\$ 1,449,667	\$3,699,213	
Skilled agricultural, forestry and fishery workers	\$250,866	\$ 135,693	\$ 115,333	\$ 251,026			\$1,356,928	\$1,153,332	\$2,510,259	

Netherlands	\$ 81,502	\$ 55,540	\$ 28,598	\$84,138	\$41,812		\$555,397	\$285,981	\$ 841,377	\$418,121
Elementary occupations	\$ 44,000	\$27,328	\$17,230	\$44,558			\$273,280	\$ 172,303	\$ 445,582	
Spain	\$199,013	\$ 99,147	\$ 90,498	\$189,644	\$81,163		\$991,467	\$904,976	\$1,896,443	\$811,626
Elementary occupations	\$109,152	\$ 60,766	\$ 48,501	\$109,267			\$1,091,521	\$607,663	\$485,006	\$1,092,669

Modelled Estimations for Reduced Employment Probability

	Model a - 0.1% of informal carers					Model b - 1% of informal carers				
Country	Gross average income (2011)	Males average income	Females average income	Aggregate d males and females	Minimum wage	Gross average income (2011)	Males average income	Females average income	Aggregate d males and females	Minimum wage
Austria	\$ 61,627	\$ 36,170	\$ 15,282	\$ 51,452	N/A	\$ 616,268	\$ 361,704	\$ 152,818	\$ 514,522	N/A

Denmark	\$ 66,544	\$ 35,132	\$ 20,884	\$ 56,016	N/A	\$ 665,436	\$ 351,319	\$ 208,838	\$ 560,158	N/A
France	\$ 368,617	\$ 157,571	\$ 133,531	\$ 291,102	\$ 195,163	\$ 3,686,173	\$ 1,575,708	\$ 1,335,314	\$ 2,911,022	\$ 1,951,629
Germany	\$ 934,413	\$ 514,102	\$ 266,651	\$ 780,754	N/A	\$ 9,344,127	\$ 5,141,023	\$ 2,666,512	\$ 7,807,535	N/A
Greece	\$ 50,190	\$ 19,762	\$ 17,176	\$ 36,939	\$ 22,225	\$ 501,902	\$ 197,625	\$ 171,762	\$ 369,387	\$ 222,251
Italy	\$ 540,676	\$ 276,867	\$ 156,118	\$ 432,985	N/A	\$ 5,406,764	\$ 2,768,672	\$ 1,561,180	\$ 4,329,851	N/A
Netherlands	\$ 115,983	\$ 68,357	\$ 30,798	\$ 99,154	N/A	\$ 1,159,833	\$ 683,565	\$ 307,979	\$ 991,544	N/A
Spain	\$ 283,211	\$ 122,027	\$ 97,459	\$ 219,486	N/A	\$ 2,832,111	\$ 1,220,267	\$ 974,590	\$ 2,194,857	N/A

17 Annex 10: Substitution Effects Variables and Evidence

Impact question	Data utilised
3. Who experiences changes?	Wolff and Kasper (2004) - Female care recipients aged 65+
4: What changes are experienced?	Data taken from Wolff and Kasper (2004) - Delayed discharge from hospitalisation (more than 2 days above average stay for same diagnostic related group).
5: How can we measure the changes?	Wolff and Kasper (2004) - Data supplied based on assessment against control groups.
6: How much of each change has happened?	Wolff and Kasper (2004) - Results segmented by informal carer characteristics; Informal carers characterised as competent = care recipient 53% less likely to suffer delayed discharge. Informal carers characterised as overloaded = care recipient 75% more likely to suffer delayed discharge. Informal carers characterised as captive = care recipient 99% more likely to suffer delayed discharge.
7: How much of each change is caused by our activities?	Evidence of counterfactual provided by the study – using control groups. No evidence of attribution included.
8: How long do we need to measure the changes for?	Results calculated for one-year, based upon results presented in Wolff and Kasper (2004)
9: What is the relative importance of the different changes?	WHO (2011) unit cost of bed per day at primary-level hospital = \$628.43

18 Annex 11: Monetized Results of Substitution Effects Based Upon the Provision of Informal Care Based Upon Sample of 420 Care Recipients

Informal carer segment	Average of hospitalized care recipients	Average delayed discharge	Relative odds compared to average	Average cost of delayed discharge (2 days at 2008 prices)	Value based upon quantity of people affected compared to average
Competent informal carer – based upon a 4-item scale, describing caregivers’ confidence and belief that they are good caregivers	24.3% = 102.06 people	29.6% = 30.21 people	53% less likely = 16.01 people	=\$628.43 * 2 = \$1,256.86	16.01 * \$1,256.86 = \$20,122
Overloaded informal carer – based upon a 3-item scale, describing the perceived discrepancy between the tasks involved in caregiving and the caregiver’s available time			75% more likely = 52.87 people		52.87 – 30.21 * \$1,256.86 = (\$28,420)
Captive informal carer – based on a 3-point scale highlighting the strain associated with being an unwilling or involuntary incumbent of their role as carer			99% more likely = 60.12 people		60.12 – 30.21 * \$1,256.86 = (\$37,593)

19 Annex 12: Willingness to Accept Values for Having Unmet Needs

Domain	Annual value
Food and nutrition No/low unmet need	£13,104
Personal care No unmet need	£21,424
Low unmet need	£12,324
Safety No/low unmet need	£5,824
Social participation No unmet need	£14,976
Control over daily living No/low unmet need	£13,870
Annual value	£60,008 to £69,108

Adapted from Netten *et al.* (2002) Table 5.8 page 42

20 Annex 13: Life-satisfaction Values of Levels For All ASCOT Domains

Domain of SCRQoL	Preference weightings and corresponding life-satisfaction values							
	Level 1		Level 2		Level 3		Level 4	
Control over daily life	1.00	£16,427	0.919	£15,096	0.541	£8,887	0	£0
Occupation	0.962	£15,803	0.927	£15,228	0.567	£9,314	0.170	£2,793
Personal care	0.911	£14,965	0.789	£12,961	0.265	£4,353	0.195	£3,203
Safety	0.880	£14,456	0.452	£7,425	0.298	£4,895	0.114	£1,873
Food & nutrition	0.879	£14,439	0.775	£12,731	0.294	£4,830	0.184	£3,023
Social participation and involvement	0.873	£14,341	0.748	£12,287	0.497	£8,164	0.241	£3,959
Accommodation cleanliness and comfort	0.863	£14,177	0.780	£12,813	0.374	£6,144	0.288	£4,731
Dignity	0.847	£13,914	0.637	£10,464	0.295	£4,846	0.263	£4,320
Totals	£118,522		£99,005		£51,433		£23,902	

21 Annex 14: Health and Wellbeing Reference Value Options to Calculate the Worth of LTC Initiatives

Measure of health/wellbeing	Monetisation approach	Source of monetisation approach	Value	Value standardised to 2017 (US\$)
QALY	WTP – UK government	NICE (2017)	£30,000	\$42,674
QALY	WTP – patient identified lower boundary	Gyrd-Hanse and Kjær (2012)	€2,720	\$3,200
	WTP – patient identified upper boundary		€96,366	\$145,061
ASCOT	Wellbeing valuation	HACT (2016)	£118,522	\$172,549

22 Annex 15: Exchange Rates of Different Measures of Health and Wellbeing

Measure	ICE-CAP-O	Life-satisfaction	SF-36	SWEMWBS	EQ-5D
ICE-CAP-O		0.88	1.13	1.06	0.70
Life-satisfaction	1.14		1.29	1.21	0.80
SF-36	0.88	0.77		0.94	0.62
SWEMWBS	0.94	0.83	1.07		0.66
EQ-5D	1.43	1.26	1.62	1.52	

Acknowledgments

Thank you to all colleagues at Social Value UK for their contributions and continued support. Thanks also to all SPRINT partners for their valued contributions, and those of the internal and external reviewers who have helped to refine this report.



The SPRINT Project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 649565



Social Protection

Innovative Investment in Long-Term Care

www.sprint-project.eu



The SPRINT Project has received funding
from the European Union's Horizon 2020
research and innovation programme under
grant agreement No 94995.