Valuing Non-Market Goods using Subjective Wellbeing Data

Daniel Fujiwara

A thesis submitted to the Department of Social Policy of the London School of Economics and Political Science for the degree of Doctor of Philosophy, London, September 2019

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).
 I declare that my thesis consists of 65,802 words (excluding the Addendum).

Acknowledgements

I would like to thank Paul Dolan for all of the support provided during my PhD. His support and advice has been invaluable throughout my PhD. A special thanks also go out to Richard Layard, Robert Metcalfe, Dominik Hangartner, Tessa Peasgood and colleagues at the Department for Work and Pensions and the UK Government Economic Service who provided guidance and advice. I would also like to thank my family for all of the support and encouragement they gave me.

Abstract

Attaching monetary values to non-market outcomes, goods and services has become a critical part of policy evaluation across OECD countries. The HM Treasury Green Book, the core policy evaluation guidance in the UK, requires that projects and policies be assessed using Cost-Benefit Analysis (CBA), which compares the benefits and costs of a policy in monetary terms and hence requires valuation of the outcomes of a policy. Outside of public policy, the private sector is also increasingly interested in valuing the outcomes of their activities to measure the social value that they generate. However, valuing non-market goods such as education, health, crime, environment, and heritage is difficult because they are not traded in markets. Wellbeing Valuation (WV) is a relatively new method, first developed in 2002. There are a number of technical problems with the method related to the statistical estimation methodology and a number of issues that have not been explored in full such as how to interpret the values. This has restricted the method's use in policy evaluation to date.

The aim of this thesis is to develop a comprehensive understanding of the WV approach and to improve the methodology so that it can be applied robustly in CBA, policy evaluation and in social value studies. I do this by developing a complete theory of WV and a new set of technical criteria to be used to assess the rigour of WV studies. I then develop a new statistical method for WV, the Three Step Wellbeing Valuation (3S-WV) method, and demonstrate how it solves for the main technical issues and improves the values and results derived from the method. I also provide a new framework for interpreting values derived from WV. I showcase the new 3S-WV method on a case study to value the non-pecuniary benefits of employment.

This thesis also contains an Addendum that was requested by the examiners, which should be read together with the thesis. The Addendum contains clarifications, changes and additions to Chapters 2, 3, 4, 5, and 6.

Contents

List of figures

List of tables

List of abbreviations

2SLS – Two-stage least squares

3S-WV – Three-Step Wellbeing Valuation

ATE – Average treatment effect

ATNT – Average treatment effect for the non-treated

ATT – Average treatment effect for the treated

BHPS – British Household Panel Survey

CBA – Cost-Benefit Analysis

CS – Compensating surplus

DRM – Day reconstruction method

EQ – Experience quality

ES –Equivalent surplus

ESM – Experience sampling method

GSOEP – German Socio-Economic Panel

IV – Instrumental variable

LATE – Local average treatment effect

LS – Life satisfaction

MRS – Marginal rate of sub

NIMBY – Not in my back yard

NOAA – National Oceanic and Atmospheric Administration

OLS – Ordinary least squares

PR – Preference realisation

QALY – Quality adjusted life year

RDD – Regression discontinuity design

RP – Revealed preference

S.E. – Standard error

SP – Stated preference

SWB – Subjective wellbeing

WTA – Willingness to accept

WTP – Willingness to pay

WV – Wellbeing valuation

Chapter 1

1. Introduction

1. Setting the scene: the role of monetary valuation in policy analysis

This thesis is about monetary valuation and how we can value non-market goods and services such as education, the environment, health, crime, and social capital. This is an important question for research because it tells us how these types of goods, outcomes and services benefit the public. The fundamental aim of government policy evaluation is to assess whether public funds are spent on activities that provide the greatest benefits to society (Hausman and McPherson, 2006). Policy evaluation makes up an important part of the activities and budget of most OECD governments. Arguably, nowhere is this more so than in the United Kingdom (UK), where HM Treasury plays a key role in verifying the effectiveness of different policy interventions and provides formal guidance on policy evaluation.

In the UK, all new policy proposals generally require HM Treasury approval, usually given on the basis of a formal Business Case. The business case is a management tool which synthesises the results of all the necessary research and analysis needed to support decision making in a transparent way[1]. Business cases are composed of five aspects:

i. The Strategic case sets out the rationale for the proposal, it makes the case for change at a strategic level.

ii. The Economic case assesses the costs and benefits of the proposal to society as a whole, and spans the entire period covered by the proposal.

[1] HM Treasury. Assessing Business Cases 'A Short Plain English Guide'. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/190609/Green_Book_guidance_short_plain_English_guide_to_assessing_business_cases.pdf

iii. The Commercial case is concerned with issues of commercial feasibility and sets asks whether the proposed solution be effectively delivered through a workable commercial deal.

iv. The Financial case looks at issues of affordability, and sources of budget funding for the project.

v. The Management case is concerned with the deliverability of the proposal in terms of the project management involved.

Although all five aspects are important, the economic case is "the essential core of the business case" [2] and this is because it is where the outcomes of the intervention are explicitly analysed and policies are assessed or ranked in terms of their worthiness (the other four cases relate to the issue of viability of the policy rather than to its worthiness). In theory, under this framework policies are determined through the economic case and then assessed whether they are viable (and hence can be implemented) using criteria set out in the other four aspects of the business case.

Guidance for assessing the economic case is set out in the HM Treasury Green Book manual (2018), which stipulates that **cost-benefit analysis (CBA)** should be used. This entails measuring all of the benefits and costs to society associated with the policy intervention in monetary terms. The preferred option is the one that has the highest net social benefits over the full life of the policy and its legacy.

CBA also has a significant and often dominant role in policy analysis in many other OECD countries. For instance CBA has a long history in the US, where it was first implemented in the 1930s by the US Army Corp of Engineers. Until that point evaluations of pubic investments were almost completely ad-hoc. The Flood Control Act of 1936 mandated that projects be assessed in terms of their benefits and costs and only those with positive net benefits should be implemented. In the US, the use of CBA at the federal level significantly increased with the issuance of two Executive

[2] HM Treasury (p.5). Assessing Business Cases 'A Short Plain English Guide'. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/190609/Green_Book_guidance_short_plain_English_guide_to_assessing_business_cases.pdf

Orders in 1981 and 1994 (respectively by President Reagan and President Clinton), confirming the government's commitment to CBA in regulatory decision-making[3] and its key role has continued to this day (Sunstein, 2018). Elsewhere central financing departments in Canada, Australia and New Zealand have produced their own technical guidance on CBA and the Green Book manual has been translated in to other languages by many foreign governments.

The demand for valuation is now also increasingly coming from the private sector, where social value is a hot topic. Organisations such as Marks and Spencer and Siemens have made efforts to quantify their impacts on local communities and through corporate social responsibility (CSR) and this often includes assessing how much value they have generated for society. Elsewhere the Social Value Act, which came into force on 31 January 2013, requires people who commission public services to think about how they can also secure wider social, economic and environmental benefits. The Act is intended to help commissioners get more value for money out of procurement. The result is that all major construction companies now need to demonstrate the social impact and value of their projects alongside the economic benefits when bidding for work and all key companies in the sector now have dedicated social value teams (e.g. Morgan Sindall, Lendlease, and Kier). Whilst private sector companies do not usually use the Treasury Business Case model and CBA, they do require information on the value of non-market goods, services and impacts and hence require robust valuation methodologies. Since the private sector effectively follows the Government's lead in this area I will focus on Government guidelines and in particular the Green Book and the methods that underlie it, namely CBA, but recognise that the issues raised and the contributions developed in this thesis are also highly relevant and applicable to the private sector as well.

CBA has its roots in welfare economics, a branch of economics that uses microeconomic theory and techniques to evaluate questions surrounding optimal resource allocation from a full societal perspective. CBA starts from the premise that

[3] http://community.amstat.org/chicago_chapter/calendar/20052006/may52006conference/downloadpresentationshistoryofcostbenefitanalysis

the only *morally-relevant good* is human welfare – welfare is the only thing that can make claims on our resources and hence decisions should aim to maximise welfare. In this sense CBA is a *welfarist* approach to policy analysis (Hausman and McPherson, 2006). Furthermore, CBA is *consequentialist*. That is, it stipulates that it is the outcomes of an action that matter and get counted, rather than anything to do with the intention or process of the action, in so far as the intentions and processes have no impact on the outcomes of an action.

CBA, thus, assesses policies in terms of their outcomes for welfare. This is done by comparing the negative outcomes (costs) to the positive outcomes (benefits), where costs relate to losses in welfare and benefits relate to gains in welfare due to the intervention. The key process in CBA is to convert all outcomes related to the intervention in to the same metric so that the costs and benefits can be compared on a like-for-like basis. CBA does this through conversion of all outcomes on to a monetary scale. In theory any metric could be used but monetisation allows us to compare outcomes to the implementation costs, which are in financial monetary terms from the start and which tend to make up a large part of the costs of an intervention. Monetisation is also useful as it allows decision makers to assess the overall impacts of an intervention in terms that they will be familiar with since return on investment figures and other metrics used in business and organisational decision making are usually set out in financial or monetary terms.

Monetary valuation is therefore a core component of modern-day policy evaluation and so it is critical that we have robust methods for valuing outcomes, goods and services and that we continue to develop new methods. Since CBA is welfarist and absolutist about this, the key point in CBA is that the money metric/value *must only* represent impacts or changes in welfare associated with the outcomes of the intervention. In theory a monetary value is, therefore, simply a measure of how people's welfare changes. How this could be measured in theory and in practice was the centre of debate in policy evaluation in the 1800s when the idea of valuation was first put forward. It found its solution when CBA was formally conceptualised as an offshoot of welfare economics (Backhouse, 2002). John Hicks' theory of valuation – first developed in the 1930s - now sets the fundamental theoretical basis of valuation in economics and CBA. The theory, in a nutshell, states that the monetary value of

some outcome, good or service is the amount of money that would be required to have the same effect on someone's welfare. Hicks developed two basic measures of value known as compensating and equivalent change measures. I will discuss these in detail later in the thesis. The theory is a purely abstract one - Hicks did not provide a methodological framework for measuring values using this theory. Indeed, initially Hicks did not even provide or stipulate what welfare is and how it should be measured.

This all came later in economics through the important theoretical work of Paul Samuelson's utility theory, which set preference satisfaction as the central measure of welfare. This led to the development of valuation methods based on people's choices and preferences such as the stated preference approach and the revealed preference approach. A third distinct category of valuation methodology is wellbeing valuation which was first introduced in 2002. This method moves away from defining welfare in terms of preference satisfaction to defining welfare in terms of people's self-reported feelings through subjective wellbeing (SWB) measures. It is a method that is still in its infancy and is in development, but it promises many key advantages over the traditional preference valuation methods.

This thesis focuses on the wellbeing valuation method as an alternative to preference-based valuation methods with the ultimate aim of developing and improving the wellbeing valuation methodology so that it can be applied robustly in CBA, policy evaluation and in social value studies conducted outside of Government in the private sector.

1.2. Purpose and scope of the PhD thesis

The literature on wellbeing valuation (WV) to date has focused almost entirely on the application of the method to new areas – in other words using the relatively new and novel WV approach to value more and more different types of non-market goods and services, such as safety (low crime), health, environment and education. Whilst this has thrown up many interesting findings, this research trend has come at the expense

of a serious attempt to really understand, develop and improve the theory and methodology behind WV.

There are a number of important questions and issues regarding WV which have not been discussed let alone solved in the literature; it is still unclear how WV relates to the economic theory of valuation, how it should be seen against other (traditional) valuation methods such as stated preference and revealed preference methods, what the key validity and robustness criteria and conditions are, and what the values derived from WV really mean. Add to this the significant problem that has been highlighted in the literature that the values from WV seem to be implausibly too high and inaccurate.

Given the importance currently placed on valuing non-market outcomes, goods and services, this thesis examines the wellbeing valuation approach with the aim of developing a comprehensive understanding of the method and a new and improved methodology to solve for many of the current technical issues related to the approach to increase the rigour and robustness of the method and ultimately its application in policy evaluation. The aim is to provide a platform for using wellbeing valuation in CBA and policy analysis and in social value assessments undertaken by private sector companies, which will ultimately itself have a social impact since it will allow us to better understand how our actions and policies benefit society and to make decisions according to that evidence.

1.2.1. Original contributions of this thesis

This thesis makes a significant contribution to the literature by building on previous research in the following ways:

i. **Valuation theory.** Hicks' theory of compensating and equivalent surplus is the accepted theory for valuing non-market goods in economics. To set the context this thesis takes Hicks' theory as a given and contributes to the literature by providing a clear description of how WV aligns with Hicks' valuation theory and under what conditions. I am not aware of any study to

date to have done this. This is a critical first step in developing and improving the WV method because it provides a theoretical benchmark that values derived from WV need to align with in order to be robust.

From this I will develop a full theoretical exposition of the wellbeing valuation approach to valuing non-market goods. This will allow me to create a set of theoretical conditions and criteria which the WV method is required to adhere to in order to produce robust values in line with Hicks' valuation theory. This is a significant contribution to the literature in WV because there does not currently exist a set of technical conditions and criteria against which to assess WV studies. It is the crucial first step in improving the WV method so that it can be used to produce robust values.

ii. **Rationale for wellbeing valuation.** I build a rationale for using WV to value non-market goods and services. This is partly reliant on some of the problems with the more traditional preference-based valuation methods as well as the unique advantages of WV itself. The problems associated with preference-based valuation methods are numerous and have been well documented and rehearsed in many previous publications. I will provide a summary of the key problems which are relevant to the rationale for WV, but also provide some new insights and problems related to preference-based valuation methods which have not been discussed before.

In developing the rationale for wellbeing valuation I will also use the new estimation criteria developed in this thesis for WV to provide a full critique of the current wellbeing valuation methodology and studies to date. Whilst there have been a number of papers in the past that have provided some critical assessment of the method, the critique in this thesis will be more extensive and comprehensive developing new ideas and critiques that have not been discussed, recognised or understood before.

iii. **Development of a new approach to wellbeing valuation – *The Three-Step Wellbeing Valuation* method**. The key and main contribution of this thesis is the development of a completely new statistical/technical approach for

wellbeing valuation. Three-Step Wellbeing Valuation provides a solution to the technical problems associated with the method as described in this thesis and ensures that the values derived from WV are robust and in line with Hicks' valuation theory. Three-Step Wellbeing Valuation represents a significant modification to the WV methodology and no other study in the literature to date has used a similar methodology. It is, therefore, a significant original contribution to the literature. I demonstrate the new method with a labour market case study where I value the non-pecuniary aspects of employment.

iv. **Interpretation.** I provide a full interpretation of values derived from the wellbeing valuation approach. This covers many new areas that have not been discussed in the literature to date and is an important contribution as it is crucial to understand in order to apply the values in the correct way in policy evaluations.

Given that the field of research in monetary valuation is wide and ever-growing, it is important to state what this thesis will not cover. The thesis will work within the current theoretical framework underlying welfare economics and CBA. That is, I do not provide a defence of CBA or of Hicksian valuation theory. Although this is an important area of research it is out of the scope of this thesis. Hicksian value theory is a welfarist approach to valuation and as a consequence in this thesis I do not attempt to defend the welfarist paradigm. I take the current methodology as given and discuss the role and relative advantages and disadvantages of the wellbeing valuation approach within this context. Also, at a broader or deeper level I do not discuss the merits of valuation per se – i.e. whether it is right or wrong to place monetary values on things. There is an extensive literature in this area and many of the arguments have been captured well in books by Anderson (1995) and Sandel (2013). I will take as my starting point that there are good reasons to value things as I have set out in the first part of this Chapter. What this means is that I will also not cover recently developed proposals to use wellbeing as the evaluation metric rather than monetary values in a form of cost-effectiveness analysis whereby costs are compared against the wellbeing impacts of a policy or programme (for example see Wright et al., 2017). This thesis is

intended to contribute to methodological issues in valuation rather than first-order theoretical issues related to monetary valuation.

1.2.2. Structure of the thesis

The thesis is set out as follows. In Chapter 2 I set out the theory of monetary valuation in economics. This will act as the core theoretical framework for the whole thesis. As we will see valuation theory in economics is based on wellbeing, but it is agnostic about the measure of wellbeing used and thus to this end I will follow the theory with a discussion of the key measures of wellbeing. I end this chapter by then discussing the traditional valuation methods in economics which specifically use what is known as the preference satisfaction account or measure of wellbeing. I will review the critical literature relating to these methods. Chapter 2 will therefore set the scene in terms of what valuation is and how it is defined and where the methods are currently. Discussing the key problems of the traditional valuation methods provides important rationale for the wellbeing valuation method as an alternative approach to non-market valuation; I will show that there are significant problems with the current methods and that the field would benefit from a new approach that could solve for these problems.

Chapter 3 is devoted to the wellbeing valuation method. I first set out the underlying theory and methodology of the wellbeing valuation method. I then discuss the main advantages and disadvantages of the wellbeing valuation approach in comparison to preference-based valuation methods and provide the full rationale for wellbeing valuation. The chapter ends with a defence specifically of the life satisfaction measure that is most frequently employed in wellbeing valuation. Although as I have stated in Chapter 1, I do not aim in this thesis to provide a defence of welfarism per se it seems appropriate to spend some time discussing the validity of the SWB measure that has formed the basis of wellbeing valuation to date. Chapter 3, therefore, builds logically from Chapter 2 to introduce a new method - the wellbeing valuation approach - and to provide rationale and support for its use in light of the problems with the current more traditional valuation methods.

Chapter 4 then assesses the current wellbeing valuation method with the aim of improving it. I develop a framework for assessing the validity of wellbeing valuation studies based on a set of new criteria that I establish in this thesis. I then develop a full theoretical approach to measuring values in wellbeing valuation. I end the chapter by assessing the current wellbeing valuation methodology against the validity criteria that I have developed and demonstrate that there are a number of problems with the current wellbeing valuation methodology.

This sets the background for Chapter 5, where I set out a new approach to wellbeing valuation, the Three-Step Wellbeing Valuation approach. I derive the estimation framework and procedure for Three-Step Wellbeing Valuation and discuss how it provides a solution to the key technical challenges in wellbeing valuation and how it adheres to the validity criteria set out earlier in the chapter. A key part of the estimation process in all wellbeing valuation studies is the estimation of the impact of income on SWB. In this chapter I set out a generic model to estimate the impact of income in a robust way. This model sits at the core of the Three-Step Wellbeing Valuation approach. I argue that Three-Step Wellbeing Valuation provides a more robust method of valuing non-market outcomes using subjective wellbeing data and one which aligns with Hicks' value theory. The chapter also provides a detailed discussion of how values from wellbeing valuation should be interpreted and used in CBA and what their relation is to preference values.

In Chapter 6 I demonstrate the new Three-Step Wellbeing Valuation methodology with a case study of employment-related values. Employment outcomes have been systematically undervalued in traditional valuation methods and CBA, which have tended to focus only on the income-related benefits of employment at the expense of missing the important impacts on health, relationships, self-esteem, social stigma and personal identity. Wellbeing valuation provides a highly suitable framework for picking up and valuing these non-financial benefits, which as Greenberg and Robins (2008) argue should be part and parcel of any CBA on employment. I derive values associated with moving from unemployment to employment and I show the improvements in estimation gained from using Three-Step Wellbeing Valuation in comparison to current wellbeing valuation methods.

Chapter 7 is the concluding chapter which focuses on the policy implications of the findings in this thesis and concluding remarks. I also provide some recommendations for future research.

Chapter 2

2. Valuation methods

2.1. Introduction

This chapter provides the important opening setting for the thesis. It provides a description of where the field of valuation currently sits, what is wrong with it and why it needs improving.

I start by setting out the theory of valuation in economics. This is an uncontested theory and is generally the consensus in economics. It provides a barometer or target against which the rigour of valuation methods can be assessed and will be referenced and discussed numerous times during the thesis. The chapter then goes on to describing the traditional valuation methods in economics that use preference as the measure of welfare: stated preference and revealed preference valuation methods. It finishes with a critique of these methods, which are well-known in economics. This provides some of the rationale and reason for exploring and using a different method for valuation, wellbeing valuation. Chapter 3 will then go on to set out a full rationale for wellbeing valuation, borrowing heavily from the critiques discussed in this chapter.

2.2. The theory of valuation in economics

The theory of monetary valuation developed in economics is fully consistent with the underlying welfarist paradigm in economics and CBA and is therefore the underlying theory of the Green Book and other related policy manuals.

The value of a good or service relates to the impact that it has on human welfare (Champ et al., 2003). The theory is rich and there exist a number of possible ways of measuring welfare in monetary terms, which all derive from the same fundamental

welfarist premise. A good place to start is the **money metric** measure of welfare. This defines some level of welfare, or utility as it is known in economics, in terms of the expenditures needed to attain that level of welfare. The standard assumption in economics is that individuals maximise a utility function with respect to market and non-market goods subject to a budget constraint:

$$(1) \qquad \max_{X} \; U(Z, Q) \qquad \text{s.t.} \;\; P \cdot Z \leq M$$

where Z = market goods; Q = non-market goods; P = prices and M = income.

The dual problem for the individual is to minimise expenditures subject to obtaining a given level of utility (U^*):

$$(2) \qquad \min_{X} \; P \cdot X \qquad \text{s.t.} \;\; U(Z, Q) \geq U^*$$

This produces the expenditure function:

$$(3) \qquad e = e(P, Q, U^*)$$

which shows how expenditure changes as a function of the prices of market goods and provision of the non-market good, such that the individual continues to maximise utility at the level U^*.

The expenditure function provides a money metric measure of welfare as it shows the minimum expenditure required to obtain the same level of welfare (U^*) as with Z and Q. This measure of the monetary equivalent of some *level* of welfare can be used to assess the monetary value of *changes* in welfare due to non-market goods and services, which represents the monetary value of the non-market outcomes themselves.

Hicks (Hicks and Allen, 1934) set out two measures of monetary value, known broadly as **compensating and equivalent welfare measures**. These can be derived from the expenditure function as in equations (4) and (5). Here I will focus on

compensating and equivalent surplus rather than variation measures. Surplus measures differ from variations in that the latter are calculated after the individual has made adjustments to his consumption set (Randall, 1982) and hence relate to price changes. Thus compensating/equivalent variation relates to price changes, whilst compensating/equivalent surplus relates to quantity or quality changes, which applies to non-market outcomes.

(4) CS = $e(P^0, Q^0, U^0)$ - $e(P^0, Q^1, U^0)$

(5) ES = $e(P^0, Q^0, U^1)$ - $e(P^0, Q^1, U^1)$

Where CS = compensating surplus, ES = equivalent surplus and the 0 and 1 superscripts refer to before and after provision/consumption of the non-market good (Q).

In words, CS is the amount of money, paid or received, that will leave the agent in his initial welfare position following a change from the status quo. And ES is the amount of money, to be paid or received, that will leave the agent in his subsequent welfare position in the absence of a change from the status quo. Here the change is in the form of changes to the quantity or quality of non-market goods represented as $Q^0 \rightarrow Q^1$.

There are two important points to note here. First, this is a theory about value to the individual. This is known as the *primary benefits* of non-market outcomes and relate to the value of impacts directly on an individual's welfare. There are also *secondary benefits* that can be valued separately. Secondary benefits relate to impacts that benefit society more widely which at some point may be an indirect benefit to the individual as well. This mainly encompasses impacts on the economy and public purse. This could be, for example, reductions in medical service usage due to improved health or increases in tax receipts due to rises in employment.

The two types of benefit are important for different sectors of society and both are included in CBA. As individuals it is highly unlikely that we make any personal decisions based on secondary benefits – we go to the doctor to get better and not to

reduce medical expenditures to the state later on (in fact we increase medical expenditures by going to the doctor in the first place) and we care about safety primarily because crime has significant adverse effects on our wellbeing and not because crime incurs costs to the state (policing, courts, prisons etc). But as policy makers or as (civic) individuals in instances where we are making decisions for the good of the community or society we also care about secondary benefits since this allows us to provide more or better services to people.

The focus of this thesis and of the valuation theory discussed above is on the primary benefits to the individual as this is where wellbeing valuation can be employed.

The second point of note is that the theoretical framework can be derived without recourse or reference to any concrete measure of welfare as equations (1) to (5) demonstrate. And indeed Hicks, in his pioneering work on the theory of value (1934), did not initially propose a specific measure of welfare to be used in calculations of CS and ES. And so how these measures of value and welfare change would be assessed in reality was not clear until economists started to adopt a standard measure of welfare in empirical work. This came to be the *preference satisfaction account of welfare*, to which we now turn in the next section. This is in contrast to other forms or theories of wellbeing; broadly speaking there are three accounts of human wellbeing (Parfit, 1984):

1. Mental states and the self-reported experience of the individual.
2. Preference satisfaction.
3. Objective lists encompassing normative ideals.

These are discussed in detail in Chapter 3. Briefly, mental state accounts of welfare are based on people's self-reports about how their lives are going, whilst the objective list account is based on normative assumptions about basic human needs and rights (Dolan et al., 2011). The WV method uses the mental state account of welfare and hence the basis of this thesis is a comparison between valuation methods that use the preference account and those that use the mental state account of welfare. We start with an assessment of the preference-based valuation methods before moving on to introducing the WV method in the next chapter.

2.3. Traditional methods: Preference-based valuation methods

Although Hicks' pioneering work on the theory of value did not specify how 'welfare' should be defined and measured, work in the early twentieth century by economists such as Paul Samuelson and Roy Allen directed economics to what is known as the preference satisfaction account of welfare and consequently the theory of value followed suit (Hicks, 1934). This move to a preference account of the world is termed the 'Paretian turn' by Bruni and Sugden (2007. p.146) in recognition of Vilfredo Pareto who had initiated the transition of economics to a theory of rational choice. The preference satisfaction account states that "what would be best for someone is what would best fulfil his desires" (Parfitt, 1984. p.4).

In what Wong (2006) terms the 'Samuelson Programme' we see various economists contributing to the revealed preference approach to the theory of consumer behaviour. This is a well-known theory that forms the basis of modern economic theory and hence will not be discussed in great detail here, but the fundamental premise is that under a small set of rationality assumptions embodied in the Axioms of Revealed Preference we are able to map choices over a number of binary options on to a well-defined utility function. Rationality here implies that preferences are:

i. **Complete** – individuals are able to express a preference for any good or say they are indifferent between any pair of goods;
ii. **Transitive** – individuals who prefer (or are indifferent to) good x over good y, and who prefer (or are indifferent to) good y over good z, must also prefer (or be indifferent to) x over z; and
iii. **Reflexive** – individuals are indifferent between x and x.

If these assumptions are met then people will behave *as if* they are maximising some utility function. And it is important to note that these are 'assumptions'. Indeed, Pareto's *integrability problem* was that it was not possible to prove that preferences are transitive in some commodity space and so transitivity is only a mere "speculative hypothesis" (Bruni and Sugden, 2007. p.160) Economists are generally very reluctant to make normative claims about agents but if we add a further substantive assumption that people act to maximise their own welfare then preference satisfaction, utility and

welfare all become synonymous with each other and preferences can be used as the basis of welfare in valuation. For the purposes of valuation we need to add two further assumptions (Champ et al., 2003):

iv. **Non-satiation** – that preferences are never fully satiated such that the individual always places a positive value on more consumption; and
v. **Substitutability** - if the quantity (or quality) of one good decreases it is possible to increase the quantity (or quality) of another good sufficiently to make the individual indifferent between the two states of the world.

In CBA preferences are usually taken as they are - as actual non-idealised preferences (Adler, 2012; Champ et al., 2003), but in practice policy makers may require that preferences be informed to some extent for the purposes of policy analysis. Apart from these assumptions economics makes no further normative claim about how preferences should be. In contrast to Kahneman's (2000) substantive rationality criterion - whether preference and choice maximise wellbeing as experienced by the individual (Kahneman states that this measure of wellbeing should be a hedonic measure) - the early founders of the current approach to rational choice (through preference) in economics were not interested in the basis or reasons for preference (Bruni and Sugden, 2007).

As Samuelson states welfare economics rests on "one fundamental ethical postulate", that "the preferences of individuals are to count in the allocation of resources" (from Sagoff, 2003. p.588). Under the preference satisfaction account of welfare, higher levels of utility denote a greater number of preferences satisfied. Utility in this sense is not observable, but a number of methods exist for measuring welfare change and monetary value using preferences.

Under preference valuation methods CS and ES are often rephrased in terms of **willingness to pay (WTP)** or **willingness to accept (WTA)** and Table 1 describes the relationship between these concepts of value.

Table 1. The relationship between CS, ES, WTP and WTA

	Compensating Surplus (CS)	**Equivalent Surplus (ES)**
Welfare gain	*WTP for the positive change*	*WTA to forego the positive change*
Welfare loss	*WTA the negative change*	*WTP to avoid the negative change*

As I discuss in more detail below it is important to recognise that technically speaking the concepts of WTP and WTA are only relevant to preference-based valuation methods. As I will set out later we should not use these terms in WV.

Conceptually there are two distinct ways of measuring CS and ES using preferences: (1) ***Direct approaches*** (termed *income compensation approaches* by Randall (1982)) measure value directly in terms of the money required to restore some level of welfare, whereas (2) ***Indirect approaches*** use data on expenditures on related marketed goods to infer values for non-market goods and services (Randall, 1982). The latter uses the expenditure function framework to achieve this, whilst direct approaches can be derived using the direct utility function as I demonstrate below. Revealed preference and stated preference methods make up two of the recommended valuation methods in the Green Book (2018) and OECD (2018) guidelines.

2.3.1. Revealed preference valuation methods (indirect approaches)

Generally speaking where proxy markets exist the favoured approach is to estimate WTP or WTA from people's market behaviour using **revealed preference valuation methods**. Revealed preference methods uncover estimates of the value of non-market goods by using evidence of how people behave in the face of real choices. The basic premise is that non-market goods affect the price of market goods in other well-functioning markets and price differentials in these markets can provide estimates of WTP and WTA. This exploits the expenditure function under the assumption of weak complementarity whereby the demand for a market good depends on the level of the non-market good. There are a number of methods that can be used here and two of the most common revealed preference methods are:

(i) Hedonic Pricing Method. This involves examining people's purchasing decisions in markets related to the non-market good in question. The hedonic pricing method has most commonly been applied using data from housing and labour markets. In the former, the intuition is that the price differential between otherwise identical houses that differ in their exposure levels to non-market goods and bads such as pollution, noise, crime or education facilities reveals information regarding individuals' WTP for such goods. Labour market applications follow a similar logic, though the focus is typically on the compensating wage differentials that are paid in relation to job characteristics such as health and safety risks or job security.

(ii) Travel Cost Method. This involves observing costs incurred in the consumption of the non-market good in question. The travel cost method has most predominantly been used to estimate the value of recreational sites (e.g. a river, a park, or a beach). It has also been used to value changes in the characteristics of sites (e.g. ease of access). The number of visits to a site by an individual over a period of time is likely to be related to the price they have to pay to visit the site, the travel costs incurred, the price of substitute sites available to them, and their income. This information can be used to model demand curves and hence WTP for the sites.

Behaviour can also be observed through the actions people take to insulate themselves from things that lower their utility and this forms the basis for the **defensive expenditure method** for non-market valuation. For example, in response to traffic noise or air pollution, households may purchase double glazed windows or hire window cleaners. Therefore, expenditures on market goods can be related to levels of non-market bads.

Revealed preference is an indirect method because it utilises the expenditure function to circumvent the need to observe or measure utility. Under the assumption that the non-market good is an argument in the demand function for market goods, market data (prices and quantities consumed) can be used to "reveal the welfare impact of changes in Q [the non-market good]" (Randall, 1982. p. 152).

2.3.2. Stated preference valuation methods (direct approaches)

Although economists are generally very wary of self-reported or non-behavioural data due to incentive compatibility issues very often proxy markets do not exist for the non-market good in question and hence we may need to ask people about their WTP or WTA instead. This makes up the basis of so-called **stated preference valuation methods**, which use surveys to ask people directly about the value they place on a good or some attributes of a good.

Contingent valuation methods construct and present a hypothetical market to questionnaire respondents. A detailed description of a good, how it will be provided, and the method and frequency of payment are usually highlighted. Following this, questions are posed in order to infer a respondent's WTP or WTA. These valuation questions can be presented in a number of different ways, including open ended, bidding game, payment card, and dichotomous choice elicitation formats.

Contingent valuation questionnaires also normally contain additional questions to gain information on a respondent‘s socioeconomic and demographic characteristics, their attitudes towards the good, and the reasons behind their stated valuations. The responses to these questions are typically used to model the determinants of stated valuations so that econometric analysis can be used to filter out the proportion of WTP (or WTA) related directly to the good.

Non-market goods can also be described by their attributes. For example, a scheme to reduce sewage overflows into the River Thames could be described by the resultant reduction in fish deaths, health risks, and visual disamenity (Mourato et al., 2005).

Choice modelling methods focus on a good‘s attributes and their values. To uncover valuation estimates, choice modelling questionnaires present respondents with a series of alternative descriptions of a good. The alternative descriptions are constructed by varying the levels of the good‘s attributes. Depending on the specific choice modelling method adopted, respondents are either then asked to rank (contingent ranking), chose (choice experiments), rate (contingent rating), or choose then rate (paired comparisons) the descriptions presented (Hanley and Shogren, 2005). For

these methods, as long as cost or price is included as an attribute, statistical techniques can be used to recover WTP estimates for the other attributes of the good.

Stated preference is a direct method in that it seeks to directly estimate the amount of money that compensates or equates to the change in welfare due to the non-market good. For example, the following functions can be used to estimate CS for a non-market good:

(6) $U(Q^0, M^0) = U(Q^1, M^0 - CS)$

(here CS = WTP for the positive change $Q^0 \rightarrow Q^1$)

(7) $U(Q^1, M^0) = U(Q^0, M^0 + CS)$

(here CS = WTA for the negative change $Q^1 \rightarrow Q^0$)

In stated preference CS (the solutions to equations (6) and (7)) is estimated from people's stated WTP and WTA. Although we cannot directly observe utility (U) in equations (6) and (7), stated preference represents a much more direct approach targeting estimates of WTP/WTA without recourse to market data. In this set up we are reliant on the respondent accurately estimating the welfare change due to the non-market good *and* the amount of money required to produce the equivalent impact on wellbeing. This is an *identifying assumption* in stated preference because it is not something that can be tested.

2.4. Critiques of preference-based valuation methods

Preference-based valuation methods have been found to suffer from a number of problems. The literature in this area is extensive and includes whole special edition journal publications devoted to it. Since these problems have been well-documented in the literature I will cover the main critiques that are relevant to this thesis and the comparison ultimately to wellbeing valuation. These critiques form the first part of the rationales for wellbeing valuation that will I develop in subsequent chapters.

The problems related to preference-based valuation methods can be categorised in to those that (a) critique the validity, rigour and usefulness of preference as a measure of welfare per se, which has implications beyond valuation for any application of preference such as quality adjusted life year analysis in health economics, and those that (b) have found technical problems and issues in the methods themselves. The first category of critiques fundamentally questions whether it is right to use preferences in valuation and the second category demonstrates that even if we accept preference as a measure of welfare for valuation we run into a number of technical problems when applying preferences in RP and SP approaches.

2.4.1. Substantive critiques of the preference satisfaction account of welfare

The preference satisfaction account of welfare has come under increasing attack from the behavioural economics sub-discipline and the fields of the psychological sciences that underpin it. The following critiques focus on the extent to which preference can be relied on as a measure of welfare. This is clearly an important issue for valuation methods that use preferences, but it also has wider implications in relation to the use of preference per se in policy analysis. This thesis does not cover the latter as the focus is on valuation methods.

Sagoff (2003) argues that preferences may not fundamentally align with individual welfare. An individual's preference may have all sorts of motives. Findings from experimental research suggests that people are committed to goals other than their own welfare, especially when making choices about policy (Sagoff, 2003). Similarly, Sen (1977) discusses the issue of choice based on 'commitment values', whereby a person chooses an act that yields a *lower* level of welfare in order to fulfil and commitment.

RP and SP methods both suffer from the fundamental problem of the *context sensitivity* of preferences. Although traditionally economists have tended to see preferences as stable, consistent and uniform, a large and growing literature in the decision sciences has shown that preferences can be highly context-dependent (see Slovic and Lichtenstein, 2006); they can often be biased by irrelevant factors, which

mean that what people want may not always align well with what is best for them. Experimental evidence suggests that context sensitivity is equally problematic for RP and SP methods (Carlsson, 2010).

There are countless experimental studies in this area and so it is out of the scope of this thesis to provide a full dissection of the results here. Instead, I provide a discussion of the main findings, which will provide the basis and rationale for the new approach proposed in this thesis, which uses measures of people's self-reported wellbeing to value non-market goods rather than their preferences.

1. If preferences are to be accurate indicators of our welfare it is obvious that people need to accurately predict how much they will like in the future the thing that they show a preference for now (Kahneman, 2000). But numerous experiments have shown that people are unable to accurately predict the pleasure or benefits they will get from different goods and services (Loewenstein and Schkade, 1999; Loewenstein and Adler, 1995; Read and van Leeuwen, 1998; Simonson, 1990; Loewenstein et al., 2003; Wilson and Gilbert, 2003) and this is true even for everyday goods such as music and ice cream (Kahneman and Snell, 1992). In Kahneman and Snell's (1992) study participants were asked to consume yogurt and their favourite flavour of ice cream and to listen to their favourite music each day for a week. They rated their liking of the goods after each consumption and also predicted their liking and enjoyment of the goods for the following day. Correlations between predicted and actual enjoyment were negligible even in relatively large sample sizes. Nisbett and Kanouse (1969) and Read and van Leeuwen (1998) find evidence that shoppers who have recently eaten cannot forecast their future food consumption and appetites accurately.

Prediction errors are exacerbated when the temporal gap is long (ie, when they try to predict preferences far into the future) and when the agent’s circumstances vary over the period (Kahneman and Thaler, 2006). One of the drivers of this phenomenon is that people are unable to predict how much they will adapt to different things and circumstances in the future. They therefore tend to over-estimate the utility gain that will result from events, circumstances or outcomes (Kahneman and Thaler, 2006; Loewenstein and Adler, 1995). Frey and Stutzer (2004), for example, argue that

people underestimate how quickly they will adapt to extrinsic goods, such as money. They, therefore, end up sacrificing too many intrinsic goods, such as time with family and friends, for time spent at work and commuting. Schkade and Kahneman (1998) present evidence that people are not able to predict the satisfaction they would derive from moving from the Midwest to California. Individuals tended to focus on one or two salient aspects associated with California, such as the weather (which in reality does not feature so saliently in people‘s actual day-to-day lives), when forecasting utility.

Asking people about how something will affect their lives or about their preferences between different states of the world often leads to a *focussing illusion* (Kahneman et al., 2006; Schkade and Kahneman, 1998), whereby at the time of preference elicitation people are focusing only on the salient aspects of the condition and this may not reflect in any way how people would actually experience these conditions or states in real life. The fundamental problem is that what we focus on in a preference question is often not what we focus our attention on in the actual experiences of our lives, where lots of other phenomena vie for our attention and we may adapt to certain things. (Dolan and Kahneman, 2008). As Kahneman (2012) puts it "nothing is as important as you think it is when you are thinking about it."

Gilbert puts these welfare or utility mis-predictions down to a *presentism* heuristic, whereby people generally find it hard to predict how much they will like something and use a short cut method of simply projecting current tastes and desires on to their predicted future preferences.

Finally, Kahneman (2000) states that retrospective evaluations of previous experiences are the main sources of predictions of future outcomes and quality of experience. But these memories are fallible because of the peak-end rule, whereby people tend to remember the quality of an experience by the most extreme affect and the experiences during the final moments of the activity. This can lead to duration neglect, whereby the remembered quality of an experience is not dependent on the duration of the episode. "Affective peaks and endings are more salient than duration in the cognitive representation of events" (Kahneman, 2000. p.769).

Whatever the driver behind these mis-predictions may be, it is clear that current contextual factors will have big effects on how much we state or reveal we prefer something.

2. The fundamental assumption behind the use of preferences as indicators of welfare is that they are rational preferences. As discussed, this allows us to map people's choices on to well-defined utility functions. Adding a non-satiation assumption then allows us to use rational preferences for the purposes of valuing non-market goods. A key axiom of the rationality assumptions is that preferences are transitive. This simply means that individuals who prefer (or are indifferent to) good x over good y, and who prefer (or are indifferent to) good y over good z, must also prefer (or be indifferent to) x over z.

The transitivity and rationality axioms are severely challenged by the phenomenon of *preference reversals*. This is when someone or a group of people initially shows a preference for A over B, but then when the same information about A and B is presented in slightly different ways they prefer B to A. Preference reversals violate the rationality assumptions making it difficult to judge which state of the world ultimately makes the individual better off.

One of the first and certainly one of the most famous examples of preference reversals was found in Slovic and Lichtenstein's 1971 (see Slovic and Lichtenstein, 2006) experiments on preferences over different gambles. People were offered two different bets of the same expected value; a probability bet (high probability of winning a small amount – eg, an 80% chance of winning $5) and a dollar bet (low probability of winning a large payout - eg, a 10% chance of winning $40). In lab experiments as well as field experiments in casinos the overwhelming majority of people chose to play probability bets over dollar bets, but when both of the bets were given to them and they were asked to sell them back to the House, the majority assigned higher prices (higher WTA values) to the dollar bet - ie, they preferred the dollar bet when selling back to the house. This was explained by people using different information under the two tasks; people put an emphasis on probabilities when making a choice and then conversely they put an emphasis on payout when stating a price.

Another example is preference reversals under separate versus joint evaluation modes, made famous by Hsee's work (e.g. Hsee, 1998; Hsee and Hastie, 2006). Here people use different aspects of the same information set when jointly evaluating a good (say two different TV sets) rather than evaluating it on its own. Hsee's (2000) music dictionary study asked students to state their WTP values for the following two music dictionaries.

	Dictionary A	**Dictionary B**
Year of publication:	1993	1993
Number of entries:	10,000	20,000
Any defects?	No, it's like new.	Yes, the cover is torn; otherwise, it's like new.

Respondents were assigned to three different groups: i) subjects who were shown both dictionary descriptions and asked to state their WTP for each (joint evaluation mode); ii) subjects who were shown dictionary A only and asked to state their WTP for that dictionary (separate evaluation mode); and iii) subjects who were shown dictionary B only and asked to state their WTP for that dictionary (separate evaluation mode).The mean WTP values for the two dictionaries are shown in the table below.

Preference reversals in joint and separate evaluations

Evaluation Mode	**Dictionary A**	**Dictionary B**
Joint	$19	$27
Separate	$24	$20

Source: Hsee (2000).

Under joint evaluation, people state a higher value for dictionary B. However, under separate valuation, dictionary A attracts the highest stated value. These joint evaluation– separate evaluation preference reversals can be explained by some simple heuristics. In separate evaluation people focus on the categorical attributes of the good, in this case 'whether the dictionary has any defects'. In joint evaluation, attention is focused on the incremental aspects or differences in the goods, in this case 'the number of additional entries'.

In a well-documented study two economists, Grether and Plott (1979), criticised the previous work on preference reversals by psychologists and replicated the experiments introducing improved incentive compatibility, a wider and more varied range of participants and more information for participants. However, the preference reversal phenomenon did not disappear.

Importantly, preference reversals have also been found in contingent valuation surveys for environmental goods and amenities (Irwin et al., 1993; Gregory et al., 1993; Brown, 1984) and safety programmes (Slovic et al., 2002).

Slovic et al. (2002) employed a phenomenon known as *proportion dominance*. People attach greater weight to information formats that use proportions, percentages or probabilities, rather than absolute figures because these formats put the outcome dimension into perspective. Proportional formats have upper and lower bounds which allow people to place where a given value falls (Slovic et al., 2002). In a study on airport safety equipment, people in different groups were offered equipment that, in the event of a crash landing, would (i) save 150 lives or would (ii) save 98% of 150 lives (147 lives in total). In general people placed a higher willingness to pay or value on the equipment that saved less lives. In fact it was found that saving 98%, 95%, 90% and 85% of 150 lives were all more valuable options than saving 150 lives.

The manner in which information is presented has a huge impact on people's preferences and their willingness to pay for an outcome/good. The presentational issues described here should be irrelevant to the choices that people make and the value that they place on outcomes because fundamentally people have access to the same underlying information. The problems related to preference reversals means that it is quite possible for survey developers and enumerators to force or induce the results they want in stated preference studies and this is a huge problem for preference valuation methods.

Smith and Moore (2010) discuss some evidence (albeit from laboratory experiments rather than real-world market settings) to show that the presence of irrational agents (those holding inconsistent preferences) can distort market efficiencies and

performance. This has implications for using revealed preference methods with market data.

It is important to note that even if preferences are rational (coherent), we may still want to reject them if they do not align with welfare and this could be due to the effects of utility mis-predictions discussed above. Kahneman calls this a new *substantive rationality* challenge for preferences.

3. The environment in which people are placed can also provide some other cues or nudges in regards to their preferences. People tend to systematically anchor their values for non-market goods on irrelevant numbers or cues that appear in the environment at the time. This applies to both real market scenarios and to SP. For example, real estate agents are influenced by random house listing price anchors when valuing a property (Northcraft and Neale, 1987).

Ariely et al. (2003) found, for example, that people's WTP for a range of everyday consumer goods and their WTA values for small annoyances, such as high pitched sounds, were heavily anchored around their social security (SS) numbers. People were asked to write down the last two digits of their SS number and were then asked whether they would be willing to pay or accept a value equal to that number. Values were then increased or reduced from the initial SS number anchor until the respondents' maximum (minimum) WTP (WTA) values were derived. US SS numbers are randomly generated, which means that they could not provide any information on the quality of the good. In general, people with higher SS numbers were willing to pay significantly more for the goods. An interesting second finding was a marked stability of relative preference. For example, although people's absolute valuations of a superior and inferior wine were subject to normatively irrelevant number anchors, the vast majority of people valued the highly rated product more than the inferior product. Therefore, the evidence suggests that people did not know how much they were truly WTP for each of the wines, but they did know that they were WTP more for the superior wine. This, and other evidence, lead the authors to claim that people's preferences and valuations were coherently arbitrary; "consumers' absolute valuation of experienced goods is surprisingly arbitrary, even under "full information" conditions. However, consumers' relative valuations of different

amounts of the good appear orderly, as if supported by demand curves derived from fundamental preferences" (Ariely et al., 2003. p.74).

In wine tasting experiments, Plassmann et al. (2007) found that actual (neurological) experience of the good (wine) measured under functional magnetic resonance imaging (fMRI) was primed by price anchors. Plassmann et al. (2007) gave the same wine to different groups, manipulating only the price tag across the groups. They found that reported experience/pleasantness *and* activity in the medial orbitofrontal cortex (mOFC) (a key area of the brain associated with experienced pleasantness) both increased with price (although the wine was identical).

WTP values have also been shown to be affected by the pleasantness of the room, smells and moods of the respondent (Poundstone, 2010). Sadness leads to higher WTP values because the emotion signifies that things are not great and we need to change our circumstances, whereas disgust leads to lower WTA amounts as it tells us to get rid of current possessions (Hastie and Dawes, 2010; Bleichrodt, 1997).

The anchoring effect leads to a number of well-documented problems in SP valuation. Firstly, estimates derived through the bidding game format have been found to be subject to starting point effects: The higher the opening offer is, the larger the valuation estimates are. And second, estimates found under the payment card elicitation format have been found to be sensitive to range effects: A presented range of £0-100, for example, would attract higher valuation estimates than a range of £0-50.

Duborg et al. (1997), for example, report results from an SP study for the UK Department of Transport that looked at the value people attach to reductions in the risk of road injuries. In an elicitation format similar to the bidding game, they found that a £75 starting point resulted in mean WTP estimates around 1.89 to 2.87 times as large as those elicited with a £25 starting point. They also employed a payment card elicitation format; using a range from £0 to £500 for one sample and from £0 to £1,500 for another. They found that the latter range generated higher WTP estimates.

4. Contextual factors can also help to explain the odd findings that have emerged under the broad title of *embedding effects* in the SP literature. There are three types of embedding effect:

i) Insensitivity to scope

This refers to when the estimated WTP for a non-market good is insensitive to the size of that good. For instance, Desvousges et al. (1992) found no significant difference in the mean levels of WTP to save 2,000, 20,000 or 200,000 migrating birds from death. Scope insensitivity has been discovered in a number of other applications. Schulze et al. (2018) discover little difference in the estimated WTP for a partial or complete clean-up of a contaminated area; McFadden and Leonard (1993) find that residents in four western states are willing to pay only 28% more to protect 57 wilderness areas in those states compared to the protection of a single area; Jones-Lee et al. (1995) find that reducing the number of non-fatal road injuries by a factor of three only increases the stated WTP for a programme by 29%; and Hutchinson et al. (1995) find insensitivity to WTP for increases in life expectancy in normal health for the respondent and all members of their immediate household. The mean WTP for an extra 6 months was just over 30% higher than an extra 1 month.

Ariely et al. (2003) claim that scope insensitivity is further evidence of coherent arbitrariness because insensitivity to scope is most dramatic in studies that use between-subject designs. Within-subject design studies produce valuations that are far more responsive to scale. Kahneman and Knetsch (1992) argue that insensitivity to scope is explained by respondents putting forward their WTP for the moral satisfaction of contributing to public goods, rather than their true valuation of the good. Another explanation (Kahneman et al., 1999) is that insensitivity to scope reflects respondents expressing an affective valuation of a prototypical exemplar. Here, affective valuation refers to assessments of preference on the basis of "the sign and intensity of the emotional response to objects". (Kahneman et al., 1999. p.204). In the study by Desvousges et al. (1992) cited above, for example, under this psychological hypothesis respondents would have formed a "mental representation of a prototypical incident, perhaps an image of an exhausted bird, its feathers soaked in black oil, unable to escape" (Kahneman et al., 1999. p.213) and they would have then responded on the basis of their affective valuation of this image.

ii) Sub-additivity effects

These effects occur when the estimated WTP for one good plus the estimated WTP for another good is greater than the estimated willingness-to-pay when respondents are asked to value both goods together (Kahneman and Knetsch, 1992).

iii) Sequencing effects

These effects have been found when more than one good has been valued in a survey and the estimated value of a good differs according to where in the sequence it is presented to the respondent. The estimated WTP for a good has been found to fall the later in the sequence that it is presented (Samples and Hollyer, 1990).

It has been argued that insensitivity to scope findings are idiosyncratic and/or that the studies that have obtained such results are flawed in terms of survey design (Whittington et al., 1992; Carson et al., 2001). For example, the finding of insensitivity to scope should not be surprising if the description presented is not adequate to enable the respondent to distinguish between the smaller and larger good or if the survey emphasises they symbolic nature of providing the good. Another potential explanation is that individuals are running up against a budget constraint, so that they value the larger good more but they are unable to pay required multiple. However, Loomes (2006) notes that contingent valuation studies formed with WTA questions have also found insensitivity to scope.

Sequencing effects and sub-additivity effects have also been argued to be explainable with reference to income effects (Hoehn and Randall, 1989; Carson et al., 2001). Intuitively, each new good obtained reduces the income available for respondents to spend on other goods. Given this, the later in the overall package that a good is offered, the less people can spend on it.

This all further supports the notion that context matters for preferences and in preference valuation studies like contingent valuation. Context is likely to matter because the task of stating a WTP or WTA value in SP is a cognitively demanding one. Recall that in SP we are reliant on the survey respondent calculating the point at which he is indifferent between the money amount and the good. This will provide us the exact amount of money that produces the same effect on his welfare as the good in

question. In a seminal piece of work Mosteller & Nogee (1951) offered bets with different probabilities of winning. Participants had to accept or reject a simple binary gamble with a probability of 2/3 to loose 5 cents and a probability of 1/3 to win a particular amount. The winning amount varied between 5 cents and 16 cents. As the expected value of the win increased more people opted in to the bet. The study tested many different psychological phenomena, but the interesting finding for valuation was that where the expected losses and gains were close or nearly equal participants took a lot longer to make their minds up as to whether they were going to take the bet found. In other words, when we approach a point of indifference people take longer to decide implying that the cognitive burden increases. This suggests that if people are answering SP survey questions like we want them to (in accordance with economic theory), then this is a very cognitively demanding task - it is likely to be even more demanding than the experiment in Mosteller and Nogee (1951) because we are often dealing with non-market goods that people know very little about in SP. In support of this theme, (Whittington et al., 1992) undertook a contingent valuation study in Nigeria asking about WTP for communal water supply in rural areas. They found that that giving participants time to think in SP studies (a full day instead of a few minutes) significantly impacts on WTP figures because it allows people to think over the issue in more detail.

A well-known conclusion from the psychological literature is that people tend to use heurisitcs or shortcuts to aide the decision-making process when faced with novel and complex problems and data. Stating or deriving WTP/WTA figures is complex and demanding and it can be argued that people use contextual cues and primes in heuristic decision-making processes to respond in SP studies (and to some extent in their market decisions).

5. Related to the context-sensitivity critiques, Sagoff (2003) claims something even more fundamental - that it is not possible to 'observe' preference. In welfare economics preference does not "cause or pre-exist choice; rather it is derived and inferred from it". The key point is that preference is seen as a "theoretical construct" that is inferred from the selections and choices individuals make among a set of alternatives (Sagoff, 2003. p.591), but the alternatives or "choice sets" available to the individual at the time of choice are *unknown* to the economist (Sagoff, 2003. p.594).

Sagoff (2003) describes a number of examples to put over the point, the most vivid of which is his description of his recent purchase of eight boxes of cookies from a girl scout for $20. An economics perspective would take this to signify that Sagoff prefers (and values) a box of cookies at least as much as $2.50. It turned out, however, that Sagoff actually dislikes cookies and gave them all to his office colleagues; the $20 had nothing to do with his preferences for cookies. Sagoff claims that the $20 may instead have signified the value he holds towards goodwill in the local community or towards supporting the scouts. Thus, from this behaviour we cannot decipher whether the choice for Sagoff was between (i) the cookies versus $20; (ii) community goodwill versus $20; (iii) supporting the girl scouts versus $20; or (iv) a mixture of all of the above versus $20. The value is dependent on the alternatives present to the individual at the time of the decision and these alternatives are not observable to the analyst. A choice, therefore, can reveal the preference and value of anything the analyst stipulates it to be and Sagoff concludes that, therefore, preference cannot provide the basis for CBA.

This context-sensitivity of preference in relation to the alternatives and choice sets on offer leads Koszegi and Rabin (2008) to conclude that welfare analysis requires additional data in the form of happiness or subjective wellbeing metrics.

There are also some issues outside of context-sensitivity. For example, although preferences are generally left as they are and are not laundered in any way in RP and SP studies (i.e. there is no requirement on idealised preferences), we expect people to have a sufficient level of information about the good, but this is not always the case (Frey and Stutzer, 2005; Stutzer and Frey, 2004a; Stutzer and Frey, 2004b; Robinson and Hammitt, 2011). Also, specifically for SP, people may not fully understand the details of the payment system (Braga and Starmer, 2005). This all means that respondents can be manipulated by the information provided during the SP survey. The bias generated by non-neutrality in presentation is termed *information bias*.

Second, face-to-face or telephone surveys also create the potential for *interviewer bias* if respondents deviate from their true preferences under influence exerted by the interviewer. Information problems will lead people's preferences and their welfare impacts to misalign.

And finally on the issue of preference and welfare misalignment, Sagoff (2000) makes the case that preference cannot be used for policy analysis and decisions because "if the preference – or the associated WTP – has no clear relation to wellbeing…, society has no prima facie reason to seek to satisfy it" (p.1428). He goes on to conclude in emphatic terms stating that "that people are willing to pay more for one outcome than another… tells us nothing beyond that fact – nothing further and therefore nothing whatever about the relative value of that outcome". "Since WTP correlates with no independently defined conception of the good (such as happiness), what is that point of measuring it?" (Sagoff, 2000. p.1430). In a similar manner to Koszegi and Rabin (2008), Sagoff (2000) calls for welfare analysis in the form of WTP to take account of welfare more explicitly but in a novel way; Sagoff recommends that instead of asking people their WTP for some non-market good, we should ask them to state their WTP for the welfare change associated with the provision of the non-market good.

With this recommendation in mind in Dolan and Fujiwara (2012) we surveyed 1,001 adults who had recently completed an adult learning or training course. In the survey we asked how much people would be willing to pay per month (for one year) for a course or training that gave them a range of different benefits. In total we asked about 10 different benefits, such as a course that "*Led to a certificate or qualification"* or that "*Enabled you to improve your knowledge or skills"* and so on. We also asked about the WTP for a course that "*Improved your happiness on a day-to-day basis"* and for one that "*Made you more satisfied with your life overall".* In terms of outcomes, life satisfaction ranked second out of 10 and happiness was sixth (*"Progress in work or a career"* was first), meaning that using wellbeing explicitly in the WTP question had the tendency to reduce WTP values for training courses, which is good news in light of the issues of hypothetical bias discussed below. Importantly, the number of non-responses ('don't knows') and zero (£0) WTP values – two key tests of the validity of a contingent valuation study – were not any higher for the life satisfaction and happiness WTP questions than the other 8 outcomes, which suggests that Sagoff's recommendations can be applied in practice. For Sagoff this would ensure that preference and WTP have a meaningful connection to wellbeing. This early paper acted as a first bridge for me between preference satisfaction and

wellbeing valuation and demonstrated that wellbeing data and concepts can be applied to valuation.

The findings and recommendations stemming from this literature set out some rationale for the greater use of SWB data in policy analysis to solve for some of the main substantive problems related to the preference satisfaction account of welfare.

2.4.2. Technical critiques of the preference satisfaction account for valuation

This section focuses on specific problems that have been found in the RP and SP valuation literature and therefore focuses on a more narrow set of issues than those covered in the previous section which looked at the critiques of the preference satisfaction account of welfare per se.

1. Choi et al. (2011) use an experimental setting to test the extent to which individual choice behaviour complies with economic preference rationality assumptions. Here choice behaviour complies with rationality assumptions if there exists a well-defined utility function that choices maximise. Subjects were presented with a binary choice under risk with varying levels of financial payouts. The authors found that less than half of the sample of people exhibited choices that came close to satisfying rationality assumptions. There were some respondents who have "very high error propensities" (Choi et al., 2011. p.27). High-income and high-education subjects displayed greater levels of consistency and younger subjects were better utility maximisers in their choices. They conclude that the heterogeneity in utility-maximising performance suggests that "there are circumstances when revealed preferences may not be 'true' preferences. If so, then positive predictions and welfare conclusions based on revealed preferences may be misleading" (p.8). I would also add these problems also apply to stated preference methods.

2. RP approaches are limited by the number of proxy markets available that can reveal something meaningful about the value of a non-market good and by the fact that the proxy market in question needs to be functioning well. For example, the values that we place on clean air, improved mental health and protection of endangered species

may not show up in any of our market transactions. And even if they may in theory show up (e.g. cleaner air could contribute to house prices), they may not in practice if we are unaware of the non-market good when making market decisions. CBA tends to take actual preferences (rather than some account of idealised preference), and so well-informed consumers in related markets is one implicit pre-requisite for RP valuation. The market itself also needs to be sensitive to changes or levels in the non-market good. For example, valuation of environmental amenities like clean air or noise is not possible where there are state-related interventions in the market such as caps on house rental prices. With rental caps houses in clean and quite areas may be restricted from increasing in price.

SP methods get around these problems by creating a hypothetical market, with full information about the good. This allows us, in theory, to estimate values for any type of non-market good, but with the downside that in SP we are working in hypothetical market scenarios with reduced incentive compatibility, which can create problems of its own.

3. CS and ES are estimates of Hicksian surplus. Hicksian surplus is derived from the substitution effect and is the theoretically appropriate measure because it captures the monetary amount required to hold each individual's utility constant. While some applications have made attempts to recover compensated measures, RP methods like travel cost and hedonic market methods typically estimate and report changes in Marshallian surplus, which differs from Hicksian surplus in that it picks up the income effect as well. Although, in practice, income effects are likely to be small in non-market valuation settings this is still an important point to bear in mind about RP methods.

4. There may be cases when individuals hold lexical preference orderings for non-market goods, whereby a reduction in the non-market good cannot be compensated for by an increase in income and consumption of market goods. This represents a violation of the substitutability assumption for valuation and in such cases no finite WTA (or WTP) amount exists for the non-market good (Adler, 2012; Adler and Posner, 2008).

5. Specifically to SP methods, asking survey respondents for a WTP or WTA figure can change the subsequent perceptions and values associated with the non-market good (Sandel, 2003). This means that if we ask people their WTP (WTA) and subsequently undertake the policy intervention, what they actually experience could be different to what we would expect based on the contingent valuation survey results or we may not be able to value a non-market good with contingent valuation in an (economic) theory-consistent way. One interesting example of where this happens is in 'NIMBY' (Not In My Back Yard) policy interventions, such as nuclear waste disposal, power plants, airports, prisons and so on. In the US states that have used compensation schemes for NIMBY sitings have not experienced much success (Arrow et al., 1993; Frey et al., 1996). Citizens recognise that, although there are negative externalities involved for the local community, these projects are socially desirable and Frey et al. (1996) find that in the case of a proposed new nuclear waste repository in the small village of Wolfenschiessen (Switzerland) a slight majority of the villagers (50.8%) voted in favour of the project. However, when a monetary compensation package was offered to the villagers support for the project dropped by more than a half. The negative effects of the compensation package come about because people feel that the compensation acts as a bribe (the bribe effect) and because monetary compensation deprives people of satisfying pro-social feelings and behaviours (crowding out of public spirit effect) Frey et al. (1996). The compensation on offer is essentially a WTA amount and the latter effect is a case of where monetary valuation (in the form of WTA) changes the perceptions of the good. In the case of NIMBY projects a contingent valuation survey would not be able to find a finite WTA figure for such projects, although people may often be willing to accommodate them - as Frey et al. (1996) find people show support for these projects in referendum-type voting and in-kind compensation schemes have been found to be more effective in garnering support for NIMBYs than monetary compensation. It is, therefore, the monetisation that is problematic and this has implications for stated preference methods.

Also, relatedly, economic theory relies heavily on comparative statics in partial equilibrium, holding preferences constant (Bruni and Sugden, 2007). This is certainly the norm for valuation in CBA (OECD, 2018). But preferences - and hence value - may change *as a result of* provision or experience of the non-market good, and this

would only be acknowledged properly if people were able to forecast this preference-changing outcome in advance when stating or revealing a preference. This is very similar to the concept of adaptation, but there are some differences. For example, those who adapt to severe obesity or paraplegia may do so by changing their daily activities to adapt whilst their actual preferences have not changed in any significant way - eg, they still have a preference for and would like to be able to play tennis and basketball etc .

6. People find it difficult to convert a feeling or concept of value on to a monetary scale (Loewenstein and Schkade, 2003). In this interpretation people could have strong and well-defined references, beliefs and feelings for many of the things that are not sold through markets, but these beliefs are not represented monetarily (Gregory et al., 1993). Amir et al. (2008) find a disparity between people's WTP and their predicted experience or utility of goods like music concerts. Kahneman et al. (1998) found that in a juror award experiment in which people studied a number of corporate malpractice cases and were asked to rate the defendant's (the corporation) actions on a scales of 'outrage' and 'degree of punishment justified', there were strong correlations between the level of outrage and punishment across the different jurors, but the dollar awards had very little correlation. This is supported by a study by Malouff and Schutte (1989) who find that juror awards are highly susceptible to the anchoring effect of the plaintiff's initial level of compensatory demand (presumably because the jurors had no concrete idea of what the dollar compensation amount should be).

7. There exists a set of survey-related biases inherent to SP methods. The embedding effects and interviewer and information bias discussed above are themselves survey-related biases. In addition, SP surveys may suffer from,

(i) Hypothetical bias

The hypothetical nature of the good in question and the payment mechanism can lead to inflated values in surveys and it is widely believed that individuals overstate their valuation of a good by a factor of two to three when comparing hypothetical versus actual payments for goods (Murphy et al., 2005). One reason is attributed to non-commitment bias; respondents may overstate their true WTP because they do not face

a budget constraint and do not consider substitute goods within the world of the hypothetical scenario.

There is some evidence that the magnitude of hypothetical bias is greater for public goods than for private goods (Murphy et al., 2005).

(ii) Strategic bias

Respondents in stated preference surveys may have an incentive to deliberately misrepresent their true preferences in order to achieve a more desirable outcome for themselves by influencing policy. Individuals may overstate their valuations of the good if they believe their responses influence its provision and are un-related to the price they will be charged for it. Conversely, individuals may understate if they believe that their response will not influence their desired outcome but will influence the price they are charged for it (Carson et al., 2001). Carson argues that true preferences are revealed when respondents believe that the non-market good's provision is contingent on their stated values and when they believe that they will have to pay the amount they state, but this is virtually never achievable in SP studies.

(iii) Protest values

Respondents with a positive true WTP may put forward a zero stated valuation due to, for example, ethical objections to the idea of paying for the good under consideration or to the idea of government intervening in the issue at hand. If such respondents are not identified through follow up questions, and their responses consequently excluded from the statistical analysis, then biased estimates of the value of the good will result.

(iv) Non-response bias

This will occur if individuals who feel strongly for or strongly against a good or issue are more likely to respond, which can lead to either an upward or downward bias. There is also the potential for fatigue and frustration to set in, especially in iterative bidding formats. In this situation respondents make end up making little effort to provide accurate replies of their WTP/WTA (Accent, 2010).

(v) The WTP-WTA disparity

All stated preference survey choices and questions can be presented in terms of WTP (to receive a good or prevent a loss) or in terms of WTA (to lose a good or incur a loss). In theory, WTA for most goods evaluated under Stated Preferences should exceed WTP by a few percentage points due to the fact that WTP is constrained by an individual's income (Sugden, 2005). Numerous papers have found, however, that stated WTP is often far below stated WTA for the same good (Hanley and Shogren, 2005) and the WTP-WTA disparity has become one of the most infamous survey-related biases examined in stated preference research.

Sugden (2005) argues that the most credible explanations for this relate to the psychological arguments concerning loss aversion and its derivative; the endowment effect (Kahneman and Tversky, 1979; Loewenstein and Adler, 1995). Some authors argue that the appropriate formation depends on property rights (Carson et al., 2001), others have argued that the WTP formulation should always be used (Arrow et al., 1993). One reason for this is that CV studies adopting a WTA formulation have often been unsuccessful due to an inability to convince respondents that they have the right to sell a nonmarket good (Carson et al., 2001).

The WTP-WTA disparity may also be, to some extent, a product of informational constraints and inexperience. Bateman et al.'s (2009) virtual reality survey tool (that allows survey respondents to experience environmental policy changes in a 3D environment) reduced the difference between WTP and WTA for environmental goods and List (2003) finds that experienced traders (in a number of different real markets) do not exhibit the endowment effect.

8. It is also interesting to assess what conclusions we can draw from the growing neuroscientific literature on this topic. A highly-cited example is Berridge (1996) (see Berridge and Kringelbach, 2011) who found that wanting and liking or experiencing arise in two different neurological areas or systems. Therefore, “wanting things may not be an accurate predictor of whether those things will increase subjective wellbeing” (Diener and Suh, 1997. p.190). Glimcher’s (2010) seminal book provides a comprehensive review of the neuroscientific literature, looking at decision-making and valuation, two areas that are central to CBA. Although Glimcher (2010) is tentative in that he thinks it is premature to use neural measurements explicitly for

welfare analysis, there are a number of important lessons from the fields of neuroscience and neuroeconomics.

Glimcher (2010. p.396) states that the choices people make "occur because our brains explicitly represent the economic concept of preferences in the form of cardinalised expected subjective values". The transitivity axiom requires that these subjective values are stored in *absolute* terms somewhere in the brain (Glimcher, 2010). His example discusses training somebody to make utility-maximising choices between four options: A = $1,000,000; B = $100,000; C = $1,000; and D = $100. When presented with A vs B we would need to train the individual (i.e. programme the brain) to choose A over B (A > B). And presented with C vs D we would need to train the individual to choose C over D (C > D). Now when faced with B vs C, the transitivity axioms requires that B > C, but if the chooser only represents *relative* values in the brain he will choose C over B, because C has "higher learned relative subjective value" (p.235). In order for the chooser to form transitive choices and choose B over C then he must hold absolute subjective values for these outcomes.

The problem comes in the fact that subjective values seem to actually be stored in *relative* terms in the brain. "Everything we know about the brain tells us that the value of options are encoded in a reference-dependent way.... Cortical areas do not represent the absolute values of anything". "This constraint on how our brains represent subjective values has profound implications for.... welfare economics" (Glimcher, 2010. p.417). This process has come to be known as *Heeger normalisation* (Heeger, 1992), whereby the firing rates of relevant neurotransmitters (mainly dopamine neurons) are converted from an absolute to relative magnitude.

Dorris and Glimcher (2004) found that when the values of different choices in a choice set all doubled firing rates of dopamine neurons remained roughly constant. Neurons have a firing range of about 100 Hz, with a baseline of 10 Hz. This gives us only about 90 Hz of range for signalling in decision making tasks. If a candy bar represents 11 Hz and a new computer 60 Hz, then there is not a lot of leeway left for goods and outcomes preferable to a laptop computer. Thus through the normalisation process the brain ensures the same gap in firing rates for any binary decision - around a 90 Hz difference for the choice between candy and the laptop and equally around a

90 Hz difference for the choice between a laptop and a five-bedroom Manhatten penthouse. This normalisation "maximises the discriminability of the two options in the choice set given the existing cortical variance" (Glimcher, 2010. p.244). Without this normalisation process under the restricted cortical variance the brain architecture would make a huge number of errors in choice tasks. These values are generated in the fronto-parietal areas.

Now, of course, there must be some more absolute level of value stored in the brain as in reality we are fully capable of choosing properly between different amounts of money etc, over which we have not made choices before. Glimcher (2010) states that the medial prefrontal cortex and the striatum are central areas that act as the "physical seat of valuation" (p.347), funnelling all of the subjective values that guide choice. This area "supports comparisons of all of the objects we have ever encountered, and so must store the values of all of these objects within a single common framework" (p.347). However, “the shifting baselines, or reference points, of all sensory encoding systems require that vertebrates produce some degree of irrationality in their choices – some violations of axioms such as transitivity are unavoidable…” (Glimcher, 2010. p.346).

The upshot of what we currently know about the neural basis of choice implies that the axioms that underlie preference for economic analysis are not something that aligns with human nature and our biological make-up. We are back to Pareto's *integrability problem* - transitivity to some extent *really* is mere speculation.

2.5. Summary

Hicks’ theory of valuation is the agreed approach to valuation in economics and policy evaluation. Under this theory the value of a non-market good or service relates to the impact that it has on the individual’s wellbeing. Monetary value is expressed as the amount of money which would have the equivalent impact on wellbeing. Valuation theory is agnostic about the measure of wellbeing to be used in the analysis.

For the purposes of valuation economics has traditionally used the preference satisfaction account of wellbeing. This has led to the development of the revealed preference and stated preference methods for valuation. Stated preference methods are more direct in their approach in that they more closely replicate the original theory of valuation.

There is a longstanding literature that criticises preference-based valuation methods. There are critiques of preference as a measure of welfare as well as technical problems with the valuation methods themselves, which I have discussed at some length. These critiques and problems demonstrate serious flaws in RP and SP methods, the dominant methods for valuation in economics and CBA and two key approaches in the Green Book. As such these critiques form an important part of the rationale and motivation for the wellbeing valuation approach in the next chapter because as we shall see since wellbeing valuation does not use preference data it provides the potential to avoid many of these issues in valuation.

Chapter 3

3. Wellbeing valuation

3.1. Introduction

In this chapter I develop from the arguments discussed in Chapter 2 to set out the rationale for wellbeing valuation which provides the background and context for the derivation of the new approach in Chapter 4. This chapter covers and summarises a substantial amount of literature, but also makes a number of original contributions.

I start by setting out the context and background of the use of wellbeing data in research. This will show the growing interest in wellbeing analysis in economics as well as in policy analysis. I then provide a short introduction to the wellbeing valuation method, which has come out of this increased interest in wellbeing data in economics. The discussion of the wellbeing valuation method at this stage is brief because I provide a full in-depth discussion and assessment of the methodology in Chapter 4. The description of the wellbeing valuation method in this chapter is intended to provide sufficient information for the discussion of where the method sits alongside the preference methods discussed in the previous chapter and of the main rationale and reasons for using wellbeing valuation, the core aim of this chapter.

In terms of original contributions, firstly, I pull together all of the previous discussion in this area into a succinct and complete assessment of WV. Previous studies have tended to discuss a small handful of disparate issues when making the case for WV. I formally categorise and aggregate this literature into a set of key issues and pros and cons and add some further thoughts and arguments for WV that have not been covered before. Secondly, I merge previous and separate work by Adler and Sugden to form a new framework or structure for thinking about WV and its relationship to traditional preference-based valuation methods. This is an important issue because it will determine the extent to which WV and preference-based methods are

complements or alternatives in terms of valuation methods. The literature to date has not focussed on this key question in any detail. The joining up and synthesising of these various areas of work and theories has, to the best of my knowledge, not been done before.

This chapter will provide the most comprehensive rationale and support for the WV method to date.

3.2. Subjective wellbeing measures in economics and policy evaluation

The definition and measurement of human welfare has a long academic history going back to the ancient Greeks and other thinkers such as Confucius (Bok, 2010). Welfare and happiness were the central themes in the writings of Socrates, Aristotle, Epicurus and other early Greek philosophers and much of their theories and viewpoints have shaped how we think about welfare today and some argue that we have not developed or added that much more in addition to what the ancient Greeks had said (Diener et al., 1999).

Recognition of the role of welfare surged with the work of the classical utilitarian philosophers such as Bentham, Mill and Sidgwick. These early utilitarian thinkers had a profound influence on economics. Jeremy Bentham first defined *utility* in hedonic terms, measured as the balance of the amount of pleasure versus pain experienced by an individual. Under utilitarianism this concept, which Bentham often called happiness, was the ultimate intrinsic good and hence consequently Bentham claims that “The greatest happiness for the greatest number is the foundation of morals and legislation.” (Bentham, 1983). Similarly, experience and sensation played a paramount role in the theories and work of the economists of this generation (e.g., Jevons, Edgeworth and Pantaleoni). To them psychological phenomena such as sensation, pleasure and pain were "an essential part of economics" (Bruni and Sugden, 2007. p.154). Classical utilitarianism and early economic theory was thus based on hedonic or experienced concepts of wellbeing and utilitarianism (albeit under different guises) has become the basic moral or normative tenet in modern economics.

Bentham's moral philosophy rests on two criteria/assumptions: (i) that pleasure and pain are quantifiable and (ii) the quantities can be aggregated across individuals (Read, 2007). The main stumbling block back then was the degree to which it was possible to measure pleasure and pain and it was clear that utilitarian philosophers and economists looked forward to the day when hedonic states could be measured directly using, for example, a *hedonimeter* as proposed by Edgeworth (McPherson & Hausman, 2006). This type of technology never materialised and so whilst staying true to the utilitarian framework, economists abandoned Bentham's hedonic measure of utility in favour of a preference satisfaction account of welfare in the early twentieth century (Bruni and Sugden, 2007). No clearer can this been seen than in the work of Edgeworth, a devoted Benthamite who became the "pathfinder of ordinalism" (Read, 2004. p.5). Under the ordinal utility approach, as Paul Samuelson showed (as discussed above), it is possible to map people's choices over different bundles of goods on to a complete map of utility.

This move represents a normative transition in economics between different definitions of welfare that themselves have a long tradition in philosophical thinking. Although many permutations exist, we can think of three broad accounts of human welfare (Parfit, 1984):

4. Mental states and the self-reported experience of the individual.
5. Preference satisfaction.
6. Objective lists encompassing normative ideals.

Mental state accounts of welfare are based on people's self-reports about how their lives are going. I shall use the terms subjective wellbeing (SWB) and mental states interchangeably here. Broadly speaking, mental states can be,

- Evaluative SWB, which are global assessments of people's wellbeing such as life satisfaction.
- Experience SWB, which are measures of people's feelings or *affect* over a period of time. This could be measures of happiness, worry, anxiety, sadness,

fatigue, vitality and so on. This account of welfare is what Bentham originally held human welfare to consist of.

- Eudemonic SWB, which relates to people's psychological needs, such as autonomy and the feeling of things being worthwhile, which could contribute to welfare independently of any pleasure they bring (Hurka, 1993).

The preference satisfaction account has been described in detail in Chapter 2. It equates the degree to which people's preferences are satisfied with their level of welfare and is the measure used traditionally by economists since the early part of the twentieth century. As we have seen this account is dependent on a number of assumptions regarding the validity of preferences.

Finally, the objective list account is based on normative assumptions about basic human needs and rights (Dolan et al., 2011). In objective lists "certain things are good or bad for us even if we would not want to have the good things or avoid the bad things" (Parfit, 1984. p.499-502). This may be measured, for example, by the literacy and morbidity rates in a country and the items on the list often cover the items set out in Nussbaum's *Capabilities* approach (Bok (2010) and Dolan and White (2006) argue that Nussbaum's capabilities are objective lists).

The first two categories are subjective in that they allow the individual in question to determine or reveal what is important for his welfare, whereas the objective list account represents a list of factors that are determined externally to someone's self-reported wellbeing or their preferences and wants although of course many of the items on any list are important determinants of subjective measures of welfare. (Veenhoven, 2010).

For economists there are a number of big draws associated with subjective accounts of welfare. Firstly, subjective measures placate the profession's general (self-claimed) reluctance to make strong normative statements (Bruni and Sugden, 2007). Subjective accounts (SWB and preferences) privilege the individual "as the only one qualified to assess his or her own wellbeing" (MacKerron, 2011. p.3) which is compatible with liberal political views (MacKerron, 2011; Dolan et al., 2011) and helps economists (to some extent) stay clear of making normative claims.

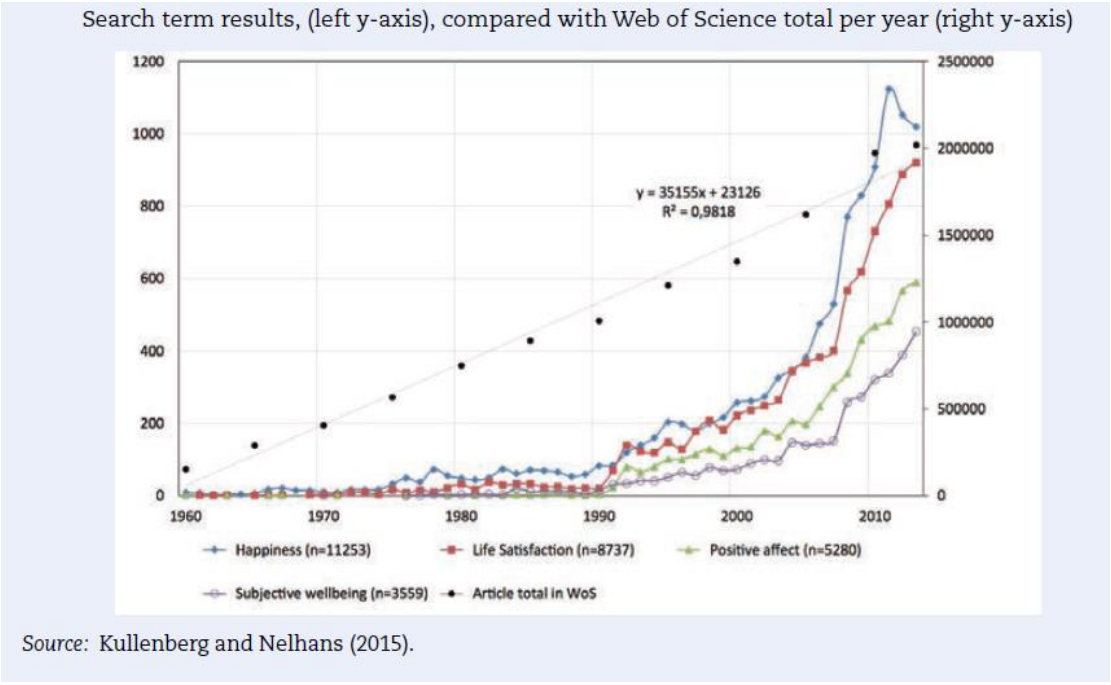
Secondly, and more importantly, subjective welfare measures provide the facility to derive trade-offs between different goods, services and life events to help direct policy. Policy evaluation needs to draw conclusions on what actions are right to take in society's interest and in a world with scarce resources this invariably means making trade-offs between different objectives. With subjective welfare measures we can be led by the individuals/citizens themselves; their preferences or the differential impacts on SWB will tell us what is worth doing and what is not. In objective list accounts of welfare, it is a hard enough task determining the items to include in the list, let alone thinking about how we can weight these items against each other (Diener and Suh, 1997). Invariably weighting systems for objective list accounts fall back to relying on subjective measures to provide clues as to which goods are more important and in doing so raises the question of why not just use subjective measures from the outset? For these reasons it is impossible to derive monetary values for different goods based on an objective list account of welfare as we lack robust objective techniques for weighting the worth of money in relation to other goods in order to derive theoretically consistent measures of value (compensating and equivalent surplus). We shall, therefore, leave the objective list account of welfare here and focus on mental states and preference satisfaction accounts from here on.

The two subjective accounts of welfare have dominated discourse in economics over the past two centuries. The transition discussed above was from a reliance on mental state accounts of welfare to the application and endorsement of the preference account. This move to ordinal measures of utility in the twentieth century was by no means at the time a mere stop-gap on the road to eventual cardinality (once - we might suppose - that scientific methods of welfare measurement had been developed to a sufficient degree) (Read, 2004).

This thesis sits at an interesting time when mental state measures are making a comeback in economic theory and applied economics. As Kahneman et al. (1997) phrases it the economics discipline is going "back to Bentham", referring to efforts in economics to revert back to mental state measures of welfare as first proposed by the early utilitarians.

The problems associated with preference satisfaction as a measure of welfare have encouraged an increasing number of economists to revert to measures of SWB in economic and policy analysis. Figure 1 shows the exponential rise in the number of SWB-related publications in economics journals.

Figure 1. The increase in wellbeing research and publications (1960 – 2015)



Source: OECD (2018).

Economists use SWB data mainly in four different ways (Diener et al., 2009; Dolan et al., 2011):

- **Economic analysis**. This research looks at the relationships between economic phenomena and SWB. For example, Easterlin (1974) and (1995) has looked at the impact of income on life satisfaction. Winkelman and Winkelman and (1998) look at unemployment and life satisfaction and Blanchflower and Oswald (2004) look at the impacts of aggregate level macroeconomic variables such as inflation and unemployment rates on SWB. A large proportion of this work also tests standard economic assumptions with SWB data.

- **Non-economic analysis.** These data have also been used to look at phenomena that are outside of the traditional economist's viewpoint. A prominent example is Metcalfe et al.'s (2011) study of the impact of the 9/11 terrorist attacks on SWB. Also SWB analysis can inform legal compensation decisions (Oswald and Powdthavee, 2008).

- **Policy evaluation.** SWB data can also be used to look at the impact on wellbeing of certain policies. For example, the Department for Work and Pensions tracked wellbeing measures (life satisfaction) as one of the outcomes of the Employment Retention and Advancement (ERA) scheme, which added a new unique combination of services to help unemployed individuals who have entered work as well as low-paid workers remain and progress in work. The impact of the ERA was assessed using a large-scale randomised trial and wellbeing outcomes were taken along with traditional labour market metrics such as wages and employment rates. Governments and other organisations may use SWB data to identify the main determinants of quality of life, such as the best places to live and work for SWB, or they may use SWB data to identify those groups who are worst-off and most disadvantaged (in terms of wellbeing) in order to determine where resources should be directed.

- **Valuation of non-market goods.** As I will discuss in the rest of this thesis SWB data can be used to derive estimates of compensating and equivalent surplus for non-market goods using the wellbeing valuation approach. These values can feed into CBA. Also, governments may use this type of analysis to determine levels of compensation to pay citizens who are adversely affected by a policy intervention.

Generally speaking, these analyses work off the large amount of data on SWB and life events that are collected by universities and national statistical offices. They are arguably all very fruitful and important areas of research and can unlock analytical mysteries that cannot be tackled with standard preference-based economic welfare analysis. The focus of this thesis is on techniques for the valuation of non-market goods using SWB data.

The development of these research areas has come hand-in-hand with an interest from policy-making institutions. The United Nations guidelines on national accounts now state that GDP should not be used to stand for wellbeing (Duncan, 2010) and both the Office for National Statistics (ONS) in the UK and the OECD have prominent wellbeing programmes in place covering both data collection and analysis.

3.3. The Wellbeing Valuation Approach

3.3.1. Background

A new method for valuing non-market goods and services, the wellbeing valuation method, has arisen out of the growth in interest in wellbeing data and analysis in economics and policy analysis. WV has been one of the main uses of SWB data in the economics literature and out of the four uses of SWB data described above it has arguably been the focus of the most amount of debate and research.

The first paper on WV was published in 2002 in *Health Economics* by Ferrer-i-Carbonell and Van Praag (2002) who looked at the valuation of various illnesses. The WV method has since then been used to value a range of non-market goods, 'bads' and outcomes, including,

- **Sports participation and the Olympics** (Fujiwara et al., 2014; Dolan et al., 2019);
- **Environment and environmental amenities, including air quality** (Welsch, 2002; Welsch, 2006; Welsch, 2007; Welsch and Kuhling, 2009; Rehdanz and Maddison, 2005; Carroll et al., 2009; Ferreira and Moro, 2009; Ambrey and Fleming, 2011; Tsurumi and Managi, 2015; Mendoza et al., 2019; Barrington-Leigh and Behzadnejad, 2017; Giovanis and Ozdamar, 2016; Krekel and Zerrahn, 2017);
- **Weather** (Fedderson et al., 2012);
- **Nuclear disasters and natural disasters** (Sarrias and Jara, 2019; Danzer and Danzer, 2011; Fernandez et al., 2019)

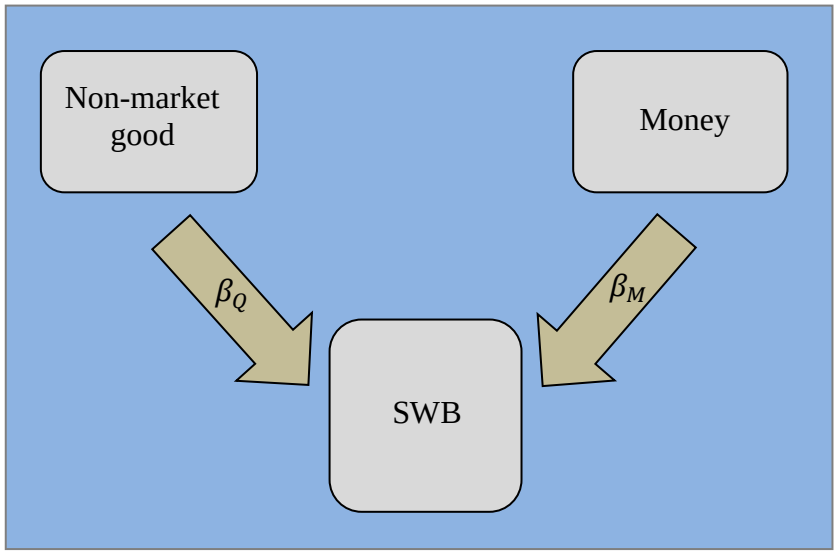
- **Health** (Groot and van den Brink, 2006; Ferrer-i-Carbonell and van Praag, 2002; Howley, 2016; Huang et al., 2018);
- **Crime** (Manning et al., 2016; Cohen, 2008; Brenig and Proeger, 2016; Moore, 2006);
- **Public sector corruption** (Welsch, 2008b);
- **Civil conflicts** (Welsch, 2008a);
- **Care-giving** (McDonald and Powdthavee, 2018; Schneider and Kleindienst, 2016; van den Berg and Ferrer-i-Carbonell, 2007);
- **Access to water** (Mahasuweerachai and Pangjai, 2019);
- **Terrorist attacks** (Frey et al., 2007);
- **Social relationships** (Powdthavee, 2008; Chandoevwit and Thampanishvong, 2015);
- **Employment and job characteristics** (Murtin et al., 2017; Clark and Oswald, 2002; Helliwell and Huang, 2005);
- **Macroeconomic events** (Blanchflower and Oswald, 2004);
- **Value of life** (Oswald and Powdthavee, 2008);
- **Commuting** (Stutzer and Frey, 2004);
- **Green space** (Aoshima et al., 2018; Tsurumi and Managi, 2015);
- **Daylight savings time transitions** (Kuehnle and Wunder, 2015);
- **Volunteering** (Becchetti et al., 2018);
- **Adult learning courses and qualifications** (Dolan and Fujiwara, 2012);
- **Income inequality** (Beja, 2011);
- **Cultural activities and events** (del Saz-Salazar et al., 2019; Fujiwara et al., 2014; Marsh and Bertranou, 2012).

In this section I set out a non-technical introduction to WV to provide the foundation for a discussion on interpretation of the method and the pros and cons of the approach. I then provide a full technical exposition of the WV approach in Chapter 4 when discussing the main technical issues of the method.

3.3.2. Introduction to the wellbeing valuation methodology

The premise of the wellbeing valuation (WV) approach is to estimate measures of welfare change (CS and ES) as set out in section 2.2. from data on people's SWB. This is depicted in Figure 2. We are interested in measuring two effects: first, the impact of the non-market good on SWB (β_Q) and second the impact of income or money on SWB (β_M).

Figure 2. The wellbeing valuation approach



Once these effects have been estimated it is possible to derive measures of welfare change by looking at the marginal rates of substitution (MRS) between the non-market good and money. This essentially measures moves around an indifference curve at a given level of SWB, where I am using the term 'indifference curve' more broadly here to mean a set of points at which an individual's level of welfare remains constant (however we may measure that level of welfare). In other words, with 'observable' welfare data – in the form of SWB data – we can measure a welfare function and the level sets of this function (provided that there are two or more arguments in the function) - which equate to the indifference curves - to see how two different goods can be traded off against each other at the margin.

A key point to note is that the WV approach, therefore, represents what Randall (1982) would call a *direct income compensation approach*. Since we can work with

an ‘observable’ measure of welfare, the WV approach actually provides a solution to McKenzie’s (1957) and Hurwicz and Uzawa’s (1971) original income compensation method, whereby one estimates the amount of the numeraire (here, money) the individual would require to have with Q^0 to achieve the same level of welfare as with Q^1 and her original level of income M^0 (McKenzie (1957) and Hurwicz and Uzawa (1971) originally worked with price changes, but following Randall (1982) I have substituted changes in a non-market good (Q) for prices here).

The solution to the original income compensation method requires observing the relevant points on indifference curves (Randall, 1982), which we can now in theory do with WV, rather than relying on market behaviour or eliciting values directly from people in stated preference. Therefore, it can be said that the WV approach represents the most direct approach to non-market valuation, which most faithfully translates the economic theory of welfare change measures into practice, with the caveat that this measure of welfare is SWB rather than preference as usually assumed in economics. One thing to note here is that WV is effectively more closely related in concept and theoretical underpinnings to stated preference than to revealed preference in terms of the directness of approach to non-market valuation, although WV is more direct than stated preference because it works with ‘observable’ measures of welfare.

In practice the elements of Figure 2 and the MRS between Q and M are estimated empirically through statistical analysis based on a model of SWB such as:

(8) $$SWB = f(M, Q, X)$$

where SWB is some measure of wellbeing, such as life satisfaction, M = income, Q = the non-market good being valued and X = other determinants of SWB.

The vast majority of WV studies to date have defined SWB as life satisfaction and employed single-equation multivariate regression models to estimate (8) - examples include, (Kountouris and Remoundou, 2011; Ambrey and Fleming, 2011; Moore, 2006; Menz and Welsch, 2012; Beja, 2011; Cohen, 2008; van den Berg and Ferrer-i-Carbonell, 2007; Clark and Oswald, 2002; Blanchflower and Oswald, 2004; Frey et

al., 2004; Stutzer and Frey, 2004b; Barrington-Leigh and Behzadnejad, 2017; Sarrias and Jara, 2019; Murtin et al., 2017; Schneider and Kleindienst, 2016; Becchetti et al., 2018). Equation (9) sets out an example of the regression models used.

(9) $$SWB_i = \alpha + \beta_1 M_i + \beta_2 Q_i + \beta_3 X_i + \varepsilon_i$$

where ε is the error term and the subscript i denotes individual i. Other statistical methods used to date include use of panel data, which adds a time subscript to equation (9) and a time-invariant term can be added in ε_i so that model (9) is estimated using fixed effects (e.g. Manning et al., 2016; Giovanis and Ozdamar, 2016; McDonald and Powdthavee, 2018).

The results from a model like equation (9) are used to estimate values as per Figure 2, where $\beta_2 = \beta_Q$ and $\beta_1 = \beta_M$. The value of Q is estimated from the MRS between Q and M. Equation (10) provides the basic format of the calculation. In practice this becomes more complex and involved as the impacts of Q and M may be estimated in a non-linear format.

(10) Value of $Q = {-(\beta_Q \cdot \Delta Q)}\Big/{\beta_M}$

Where in many cases $\Delta Q = 1$ as it's the provision of a good or service. When it is a non-market 'bad', β_Q will be negative and the result of equation (10) will be positive signifying that the individual needs to be compensated a positive financial amount.

In the WV method the value of the non-market good (Q) is therefore derived without recourse to market transaction data (as in revealed preferences) or to eliciting WTP or WTA values from survey respondents (as in stated preferences).

3.3.3. Where does wellbeing valuation sit alongside preference methods?

A key question to ask is how and where the new wellbeing valuation method fits in the current landscape of valuation methodologies made up of preference-based

methods. This question has received very little attention in the WV literature to date and the conclusions drawn tend to be fairly rudimentary.

A strong theme that runs through the current literature on SWB is the assumption that SWB data are direct measures of the economist's concept of utility (Adler, 2012) (e.g. Diener et al., 2009; Frey and Stutzer, 2009; Levinson, 2012; van den Berg and Ferrer-i-Carbonell, 2007). For example, Ferrer-i-Carbonell (2012. p.2) states that self-reported wellbeing measures serve the "purpose of better understanding individual's preferences". Diener et al., (2009. ch.2, p.4) go further to claim that "[subjective] wellbeing is essentially identical to economists' concept of utility". However, none of these types of statements have been backed up with any evidence or principled moral reasoning.

In this thesis I take a different approach that is more consistent with the thinking in philosophy and ethics, that SWB (mental state accounts) are qualitatively distinct to the preference account of welfare. This has implications for how we use and interpret values from WV. Drawing on a number of theories and propositions we can formulate a better and more comprehensive framework for thinking about WV and other valuation methods.

3.3.3.1. SWB and preferences

A useful framework for organising the different views on the relationship between SWB and preference is Adler's two defences of SWB data in policy analysis (2012) [4]:

[4] I note that there exists one further approach to interpreting SWB data in economics, which sits outside of Adler's two defences. Kimball and Willis (2006) use SWB as an argument in the standard utility function, alongside other goods and outcomes. This is what Adler (2012) calls the 'hybrid model'. This is probably the least common use of SWB data among economists and also philosophers and is problematic for valuation which requires either the PR or EQ defence. This is because in order to estimate trade-offs SWB needs to be the intrinsic outcome against which trade-offs can be made between goods and money (i.e. it needs to be the objective of the welfare function rather than an argument in it as per Kimball's approach). Under the hybrid model utility remains the intrinsic outcome and SWB only has instrumental value. Therefore, I will not explore the hybrid model any further here.

i. The **Preference-Realisation** (PR) defence of SWB adopts the view that SWB data measure the extent to which a person's preferences are being satisfied. In other words, SWB data are evidence of an individual's level of preference-realisation. As discussed above, this is the common line of thought amongst many economists.

ii. The **Experience-Quality** (EQ) defence of SWB takes the view that SWB data are evidence of an individual's mental states. This is more in line with the thinking in philosophy and amongst psychologists such as Kahneman.

Those who use SWB and preference satisfaction synonymously such as Diener (see for example Diener et al., 2009; Diener and Suh, 1997; Diener et al., 1999) are adopting the PR defence of SWB. A good example of this school of thought is Benjamin et al., (2012) where preference is 'privileged' in the assessment in that the usefulness and reliability of SWB data is questioned on the extent to which they can replicate preference data. Under the PR defence, SWB data can be used in wellbeing valuation to measure the preference-specific forms of willingness to pay (WTP) and willingness to accept (WTA). In other words, wellbeing values will represent WTP and WTA as in preference methods under this approach. For this defence to be true, it requires that SWB data evidence individuals' *ordinal* preference utility (Adler, 2012).

Many non-economists (and some economists such as Layard), however, take a broader view, which I believe better accounts for the nuanced differences between SWB and preference. For example, Adler and Posner (2008) acknowledge that wellbeing values *only* equate to WTP and WTA *if* people satisfy preferences in order to maximise SWB. Unfortunately, there is little work outside of economics on the interpretation of wellbeing values and so the direct equivalisation with WTP and WTA has seemed to stick and has become the perceived wisdom. Outside of the wellbeing valuation literature, in the broad field of normative ethics, I believe it would be impossible to find a philosopher that would take such a stance of directly equating SWB with preference and those that may would stipulate a list of important assumptions and caveats for making such a claim.

I think that Layard's (2006. p.31) conclusions, which align with the EQ defence, provide the correct way to think about this. He states that *"economics uses exactly the right framework for thinking about public policy. Policy instruments are set so as to maximise the sum of utilities,… What is wrong is the account of what makes people happy."*

I, therefore, follow Adler's (Adler, 2012; Adler and Posner, 2008) claim here that preference *should not* be equated to SWB because preferences contain more than just a reflection of SWB or mental states. Although, SWB (especially evaluative SWB measures) can predict choice and preference to some degree (e.g. see Benjamin et al., 2012), there are times when they can diverge. For example, in health where adaptation plays a significant role in SWB ratings we find that people can assign high quality-adjusted life year (QALY) values (based on preferences) to health states to which they adapt in terms of experience SWB (Adler, 2012). Likewise, Smith et al. (2006) find that current colostomy patients report reasonably high levels of life satisfaction and mood, such that adaptation is nearly complete, but at the same time they express a willingness to reduce their life-span by a substantial 15 per cent in exchange for a return to perfect health. Adler (2012. p.19) states that a utility function is simply a "mathematical device" for representing an individual's preference rankings. Since the utility function can contain non-mental state entries, then the individual can have higher utility even though her subjective experiences may not have changed.

Therefore, I am proposing that the WV approach take/align itself with the EQ defence. This is supported by Parfitt's distinction between the three different measures of welfare and by Kahneman's categorisation of wellbeing measures. Kahneman (2012) terms the preference-based welfare measures used by economists as 'Decision utility' and separates this from other accounts of wellbeing that he labels 'Experienced utility', which refers to experience SWB or affect and 'Remembered utility', which refers to evaluative measures of wellbeing (objective list measures of wellbeing are not included in Kahneman's categorisations). It is fair to say that these definitions and distinctions are recognised by a number of economists working in this area too (for example Dolan, 1998; Dolan and Kahneman, 2008).

A further reason for taking the EQ defence is that if preferences can in some cases be context-dependent, irrational and mis-informed, as some of the major criticisms of preference-based valuation methods have suggested, then strategically we do not want SWB to simply reflect preference rankings. Indeed, many of the advantages associated with wellbeing valuation would be annulled if we adopted the PR defence and claim that SWB equals or evidences preference satisfaction.

3.3.3.2. The PR and EQ defence and Sugden

Interestingly Adler's PR and EQ defence categorisation overlaps to some degree with Sugden's (2015) and (2018) discussions of issues with preferences in economics. Here we can interpret Sugden's arguments in to a set of two separate solutions for preference anomalies and problems:

Solution 1: Assume that within each agent there exists a 'rational self' that is frustrated by a 'behaviourally-susceptible' outer 'shell' and help the agent to elicit their preferences more accurately (Sugden, 2018).

In other words, inside each of us there is a calculating sophisticated individual who, given sufficient information, will consistently make choices according to the rationality assumptions in economic theory, but in reality this rational self has a hard time being heard because our environment can lead us astray. Here the environment could be something like a price prime or anchor that people end up relying on due to time and resource constraints. Sugden's theory has clear parallels with Kahneman's (2012) System 1 and System 2 framework. Within this setting Sugden is essentially saying that our System 2 brain is ok in that it works how economic theory suggests it does, but that we are often misled by System 1. The solution is, therefore, to design surveys and studies in such a way that it makes it easier for people to provide us with their real preferences and values. There is precedent for this school of thought in the valuation literature – it is essentially what Gregory et al. (1993) meant when they said that,

"designers of a CV [contingent valuation] study should function not as archaeologists, carefully uncovering what is there, but as architects, working to build a defensible expression of value" (1993. p.179).

Most of the options under this solution are preventative pre-survey methods which generally try to provide assistance before or during the survey to help participants elicit values from the 'rational self'.

Two editions of the *Journal of Environmental and Resources Economics* (in 2005 and 2010) were dedicated to methods that have been developed to deal with preference anomalies in contingent valuation studies and many of these solutions make up part of best-practice methodology in CV today. A key mechanism that sits at the heart of anomaly reduction techniques in these papers is through learning by repetition and experience. The work is based on Plott's (1996) Discovered Preference Hypothesis (DPH). The DPH argues that stable and consistent preferences are the product of experience gained through repetition. There are a number of studies that report reductions in the effects of arbitrary anchors and in the number of preference reversals as people become familiar with the good and the institutional payment arrangements in a contingent valuation context (Bateman et al., 2006; Braga and Starmer, 2005).

Contrary to the recommendations set out by the National Oceanic and Atmospheric Administration (NOAA) in 1993 (Arrow et al., 1993), which recommended a single-bound dichotomous choice format in order to mimic a market setting, Bateman et al. (2006) propose a double-bound dichotomous choice payment format for eliciting values. This, they say, is to allow for learning and experience as it gives participants the opportunity to "discover" their preferences during the survey. Alternatively, Gregory et al. (1993), propose a deliberative CV mechanism (multi-attribute utility analysis) in which a group of stakeholders, that includes the affected citizenry and technical experts, assesses the merits of the good under consideration and determines which attributes have the greatest impact on utility. Consequently, some contingent valuation surveys now employ a workshop format whereby people discuss the valuation issues with others and they can seek further information from moderators and experts again in an attempt to help them discover their true preferences (Hanley and Shogren, 2005).

Possibly the most novel approach in this area has been taken by Bateman et al. (2009) who have used virtual reality simulators to communicate environmental changes to survey respondents so that they can gain direct experience of the non-market goods and outcomes at hand. This had the effect of reducing the difference between WTP and WTA for environmental goods.

Other frequently observed problems in CV that fall under the category of embedding effects discussed in Chapter 2 have been argued to be a consequence of the survey instrument. It has been argued that insensitivity to scope findings are idiosyncratic and/or that the studies that have obtained such results are flawed in terms of survey design (Smith, 1993; Carson et al., 2001). For example, the finding of insensitivity to scope should not be surprising if the description presented is not adequate to enable the respondent to distinguish between the smaller and larger good or if the survey emphasises the symbolic nature of providing the good. Sequencing effects and sub-additivity effects have also been argued to be explainable with reference to income and substitution effects (Hoehn and Randall, 1989; Carson et al., 2001). Intuitively, each new good obtained reduces the income available for respondents to spend on other goods. Given this, the later in the overall package that a good is offered, the less desirable it will look. There may also be a similar effect if the goods are substitutes for each other.

It is fair to say that it is not fully clear how applicable these results are for preference-based valuation methods because opportunities for learning are often minimal. It is hard to provide repetitive experience for many of the public goods assessed in stated preference (Braga and Starmer, 2005) and there are likely to be constraints on the types of non-market goods and outcomes that can be simulated effectively in virtual simulators. These methods probably make good sense for environmental issues, but they are harder to employ effectively in, for example, health and education-related interventions. Where we use revealed preferences in proxy markets such as housing and labour markets, transactions are often infrequent (i.e. we don't move house, change job or negotiate wages frequently) so that few chances for learning exist at the individual level (Genesove and Mayer, 2001).

Another popular preventative method for reducing biases in CV has been the use of entreaties, which are text boxes that remind respondents of their budget constraints and to ask them to provide a truthful and accurate response. They can go as far as asking people to take an oath before answering the questions. Entreaty scripts have been shown to be effective in reducing hypothetical bias in WTP values in CV studies (Cummings and Taylor, 1999).

It is possible to align Sugden's theory of a behavioural shell with the PR defence of SWB. If SWB provided accurate information on preferences, then presumably they could be used to garner preference data without being affected by the behavioural biases and issues inherent to preference elicitation methods – in other words, they could be used to get to the individual's 'rational' preferences. Thus, in this solution preference is still king and we could measure it by either modifying and improving how we elicit preference data or by using SWB data instead because under the PR defence SWB equates to preference. Therefore, we can conclude that the general consensus to equate SWB with preference in the WV literature to date aligns those proponents with the solution set out here and with Adler's PR defence.

A different solution that is recommended by Sugden (2018) is to move away from preference in economics and to think instead about opportunity sets. We could apply a similar argument in favour of SWB and would come to the following solution.

Solution 2: Discard the preference satisfaction account of welfare in favour of self-reported measures of wellbeing (subjective wellbeing) and estimate the value of non-market goods using the Wellbeing Valuation approach.

This approach is far more drastic than Solution 1. Approaches under Solution 1 are positive (methodological/technical) in nature, whereas Solution 2 requires discussion of and reference to both more fundamental normative and positive aspects of economics. Economists have always been in the business of developing piece-wise non-substantive (in the normative sense) adjustments to their theories and methods so they can better model the way that agents and economies behave, all usually still within the standard preference view of the world and Solution 1 is typical of this tradition.

Solution 2, on the other hand, offers a more radical approach, but one that is *not* new to the economics profession (Kahneman et al., 1997). The fundamental premise of Solution 2 is a move (back) to self-reported measures of welfare, whereby the impact (and value) of non-market goods is measured in terms of the effect they have on people's self-reported wellbeing rather than in terms of people's preferences (or opportunities if we were to take Sugden's recommendations).

Solution 2 is consistent with the EQ defence and the approach I have taken here because it does not attempt or assume to equate SWB with utility and preference.

Through the development and merging of Adler's EQ defence and Sugden's solutions (or a version of it) we have developed, for the first time, a comprehensive framework for conceptualising where WV sits in the valuation landscape. As argued here, WV is not a complement to preference-based methods. WV is entirely separate from preference methods and should be seen as its own unique methodology for estimating ES and CS values. This has two major implications that I discuss in more detail later. Firstly, it means that WV values cannot be used with preference values and secondly it means that WV values should not be compared to preference values and certainly their accuracy should not depend on their ability to align to or mimic preference values (i.e. WTP and WTA) because SWB and utility are two entirely different concepts with no real reason for there to be a convergence in values when using the two different approaches. This is a substantial, but I would argue, correct divergence from the main thinking in the WV literature which has generally tended to assume that values derived from WV are the same in nature to (and can be compared with) preference-based values. This tendency to equate the two methods has been either explicit (for example Diener et al., 2009) or implicit in the literature, in the sense that studies have looked to assess the validity of WV against preference-based methods such as the results from SP studies.

In the final section of this chapter I will build a defence and rationale for WV, before then discussing methodological issues in later chapters.

3.3.4. The rationale for wellbeing valuation

In this section I set out the key advantages and pros of WV which provide the rationale for using the method. A contribution to the literature is that I categorise these pros into normative based issues and methodological issues, whilst adding some new insights. A key part of building the rationale for WV is also to consider and address any problems or cons that the method may have. I cover this in the latter part of this section and also in the next chapter. The key problems with WV relate to the validity of the measure of SWB used in the analysis and to the statistical method employed. In this chapter I provide a defence of life satisfaction, the key measure of SWB used in WV to date and in the following chapter I provide solutions to the methodological problems through a new approach to WV.

3.3.4.1. Advantages of the wellbeing valuation approach

3.3.4.1.1. Normative advantages

Probably the key advantage of WV is that it gets around the issue of incommensurability (Anderson, 1995). Many philosophers would argue that many non-market goods, services and outcomes are not commensurable or comparable with money and hence cannot be valued according to economic theory. This type of criticism is especially prevalent in the valuation of the environment, health and human life. An outcome of this is the type of lexical preference orderings discussed in section 2.4.2., in which no amount of money would equate to the non-market good. Here in stated preference no finite WTP or WTA amount would exist for the non-market good making it impossible to place a monetary value on it. The end result in CV studies is often a protest value, where despite valuing or appreciating the good in question people state a zero WTP amount.

Other forms of incommensurability can also present itself in other guises related to substitutability. Substitutability requires that a gain in one good can offset the loss in another and vice versa. Here two objects may be commensurable in the eyes of the individual but the individual has no sound internal basis or method to make the comparison between the two objects. In this context we may find that the monetary

values that are revealed in proxy markets or that are stated in surveys will be quite arbitrary and may not be uncovering what the economist really seeks to discover. As a result, this may lead to some of the anomalous preference behaviour that I have discussed in section 2.4.2. such as the anchoring effects found by Ariely et al. (2003) and difficulties in converting feelings on to monetary scales (e.g. Lowenstein and Schkade, 2003; Kahneman et al., 1998).

In the WV approach the issues of incommensurability and non-substitutability do not surface because we do not need to ask people to translate a sentiment or feeling towards a non-market good on to a monetary scale. Instead we are simply measuring the welfare impact of the good or outcome in monetary equivalent terms without ever asking the individual to make a comparison for us. A key point at the extreme is that even if individuals believe that certain goods are incommensurable with money it is possible to estimate values in WV that are precise and non-arbitrary reflections of welfare change and which align with economic theory. WV is the only valuation method that can solve for the problem of strong incommensurability and non-substitutability. Clearly this is a major advantage of the WV method, since valuation theory is highly dependent on these two assumptions. This advantage is especially relevant in policy areas that have traditionally been very problematic for valuation, such as health valuation and the environment.

A second and related advantage is there may be times when people (and markets) do not feel it morally justifiable or acceptable to trade or place a value on certain types of goods (like wildlife, health and education) in actual or in the case of SP, hypothetical markets. This is different to the issues of incommensurability and non-substitutability as it may be that we can think of a monetary amount that would equate in value to the issue at hand, but we would rather not have to think in this way for moral or ethical reasons. Preference-based valuation approaches rely on what can be called the 'commodification' of goods – that is, to value a non-market good we need to have people think and act as if the good was actually traded in markets. This may be inherently difficult at best, and morally unacceptable at worst and is likely to result in issues such as protest values and anchoring effects which were discussed in section 2.4.2. For the same reasons just discussed above, WV can get around this issue and

still provide values in these situations without having to force people to make difficult ethical choices.

And thirdly, as raised in this thesis, WV is the only direct approach to valuation. For the purest it is the method that most closely replicates Hicks' theory of valuation. It can be argued that WV is, therefore, conceptually cleaner and easier to trace and defend in terms of its normative rationale.

3.3.4.1.2. Methodological advantages

There are a number of methodological advantages associated with WV and so I shall list them here. Stutzer and Frey (2010. p.23-24) claim that the WV approach "avoids some major difficulties inherent in both stated and revealed preference methods".

1. In WV there are no rationality assumptions, like those needed if we are to use preferences in valuation. The method "does not rely on respondents' ability to consider all relevant consequences of a change in the provision of a public good. It suffices if respondents state their own life satisfaction with some degree of precision" (Stutzer and Frey, 2010. p.23; van den Berg and Ferrer-i-Carbonell, 2007) and this is probably a cognitively less burdensome task than thinking about a WTP value (Frey et al., 2009). In the critique of preference-based methods we saw how violations of the rationality assumptions lead to preference reversals which pose a major challenge for SP methods. We also saw how in studies a large number of subjects exhibit choice behaviour that is inconsistent with the rationality assumptions (Choi et al., 2011). A major draw of the WV method is that it does not suffer from this issue because it does not use preference data.

2. WV does not rely on market efficiency assumptions that are critical to revealed preference methods (Van Praag and Baarsma, 2005; Frey et al., 2004). If markets are imperfect, large transaction costs exist or people suffer from lack of information then market behaviour will reveal very little about people's WTP and WTA. In fact, WV can add further important information when market assumptions may not be met; the WV approach is able to pick up any residual effect of the non-market good after allowing for market compensation. If a non-market 'bad' like crime is not fully

reflected in lower house prices then we would expect crime to impact on SWB after accounting for house price differentials and hence in the WV approach we can value this residual impact (Stutzer and Frey, 2004a; Ferreira and Moro, 2009; Luechinger, 2009; Levinson, 2012; Van Praag and Baarsma, 2005; Cohen, 2008).

3. WV does not rely on any forecasting or prediction of preference on behalf of the individual. Market transactions and stated preferences require people to accurately predict future welfare impacts of the good, but as discussed in section 2.4.1. there is a lot of evidence that people find it difficult to predict future utility impacts even for simple everyday products. Furthermore, the value of something will depend on the likely fruition of some factors – for example the value of an umbrella depends on the likelihood of rain - and people's perceptions of future risks may not align with actual events (Luechinger and Raschky, 2009). WV relies only on people's actual experiences: we can see how the good impacts on people's welfare under the actual conditions that come to fruition.

Decisions in markets (real or hypothetical) may not accurately reflect people's experiences (Luechinger and Raschky, 2009) and in hypothetical settings this may get exacerbated by the focussing illusion, whereby people focus their attention on the salient aspects of the non-market good at the time of preference elicitation, whereas in reality these aspects have little or no consequence for how they actually experience their lives. Outside of the survey or lab, in day to day life the non-market good will have to vie for attention amongst all of the other things that are going. As Kahneman (2012) puts it nothing is as important as you think it is when you are thinking about it and hence stated preference methods are likely to overstate values all else constant. This leads to the issue of hypothetical bias in stated preference as discussed in section 2.4.2. A major advantage of WV is that we can find out the importance and values of non-market goods alongside all of the other things that affect people's lives hence eliminating hypothetical bias.

4. WV does not suffer from the broad range of survey-related biases inherent to stated preference methods because respondents are not asked directly to state a value or pay a certain price (van den Berg and Ferrer-i-Carbonell, 2007). For example, it is not possible for respondents to use strategic answers (van Praag and Baarsma, 2005; Frey

et al., 2004) or protest values. Also, information bias and interviewer bias do not exist in WV and there will not be the issues around comprehending the payment mechanism that we see in SP methods. Furthermore, there is no possibility of embedding effects or priming effects, such as anchors.

5. Related to this, since we do not elicit values from individuals in WV we do not change subsequent perceptions of the non-market good and hence can assume that the values will be accurate and stable even with provision of the good (see arguments on NIMBY policies in section 2.4.2.).

6. SWB or life satisfaction questions are relatively easier (than WTP questions) to answer for respondents (van den Berg and Ferrer-i-Carbonell, 2007). Typically the percentage of people who do not respond to SWB questions in surveys is low (van Praag et al., 2003) . This helps reduce biases due to sample selection, which will improve the valuation results from WV making them more generalizable (van den Berg and Ferrer-i-Carbonell, 2007).

7. There are two issues regarding the estimation of the indifference point. In WV we eliminate the cognitively demanding task of estimating the indifference point (as was demonstrated in the Mosteller and Nogee (1951) paper), because the analyst now does this on behalf of the respondent using statistical methods. Under wellbeing valuation it suffices that we measure the impacts of money and the non-market good on SWB and the indifference point can be measured simply from the ratio of the impacts of these two goods as per equation (10). As we will discuss there are issues related to measuring the impacts of income and the non-market good on SWB, but the actual task of estimating the indifference point is simple in WV.

In regards to the indifference point, there is an issue that has not been picked up to date in the literature and that I contribute in this thesis and that is whether people in stated preference surveys act in the way that is assumed by economic theory. That is does a willingness to pay value really represent the tipping point at which the individual is indifferent between the good in question and the money value? This is a separate issue to the one described by Mosteller and Nogee (1951). Mosteller and Nogee were interested in the cognitive burden associated with estimating the

indifference point. Let's here assume that people can actually make that calculation. The question I am asking here is then whether the figure that they state is the true point of indifference. Even if people can calculate the point of indifference (in a cognitive sense) do they estimate the real or right one? Economics assumes (as does many branches of psychology and philosophy) that human beings fundamentally act in order to maximise their own welfare. Accordingly, every purchasing decision is only made when the perceived welfare benefits of the good exceed the welfare costs associated with foregoing the money to pay for it. Taking this line of thought, we might conclude that humans are likely to be conditioned in to making some surplus in every purchasing decision based on the information at hand. In reality they may be misinformed and hence do not derive any surplus or even a dis-surplus out of the good, but this does not affect facts about their intended behaviour.

Let us assume that a given population on average seeks to derive 10% surplus on any purchase/transaction - this could be seen as a spending heuristic. That is they derive surplus to the value of 10% of the price they pay, which means that for say someone that buys a new computer for £1,000, he values the computer at £1,100 and if the computer were worth less than £1,100 to him he would not buy it (we can of course allow for different surplus requirements across different types of goods – e.g., very expensive goods may require only 3% surplus since this can be a large sum in absolute terms – but I will stick to a simple average here). Now assume that the computer is some non-market good like a nice view or lower crime rates. In theory for CBA we would want to measure the value of the non-market good as £1,100, but it would seem to be highly unlikely that this is the value that people would give us in a stated preference survey. If people are programmed to maximise welfare and hence seek surplus in any transaction then it becomes difficult to argue that they would be able to easily state a maximum (surplus exhausting) value of £1,100 in a one-off survey. Even where stated preference surveys use multiple bounded shots - whereby the survey enumerator increases values until the individual states he would not be willing to pay any higher - there is no guarantee that the WTP value at which people stop includes the surplus that they would traditionally extract in a purchasing decision.

There are very few occasions in life when people are required to state a maximum WTP or WTA amount truthfully. One occasion may be in auctions, although one could argue that supposed maximum WTP values stated in auctions still contain a purchase surplus (i.e. the maximum WTP is reduced somewhat to incorporate the surplus requirement). Plenty of evidence shows that people use a number of heuristics in stated preference surveys and a *surplus seeking heuristic* could be an important additional one and one that we have not tested before. Note that a surplus seeking heuristic would also have implications for revealed preference methods because it would mean that we would not be able to estimate correct demand curves from market data. If the heuristic legitimately exists we should conclude that WTP values in stated preference surveys and from market data are likely to be *understated* all else constant. WV is the only method that would allow us to estimate unbiased estimates of CS, ES and monetary value under the presence of a surplus seeking heuristic. This is because we can assess the full impact of the non-market good on welfare and derive the equivalent amount of money that *exactly* exhausts all surplus, such that the individual does remain at his/her original welfare position and no surplus is extracted.

8. The WV method has a broad application; indeed it could be used for any non-market good we have data on and where this may not exist we can collect primary data. Consequently, WV is far broader and of wider application than revealed preference methods and are on a par with stated preference methods in this respect. The key difference is that stated preference methods can be widely applied (you can ask a WTP or WTA question about any good you like) by virtue of the hypothetical nature of the survey instrument and scenario, where any question can be posed to respondents. An important advantage of the WV approach is its ability to be similarly wide in its application, but – and this is key - without the need to use hypothetical settings and questions: WV is based entirely on people's actual experiences (Stutzer and Frey, 2010; Luechinger, 2009; Dolan and Kahneman, 2008). This should sit as an important advantage of WV in relation to stated preference from the perspective of economists who tend to prefer revealed preference over stated preference valuation methods because the former are based on actual behaviour. Like with revealed preference, WV has some trace that can be picked up in people's behaviour and experiences, which may make it more justifiable than decisions based on hypothetical scenarios and states of the world. Certainly, an economist may tend to think this way.

3.3.4.2. Problems associated with the wellbeing valuation approach

3.3.4.2.1. Statistical methodology/technical problems

As described above, WV is a statistical approach and there are a number of technical challenges associated with the methodology. These issues mainly relate to the challenge of estimating the impacts of the non-market good and income on SWB – in other words whether we can derive robust estimates of β_Q and β_M in Figure 2. Discussion of these technical issues requires a full exposition of the statistical methodology that underlies the WV approach and cannot be understood in the context of the introduction to WV provided to date. In this thesis I develop a new framework for assessing the validity of the WV method and so I will, therefore, cover the technical problems related to WV in Chapter 4 in the process of developing the new WV approach. In Chapter 4 I will present these technical problems and provide a set of statistical solutions to address them.

In this section, therefore, I focus on the issues related to using SWB in valuation. In this regard, there are two issues: (i) general criticisms of the use of SWB measures in economics; and (ii) problems regarding the measurement of SWB. I provide a defence against these problems at the end of this chapter.

3.3.4.2.2. General criticisms of SWB in economics

One of the strongest attacks on SWB in economics has come from Gul and Pesendorfer (2008). The general gist of their argument is that economics as a discipline has no substantive element or desires. They say that economics simply "provides a benchmark for the performance of economic institutions at aggregating individual preferences" (p.4) What is relevant here are the agents' preferences as perceived by themselves and discussions of experiences "play no role in standard economic analysis because economics makes no predictions about them and has no data to test such prediction" (p.2). In sum, agents' preferences are "given" and economics merely "evaluates the performance of economic institutions" (p.33). This kind of distinction between preference in economics and measures of SWB is also made frequently elsewhere, for example by Kimball and Willis (2006).

Other criticisms come from Sen (1999) and Loewenstein and Ubel (2008) who are sceptical of SWB measures because of the possibility of adaptation to circumstances. This can be summarised by the 'happy slave' phenomenon, whereby over time a slave could adapt to his dire circumstances to the extent that he does not report a low level of SWB anymore, despite fairing badly on nearly every kind of measure we would usually associate with human welfare such as good health, freedom and dignity. Near-complete adaptation to significant life events such as paraplegia, lottery wins and end-stage kidney disease has been well-documented in the literature (Loewenstein and Ubel, 2008).

Loewenstein and Ubel (2008) also raise the problem that SWB (especially experience utility) may fail to capture the wide range of things that people deeply care about in life and hence may not be inclusive enough of a measure for policy analysis. This echoes long-standing concerns voiced in other disciplines, mainly from philosophers. Veenhoven (2004) organises the critiques on two grounds: First it does not make sense to prioritise one particular value only and second, that there are other values that rank higher than SWB and on the latter experience machine type arguments often come to the fore (Nozick, 1974).

There is also the question of the sensitivity of SWB scores. Johns and Ormerod (2007) and Di Tella and MacCulloch (2006) are concerned with the bounded nature of SWB measures. Since measures like happiness and life satisfaction are measured on scales like 1-7 or 0-10 this may make it insensitive to small impacts and after some point an effect, no matter how large or important may not be able to show up on people's self-reports as people cannot state any higher (lower) than the upper (lower) bound on the scale. Relatedly, a number of studies have found that life satisfaction is highly stable over time due to personality traits (Costa and McCrae, 1980; Chamberlain and Zika, 1992; Eid and Diener, 2004; Lykken and Tellegen, 1996). Both Eid and Diener, (2004) and Lykken and Tellegen, (1996) find that only about 15%-20% of the variation in evaluative wellbeing measures like life satisfaction is due to external factors and life circumstances. According to Lykken and Tellegen, (1996. p.188, 189) wellbeing is "largely determined genetically" and so trying to increase it is "futile". Sunstein (2015) states that SWB metrics have limited reliability for policy analysis since it is difficult to map policy and regulatory changes onto

SWB. He claims that SWB measures are too 'crude' and 'coarse' and may not pick up things outside of significant life events. In terms of the latter Sunstein does concede that SWB measures do provide useful welfare information on employment and labour market interventions where the evidence is consistent and compelling. It should be noted that his claims are mainly made on intuitive grounds, rather than being based on empirical studies.

3.3.4.2.3. SWB measurement issues

In addition to the broad challenges set out above, there are issues regarding the extent to which SWB can be measured accurately in surveys. As discussed, the main SWB measure used in WV has been life satisfaction, which is usually elicited from the following type of question (taken from the British Household Panel Survey): *"How disatisfied or satisfied are you with your life overall?"* and responses are made on a scale of 1-7 or 0-10 and so on. I will focus on the measurement problems related to life satisfaction here. This section is mainly to recognise these problems and for completeness. And it should be noted that the adverse implications that they may have for the WV approach more generally is limited for two reasons. The first is that as I shall discuss in the next section many of these problems may not be as bad as once thought and that in fact there are strong arguments for using life satisfaction in WV. The second is that the WV approach is a general approach to valuation that can be applied with any measure of SWB and the solutions and new framework that I develop in Chapter 4 can be employed with any SWB measure. Therefore, any failure related to life satisfaction as a measure of SWB does not in any way imply that the WV approach is condemned or doomed. However, if we continue to use life satisfaction in WV (as I do in the new approach set out in Chapter 4) we need to be aware of these problems and issues.

Life satisfaction is an evaluative measure of SWB, which it has been proposed, can be seen as being made up of a balance of affect (positive and negative emotions and feelings) together with a cognitive assessment of how well one's life measures up to aspirations, goals and the achievements of others (Kahneman and Krueger, 2006; Diener, 1984). A life satisfaction response is also said to incorporate to some extent a

retrospective judgement of one's life together with how one feels now (Kahneman and Krueger, 2006).

This retrospective element of the life satisfaction measure is what distinguishes it from experience SWB and is one source of the problems associated with evaluative measures because people do not always correctly remember past experiences. Furthermore, people's present feelings can be influenced by contextual factors present at the time of the interview and biases can also arise in the stage of verbally reporting life satisfaction scores and due to adaptation (Bertrand and Mullainathan, 2001; Kahneman and Krueger, 2006; Schwarz, 2012; Schwarz and Strack, 1999).

i. Remembering past experiences

Experiments have shown that people's remembered experiences can be biased due to their tendency to adopt a peak-end rule; in retrospective evaluations people place greatest weight on the peak (more intense part) and on the end of an experience. They attach less weight to the duration of an experience. There is therefore often a mis-match between people's actual experiences at the time and their retrospective evaluations of these experiences (Kahneman et al., 1993; Schwarz, 2012). Wirtz et al. (2003) for example, compare people's evaluations (satisfaction) of their holidays against their experiences during the holiday and find that people cannot accurately remember the wellbeing actually associated with holiday trips.

ii. Within-person comparisons: which information is used?

Bodenhausen and Wyer (1987) find that when responding to questions on satisfaction, "people truncate the search process as soon as enough information has come to mind to form a judgment with sufficient subjective certainty". The judgment, therefore, tends to rely on the information that is most accessible in the moment and that accessibility depends on: a) the recency of the information and b) the frequency of its use. Self-reported satisfaction scores may thus only reflect a part of the experience of the individual tainted by most recent events and experiences. For example, analysis of longitudinal data on job satisfaction from the UK and Germany finds that peak and end job satisfaction are better predictors of quitting than overall job satisfaction ratings (Webb and Sheeran 2006).

In terms of accessibility, Strack et al. (1988) famously find that question ordering can influence this. By asking students how frequently they go on dates and their life satisfaction, they find no relationship if the life satisfaction question precedes the dating frequency question (correlation = -0.12) and a strong relationship if the ordering is reversed (correlation = 0.66). This was replicated by (Schwarz et al., 1991) with questions on marriage satisfaction and life satisfaction. The authors suggest that the effect of question-order effects increase when this draws attention to information that is not chronically accessible (eg: dating frequency rather than chronic pain). Question order effects, thus, may not affect all respondents (Schwarz and Strack, 1999) – eg, respondents currently undergoing a divorce are unlikely to be affected by whether they are asked to consider their marriage before or after the general question because this information is frequently used by them (e.g. it relates to their current concerns). In other more recent studies a survey of fans of two English football clubs in the 2008 Champions League final found that fans of the losing team were less happy after the event when they had been asked about their happiness before the event (compared to those only asked after the event) (Dolan and Metcalfe, 2010). In this case, there was a contamination effect, where being reminded of one's happiness in association with the defeat produced lower levels of happiness. This is in line with other studies (for example, Bertrand and Mullainathan (2001)).

Further, conversational norms may amplify question order problems. For example, topics of conversation typically follow a logical pattern. So the respondent may assume that if they were just asked about their marriage satisfaction and then their life satisfaction, the question on life satisfaction should exclude how they feel about their marital life as they were just asked about that topic separately (Schwarz and Strack, 1999). In sum, the information made salient by previous questions can impact heavily on SWB and satisfaction scores.

iii. Between-person comparisons: Comparing self to others

How satisfied we feel about our lives can be impacted on by whom we compare ourselves against. Strack et al. (1988) find that when interviewed by individuals with a disability, respondents have been found to subdue their life satisfaction responses. In contrast, when a disabled person was present in the same room as a respondent completing their own survey, their condition was used as a standard of comparison

with the result that life satisfaction scores were inflated. Recent studies using large national datasets corroborate that life satisfaction responses depend significantly on whom we compare ourselves with (Becchetti and Pelloni 2013; Frey et al. 2014).

iv. How an assessment of one's life is constructed

Satisfaction scores also depend on what prior life events come to mind when making an assessment. Because of the accessibility effect, a positive or negative life event that comes to mind may impact on life satisfaction. Or prior life events may result in a positive (or negative) effect on life satisfaction scores because they create a benchmark (known as a contrast effect).

Strack et al. (1985) asked respondents to report either three positive or negative recent events that are temporarily accessible. Respondents reported higher life satisfaction when they thought of positive recent events. The authors suggest that people include accessible recent events when assessing their current lives but use distant events to form a standard of comparison (Tversky and Griffin, 1991). The problems are that (i) the memory of these events may not be a fair assessment of how good or bad comparators they provide, and (ii) that trivial recent events and circumstances may taint the overall life satisfaction scores of respondents, leading to a mis-alignment between life satisfaction reports and actual experiences.

v. Context effects

The research instrument itself and other contextual factors can have a large influence on responses to life satisfaction questions. Current mood can impact on responses to life satisfaction questions in two ways. Thinking about one's life whilst in a good mood may lead to the selective retrieval of positive information that leads to an affirmation of their life and a more positive evaluation. Or people may also take their current mood as a good indicator of their wellbeing in life in general (Schwarz and Strack, 1999). Some evidence suggests that the latter explanation may be more accurate and that people use a 'current-mood-heuristic' to judge overall life satisfaction (Schwarz and Clore, 2003). There are a number of factors that can impact on people's moods. In a set of famous studies Schwarz et al. (1987) (see Schwarz and Strack, 1999) show that finding money before the survey, spending time in a pleasant versus unpleasant room or seeing your football team win the night before increases

life satisfaction responses. Schwarz and Clore (1983) telephoned people on sunny and rainy days to ask life satisfaction. The results showed that on sunny days, people reported being in a better mood, being happier and having higher life satisfaction. Whilst such trivial factors are likely to influence current mood, they should not have notable effects on overall life satisfaction and it is telling that in the Schwarz and Clore (1983) study when the weather was drawn to the respondent's attention, this mood effect disappeared, meaning that it isn't salient information in the construction of one's overall life satisfaction. Relatedly, Kavetsos et al. (2014) explore the influence of calendar effects on reports of life satisfaction. They use Eurobarometer data from 31 countries over a 20 year period. They find that day and month of the interview are statistically significant, but not time of day. Their results show that, compared to June, life satisfaction increases in December, January and February and decreases in October. And that life satisfaction responses are lower on Sundays, which they suggest reflects pre-work anxiety.

vi. Reporting life satisfaction

Individuals may adjust their life satisfaction scores when reporting them in order to give more socially desirable responses. For example, reported wellbeing is higher in face-to-face surveys than in postal surveys (Smith, 1979). When interviewed by individuals with a disability, respondents have been found to subdue their life satisfaction responses. In contrast, when a disabled person was present in the same room as a respondent completing their own survey, their condition was used as standard of comparison with the result that life satisfaction scores were inflated (Strack et al., 1990). Indeed, more generally, life satisfaction ratings are likely to be determined to some extent by the comparisons people make with their own life at different times and with other people at one point in time (Diener and Suh, 1997; Dolan and White, 2006). Furthermore, Haybron (2010) notes that we probably do not generally live our lives thinking about how satisfied we are at every moment, which may make life satisfaction a difficult concept to grasp, measure and report. As a result life satisfaction scores may be *'inert'* to life circumstances and events. The general problem with these effects is that respondents may provide assessments of their wellbeing that do not reflect the true experiences of their lives (Dolan and Kahneman, 2008).

vii. The impact of counterfactuals

When constructing an assessment of one's life satisfaction one's life's circumstances may be compared relative to a counterfactual state of the world. For instance, getting a place on the podium is an aim that most athletes will hold going into major competitions, but there is evidence to show that winners of bronze medals have been found to be more satisfied than silver medallists (Medvec et al., 1995).

One explanation is that missing out on the gold medal (counterfactual for silver medallists) hurts more than missing out on a silver medal (counterfactual for bronze medallists). Or it could be that the counterfactual for bronze medalists was actually no medal (fourth place). This effect depends how respondents explain current life events and circumstances to themselves, which may differ or change depending on the context they find themselves in during the survey (Boninger et al., 1994). If the counterfactual event was inferior then present satisfaction may increase and vice versa.

viii. Adaptation

It is often cited that evaluative measures are problematic due to adaptation effects. People in dire conditions may report reasonably high levels of evaluative wellbeing because they have adapted to their conditions, whereas on closer inspection their lives are terrible when measured on any objective outcome. This is Sen's 'happy slave' problem and is especially pertinent to health conditions. For example, Brickman et al.'s (1978) famous study showed that after some time paraplegics were no less satisfied with life than able-bodied people. These problems were a major driver of Sen's and Nussbaum's approaches to endowments and capabilities (Nussbaum, 2000).

ix. Measuring experiences

Many of the criticisms above relate to the fact that evaluative measures of wellbeing, such as life satisfaction, may not be accurate reflections of the quality of our experiences at the time. Kahneman has been a proponent of using experienced utility, defined as the quality and intensity of an hedonic experience as the basis for policymaking (for example see Dolan and Kahneman, 2008; Kahneman and Krueger, 2006; Kahneman and Sugden, 2005; Kahneman et al., 1997). Experienced utility is a sum of the moment-to-moment 'utils' of an experience and it can be traced back to the work of the Classical Utilitarians, such as Jeremy Bentham and John Stuart Mill

(Backhouse, 2002; Hausman and McPherson, 2006). Experienced utility can be measured using the Experience Sampling Method (ESM) (Csikszentmihalyi, 1990) or the Day Reconstruction Method (DRM) (Kahneman et al., 2004). The ESM collects information on people's reported feelings in real-time during selected moments of the day using a Personal Digital Assistant (PDA). Respondents report their activity at the time and their subjective experiences, such as anger, happiness and fatigue.

This does not involve a cognitive assessment of well-being on behalf of the participant and is therefore a measure of peoples' positive and negative affect (Kahneman and Krueger, 2006). One criticism of the ESM has been that it is intrusive and can interrupt the flow of people's experiences. As an alternative, the DRM was developed. This method asks people to fill out diaries of their day reporting what they were doing and how they felt during those episodes in terms of positive and negative affect. The DRM is less intrusive than ESM, but does rely, to some extent on remembered utility, but the evidence suggests that over the span of one day DRM responses align neatly with ESM responses – in other words retrospective assessments covering one day or less are able to measure experiences well.

Experienced utility methods reduce reliance on remembered utility and are less susceptible to irrelevant contextual factors. ESM, in some circles, is now taken to be the *gold standard* in wellbeing evaluation and reporting (Kahneman and Krueger, 2006; Schwarz, 2012; Gilbert, 2007). An assessment of how life is going for someone can be gauged from the summation of ESM or DRM reports over a long period of time.

3.3.4.3. Arguments in favour of life satisfaction

So which measure of welfare should we use in policy analysis and non-market valuation? Although Haybron (2010) suggests that there is no consensus way of determining a good theory of welfare from a bad one, I will use a mix of empirical evidence and philosophical argument to put a case forward for evaluative measures like life satisfaction. However, as discussed it should be kept in mind that any measure of SWB can actually be used in the WV framework.

1. We should note that the correlation between life satisfaction and the supposed gold standard experienced utility (affect) measures is likely to be strong (Diener, 1993).

2. There is also a variety of evidence to suggest that overall life satisfaction is a good measure of well-being. Many studies have been unable to replicate the results from Schwarz and Strack's seminal work on the contextual biases in life satisfaction, hinting that these issues may not be such as concern as first thought (Haybron, 2010; Diener and Suh, 1997). Pavot and Diener (1993), Eid and Diener (2004), Fujita and Diener (2005) and Schimmack et al. (2002b) find mood, question order and contextual effects to be limited and problems are not so serious as to invalidate life satisfaction measures (Pavot and Diener, 1993). Eid and Diener (2004) find mood effects to be much more problematic for domain satisfaction measures than for global life satisfaction measures and that part of the mood effects may be driven by the use of a timeframe such as “these days” or “nowadays” in the evaluative wellbeing survey question wording (which helps to explain the differences between their study findings and those of Schwarz and Strack (1999). Diener et al. (1999) find social desirability influences to be minimal and in fact they claim that there is reason to believe that social desirability is a valid component of wellbeing as it taps in to important aspects of personality that are consequential for an individual’s wellbeing. Diener et al. (1999. p.53) conclude that response artifacts “do not represent any prohibitive barrier to the accurate assessment of SWB by direct self-report”. There is also some evidence that data collected from aggregated moment-to-moment experiences through ESM converge well with global or retrospective reports although this depends greatly on how ‘convergence’ is defined by the study (Scollon et al., 2003). Schimmack and Oishi (2005), Schimmack et al. (2002a) and Heller et al. (2004) find that most of the variance in life satisfaction is due to changes and impacts happening at domain level wellbeing.

Sandvik et al. (1993) and Shizgal (1999) demonstrate that there is a strong positive correlation between well-being ratings and emotions such as smiling and frowning. Research shows that Duchenne smiles (i.e. a type of smiling that involves a muscle near the eye called orbicularis oculi, pars laterali, which can distinguish between true and feigned enjoyment) are correlated with subjective well-being (Ekman et al., 1990). Urry et al. (2004) show that reports of life satisfaction are correlated with

activity in the left pre-frontal cortex of the brain, which is the area associated with sensations of positive emotions and pleasure.

3. Furthermore, many studies have found that wellbeing and life satisfaction are good predictors of future behaviour (Frijters, 2000; Clark et al., 2008; Scollon et al., 2003; Haybron, 2010) and health (Kimball and Willis, 2006), such as heart disease (Sales and House, 1971) and strokes (Huppert, 2006). Frijters (2000) finds evidence from large national German and Russian datasets that people try to maximise life satisfaction in their choices to some extent - people are more likely to try to change areas of their lives with which they are less satisfied. A number of studies have found that life satisfaction predicts suicide (Lyubomirsky et al., 2005). Benjamin et al (2012) compare people's choices and their predicted SWB under a variety of hypothetical decision scenarios, such as labour market choices. They find that SWB is systematically the best predictor of choice (compared to other life circumstances) and that among different SWB measures, such as happiness and sense of purpose, life satisfaction was the best determinant. Cohen et al. (2003) find that people who report higher life satisfaction were less likely to catch a cold and would recover quicker if they did. Kiecolt-Glaser et al. (2002) find that people with higher life satisfaction heal more quickly from wounds.

4. Life satisfaction also seems to be "observable to others" and there is strong convergence between self and third party (family members and friends) reports of one's wellbeing (Pavot et al., 1991) suggesting that in so far as we can take third party reports to be of genuine value - Lucas et al. (1996) discuss some reasons for caution - "life satisfaction is a consistent and stable phenomenon; it is not simply constructed at the moment by the subject based on short term factors" (Pavot et al., 1991. p.158).

5. Krueger and Schkade (2008) assess the test-retest reliability of life satisfaction responses. They question the same sample of women two weeks apart and find that correlation in life satisfaction responses was about $r = 0.59$, which relates closely to results from studies by Kammann and Flett (1983). Schimmack et al. (2002a) find higher retest correlations for life satisfaction over a three month period ($r = 0.73$). Krueger and Schkade (2008) conclude that these levels of test-retest reliability "*are probably sufficiently high to yield informative estimates for……research*". Other

related work of interest by Ehrhardt et al. (2000) finds that the within-subject variation in life satisfaction scores falls over time signifying that people 'learn' how to respond accurately to life satisfaction questions over time – with practice they are better able to assess their life satisfaction, which suggests that panel studies like the BHPS are able to increase the validity of their data over time.

In sum, Diener et al. (1999. p.278) claim that global self-report measures like life satisfaction "possess adequate psychometric properties" and "show moderate convergence with daily moods" and third party reports and "recall for positive versus negative life events". The evidence overall suggests that life satisfaction has reasonably high construct and convergent validity properties. Veenhoven (2004. p.7) states that although there is always potential to find some deficiencies he has reviewed the literature on the critiques of life satisfaction and has "concluded that there is no evidence that responses to these questions measure something other than what they are meant to measure".

6. The issue of adaptation seems somewhat overstated. I would agree with Layard (2006. p.29) who argues that "we should seek to work with human nature as it is". Hence if there are some experiences to which people do not adapt and others to which they do or partially do then this "information is relevant to policy". Hence, adaptation is something we should seek to understand and measure for policy rather than using the issue as an argument against some measures of SWB (Menzel et al., 2003).

7. I would also echo Loewenstein and Ubel's (2008) argument that experienced utility measures may not pick up everything that is of importance to people. This argument is also strongly made by some philosophers (e.g., Haybron, 2007; Haybron, 2000). Measures like life satisfaction will, in addition to mood, capture an evaluation of people's lives – how their life compares to their aspirations and to others (Diener, 1984; Diener, 1994; Diener et al., 2009). Although we have discussed how this may serve to bias life satisfaction responses when our judgements are strongly tainted by how well our peers are doing, the evaluative element allows us to pick up broad aspects of wellbeing such as goal attainment and the attitude towards a particular experience that on reflection may also be an important part of our wellbeing (Kelman, 2005) and many experiences may be more (or less) valuable at the point of reflection

than in the moment (Bok, 2010). Arguably life is not just about having a plurality of good moments it can be more or less than the sum of the parts and global evaluative measures are able to provide a more holistic perspective (Haybron 2007; Bok, 2010). Haybron (2007. p.120) asks the valid question "how important to me is something I care about, considered in isolation?" There will clearly be higher fidelity of ESM and DRM methods in reference to moment-to-moment experiences but the methods are unlikely to capture what Diener (1993) terms 'meta-moods', which concerns the conceptualisation of one's emotions.

The following thought experiment is interesting. Assume there is some organization or person(s) that have been tasked with the very difficult job of choosing what you should do at every stage of your life. The objective for them is to maximize your wellbeing. This could be for example the government, a dictator or your parents. In this scenario - which is not so far-fetched as it seems, because in effect this type of role is assumed by all parents at the early stages of a child's life – would you be happy with the decision-making organisation or person to base their decisions for you entirely on your hedonic state? I would posit that most people would want their evaluations of the events in their lives to count before the decision-making entity decides to prohibit marathon running (even if such evaluations are based to some extent on comparisons we make against others, on our current moods and so on). Indeed, if confronted by a choice between having the decision makers base their judgments for us on our hedonic states or on our evaluative measures of SWB, such as life satisfaction, then I would guess that most people would opt for the latter. I am not aware of any experiment that tests this hypothesis, but the fact that probably very few of us hold a grudge against our parents now (as adults) for constantly not letting us eat that extra chocolate bar or for not letting us play that extra hour on the Nintendo when we were kids is some supporting (if anecdotal) evidence - for if we had been fully informed as children, constantly eating that extra chocolate bar and getting to play one more hour on the Nintendo would have shown up negatively for evaluative wellbeing but positively for hedonic measures. Similar conclusions could be made for drug and alcohol abuse.

8. To this line of argument I would like to add examples of cases where an evaluative measure seems to be much more adept at picking up the wellbeing of an experience

than experienced utility measures. The first case is sleep. Take John, whose life is going well for him in all aspects except for his sleeping patterns and quality. John has experienced different variants of the same nightmare for 3 months on a fairly consistent basis. In the dream, John experiences everything as if it were real, but on waking quickly realises that it was all a bad dream and gets on with the rest of the day. For obvious reasons ESM methods cannot survey John during his sleep. In John's case he does not fear the nightmares before going to bed as he knows they are not real and are only 'mild' and on waking he can rationalise the experience and forget about them. Thus, ESM readings just before and after sleep would not pick up anything untoward. But given a choice John would surely prefer to be rid of the nightmares and the experience itself during sleep is unpleasant and on reflection John is bothered by them to some extent and does feel that overall his quality of life is reduced by the continuing nightmares. If this were a policy-related issue, then as policy makers we would surely want to help John overcome these problems and hence in this sense an evaluative measure (and indeed preferences) would provide a better gauge of John's welfare. There are experiences, therefore, that ESM and experienced utility measures cannot capture. This is quite a specific case, but as we shall see next there are others.

The second case concerns activities where the final episodes of the event have extreme contrasts. Two examples would be marathon running and childbirth. Take a professional athlete who has trained hard for years and is now in the final stages of an Olympics marathon competition, which he is to win and receive the gold medal. The marathon is of course a paramount part of the athlete's life and goals. But measuring the athlete's quality of life in relation to the marathon using ESM say by asking for hedonic wellbeing responses during the marathon would probably show the event to be neutral for wellbeing.

The starting couple of hours may be pleasant to some extent although increasingly hard-going. If ESM measures experienced utility as the hedonic states we would like it to, then we would expect the final few miles of the marathon to be devastatingly negative for the athlete's wellbeing. The win and realisation of the attainment of the gold medal would then bring euphoria, which would help to offset the negative affect. But because there are question marks about the extent to which hedonic states capture

an evaluative component it may not be possible for the athlete to rationalise his experiences over the last few miles and overall the whole experience of the marathon could converge on being neutral for the athlete's wellbeing. Now, of course if we were to take ESM data for a long enough period after the end of the race and receipt of the gold medal, then the whole event may show up as a very positive experience overall, but what if we want to survey the athlete soon after having finished the marathon or if we cannot take ESM data for an extended period of time, then we would be left with an overall neutral effect of the marathon win. This seems deeply worrying because intuitively we would expect such accomplishments to have huge positive impacts on wellbeing instantly and it could be argued that the greatest positive impacts would be right at the point of accomplishment, when euphoria and a realisation of what you have achieved set in. Also, some may argue that the final episodes of the marathon should not show up as negative for wellbeing, since although the athlete is in great pain and discomfort, it is a 'good' or 'purposeful' pain because it represents the culmination of the athlete's devoted work and training.

If, on the other hand, we were to survey the athlete using an evaluative measure like life satisfaction straight after the marathon win, then it is fair to say that we would expect a big positive effect on wellbeing. This would also be the case if we surveyed him a period of time after the event. And indeed if we were able to take a life satisfaction reading during the final stages of the run, we would expect it to show up as positive as the measure allows for the athlete to provide a more general evaluation of his life which means he can rationalise the pain and discomfort in the response.

A very similar story could be made for people giving natural birth, where the process is devastatingly (and increasingly) painful, but it is dominated by the positive euphoric outcome of bringing a new life into the world. Again, intuitively I think we would like the positive effect of childbirth to show up instantly and for the pain to be represented as 'good' or 'purposeful' pain and it could be argued that evaluative measures best capture this.

If experienced hedonic wellbeing measures were to produce wellbeing responses in line with our intuitions for these extreme contrast events, then it would suggest that they have somehow incorporated an evaluative component and hence by definition

would not be experienced utility measures any more. The three scenarios (sleep, marathon running and childbirth) set out above are of course quite specific, but they show circumstances when ESM and experienced utility measures are likely to do a poor job. There are likely to be other occasions and circumstances that we observe or can think of that are equally problematic for experienced utility measures. An interesting response to the problem is Dolan's (2014) suggestion to measure purpose in the moment as a hedonic measure. This would provide a potential solution to the types of issues discussed above, but I am not aware of any studies that have looked at the convergence between Dolan's hedonic purpose measure and evaluative SWB measures such as life satisfaction.

The upshot is that experienced utility should not always be seen as the 'gold standard' approach to measuring wellbeing. It is interesting to compare the situation with another important area of policy analysis, namely causal inference, to crystallise this conclusion. In causal inference randomised trials are taken to be the gold standard approach. There are of course many occasions when a trial cannot be undertaken due to practical, resource or ethical constraints and concerns. But still for *any* kind of intervention *if* a robust trial can be undertaken then it represents the best possible method for understanding causal effect. The same cannot be said of ESM and experienced utility measures, because as I have shown even if there were no constraints to the use of an ESM survey, the survey would not provide the best measures of human wellbeing in certain circumstances and for certain events and episodes. Thus, ESM does not attain the gold standard standing associated with randomised trials. The relative advantages of ESM and experienced utility compared to evaluative wellbeing measures are context-dependent rather than ubiquitous.

9. Although Kahneman and Sugden (2005) disagree from a theoretical standpoint in actuality when looking empirically at the large amount of academic work in this field it is extraordinary that a response to a simple life satisfaction question, which takes on average a few seconds to muster, is highly sensitive to nearly everything that we would expect and in the right direction – it varies with short, medium and long term factors and life events (Pavot and Diener, 1993; Schimmack and Oishi, 2005) – including anything from marriage to playing football or from employment to going to a library (Fujiwara et al., 2014; Fujiwara and Campbell, 2011). This suggests that

Haybron's (2010) concerns that life satisfaction may be inert to life circumstances are not supported by the available evidence. In some ways an argument for life satisfaction can be made in a manner akin to Milton Friedman's famous statement that a theory does not have to be realistic, it just needs to work/be predictive. For some, life satisfaction might not be a 'realistic' measure of wellbeing as it is a short single-item measure that may miss a lot, but despite this it certainly has shown that it has high predictive power as it aligns with everything we would assume to be of importance to wellbeing.

10. Life satisfaction permits the case where we feel life is going well although we may not feel happy at every minute and hence has an advantage over hedonic wellbeing measures in this respect. For example, in the case of training for a marathon in the rain where our watch broke which means we got home late and missed our favourite show on TV, we would probably have a low level of hedonic wellbeing throughout the gruelling training and when we got home to find that the TV programme had finished. However, some of us may not care about this at all in the grand scheme of things and may on reflection be happy with ourselves about having got in another training session under such testing circumstances.
Building on this argument an important advantage of evaluative wellbeing measures like life satisfaction is that it gives the individual the power to determine just how important their feelings are to their sense of wellbeing. Hedonic measures provide a real-time assessment of someone's feelings, but for some people those types of feelings may be irrelevant in some cases or may be more important in some situations (e.g. happiness is not an important factor for me when I am in the act of training hard in the gym or for a marathon, or when I am helping my children with their homework). People might have complex systems, processes and ideas about the importance of different feelings when thinking about how good their lives are which would be tracked and borne out in life satisfaction responses and scores, but which would not be accurately represented in hedonic measures. In sum, life satisfaction offers a meta-analysis over moods and feelings, as rated and judged by the individual, which hedonic measures cannot do at the risk of not providing accurate measures of wellbeing.

11. There are also issues related to cost and practicality that are conceded even by the strongest proponents of experience measures such as Kahneman and Schwarz (e.g. see Kahneman et al., 2004). DRM and even more so ESM methods are very costly to run as they require repetitive sampling over periods of time in order to build a picture of SWB. A typical ESM study lasts one to two weeks (Scollon et al., 2003) and so for policies that have impacts that last for only days or weeks ESM and DRM methods may be viable, but this becomes increasingly difficult if we want to know the impacts of policy outcomes like health, crime and employment over long periods of time such as a year. Evaluative measures are able to cover a much longer time frame (Scollon et al., 2003) and life satisfaction questions have a long history in large surveys – these types of questions have been asked since 1965 to more than one million people all over the world (van Pragg et al., 2003).

There is also the problem of attrition or selection bias in experience wellbeing surveys (Kahneman et al., 2004). Motivation plays an important role in whether people continue to complete experience surveys properly; Scollon et al. (2003) find that people in good health and those with more spare time (the unemployed and students) were more likely to complete ESM surveys. And more recent work such as Mackerron's Mappiness application for iPhone (http://www.mappiness.org.uk) clearly uses a highly self-selecting sample of the UK population (MacKerron and Mourato, 2009). Scollon at et al. (2003. p.16) conclude that "the most compliant participants for experience sampling studies will be conscientious, agreeable, non-depressed, young people who are not too busy – essentially college students". This in itself is an interesting population to study, but may not be whom policy makers are primarily concerned with. Thus, due to measurement issues, by no means is there a consensus among wellbeing scholars that experience measures are the 'gold standard' (Scollon et al., 2003) and there is uncertainty regarding whether the additional costs associated with collecting ESM or DRM data is outweighed by the benefits of experienced utility measures.

In terms of practicality and costs, then, the most viable measure of overall wellbeing for use in non-market goods valuation is likely to be the type of global life satisfaction question that is included in large national datasets like the BHPS.

12. Finally there is evidence that the general public also favours evaluative measures of wellbeing, like life satisfaction, for policy making. In a study by Dolan and Metcalfe (2011) 1,082 members of the UK public were asked a series of questions related to the importance to them of different measures of welfare. In terms of importance to people's own lives and to government resource allocation decisions SWB was clearly far more important than preference satisfaction or objective wellbeing measures. People were then asked to choose between evaluative, experienced and eudemoinic measures of SWB and in terms of government-level resource decisions evaluative wellbeing or life satisfaction came out on top. Although the sample was self-selected it was broadly representative of the UK population and hence provides support for using life satisfaction in policy evaluation and decisions. And as we work more with such measures we will surely start to see the Heisenberg principle at work – what we as society measure will influence what we seek and value (Diener and Seligman, 2004; Dolan and White, 2007).

This type of evidence is supported by Ng (2003) who claims that welfare economics is too narrow in its focus on preference since what we care about ultimately is welfare and happiness. He states that "happiness is more ultimate than preference" (2003. p.309) and that the most important question for public policies is whether they increase happiness. Similarly, Diener et al (1999) argue that social (or objective) indicators are not sufficient on their own and that policy should be based on people's subjective experiences. And similar claims have some history in economics (for a full discussion see Ng, 2003).

There are a number of high-profile proponents of evaluative measures of wellbeing such as life satisfaction. Sumner (1996) places evaluative measures at the centre of his account of wellbeing; for Sumner life goes well for someone if they have a positive attitude towards their life, encompassing both a cognitive and an affective component. Hedonic measures are problematic because they are too narrow with their focus on mental states. Sumner sets out a list of criteria for a valid measure of welfare, evidence of which I have covered above. Diener et al. (2009. ch.7. p.11) claim that if we want the broadest level of assessment of welfare, then evaluative measures "may provide the best conclusions".

In this section I have set out arguments and a strong case for the use of subjective wellbeing measures and more specifically life satisfaction in public policy and non-market valuation. These arguments would garner the support of many wellbeing scholars as I have shown above, but there will undoubtedly be those who voice significant concerns about the approach that I am taking here: Daniel Kahneman and Paul Dolan come to mind. But the important thing to note, as shall be highlighted in the next sections is that actually the wellbeing valuation approach does not rely on the robustness or validity of life satisfaction as a measure of human welfare. Indeed, policy analysis and valuation can be undertaken with any SWB metric; the wellbeing valuation methodology I set out in the next sections could use evaluative, experienced or eudemonic measures of SWB. The main issue, therefore, is more the acceptance of SWB as a general measure for public policy and there is plenty of support for this in the UK and in many other countries.

The thesis from here on focuses on the methodology behind wellbeing valuation and the interpretation of wellbeing values. The focus is on developing a new methodology that allows us to use wellbeing valuation to derive value estimates that are in line with the economic theory of CS and ES set out in Chapter 2. This general methodology can be used with any measure of SWB, but in what follows I shall use life satisfaction as the base SWB measure without any further caveats or defence.

3.4. Summary

SWB data is being increasingly used in economics and in policy analysis and evaluation. I set out a brief introduction to the WV method and argue that the WV method should be seen as distinct to preference-based methods and as such we should not compare values derived using WV against those from preference-based methods such as stated preference and revealed preference valuation methods.

I provide a number of reasons and arguments for using WV to value non-market goods and services and also provide a defence of life satisfaction, the key SWB measure used in the WV approach. The next chapter builds a new approach to WV

that seeks to address and solve for the key technical problems associated with the current WV methodology.

Chapter 4

4. A new approach to wellbeing valuation

4.1. Introduction

Chapter 3 of this thesis set out the rationale for wellbeing valuation, provided a brief introduction to the methodology and discussed the main pros and cons of the method. Chapters 4 and 5 represent the main contributions of this thesis to the literature on wellbeing valuation. The five main contributions I make in the following two chapters are as follows:

i. **A new framework for assessing the validity of wellbeing valuation (Chapter 4).**

ii. **A full theoretical exposition of the wellbeing valuation approach.** The literature to date has not adequately shown the conditions under which wellbeing valuation can provide theoretically-consistent measures of welfare change. I provide the first full theoretical exposition of the wellbeing valuation approach **(Chapter 4)**.

iii. **A detailed critical assessment of the current wellbeing valuation methodology.** Since the literature has not provided a complete theoretical exposition of wellbeing valuation it has not been possible to critique the current methods in full. I will discuss the main technical problems associated with the current methods and what this may mean in terms of biases in the current results **(Chapter 4)**.

iv. **A new methodology for wellbeing valuation.** The new method provides a framework for estimating theoretically-consistent measures of welfare change using wellbeing valuation **(Chapter 5)**.

v. **A full interpretation of the values estimated using the wellbeing valuation approach.** The literature to date has been pretty silent on this issue and where it has been discussed there have been a number of inaccuracies **(Chapter 5)**.

In Chapter 5 I will showcase the new wellbeing valuation methodology with a case study applied to valuing non-pecuniary employment outcomes. The discussion from here will take life satisfaction as the given measure, but as discussed any other measure of SWB can be substituted in place of life satisfaction.

4.2. Assessing the validity of wellbeing valuation

The key question with regards to the validity and robustness of the wellbeing valuation approach relates to the extent with which wellbeing valuation derives robust measures of welfare change as set out in economic theory. This is the task of all valuation methods in economics. In this respect, Luechinger and Raschky (2009) and Frey et al. (2009) set out a list of criteria for robust wellbeing valuation. I have put these criteria into broader categories and further developed them as I felt they were not comprehensive enough.

4.2.1. Validity criteria for wellbeing valuation

Criterion A: Construct validity

The measure of SWB used in wellbeing valuation must be a valid measure of welfare - both in terms of the normative foundations as well as technical issues related to measurement error, such that it is a true reflection of how our lives are going (Luechinger and Raschky, 2009; Frey et al., 2009).

Criterion B: Scaling of wellbeing scores

SWB scores must be interpersonally comparable and for the purpose of statistical methodology we need to determine whether life satisfaction is ordinal or cardinal in nature (Luechinger and Raschky, 2009; Frey et al., 2009).

In addition to these two criteria the following stipulations that I have developed in this thesis are required and they are areas I will cover in some detail:

Criterion C: Technical validity

The statistical methodology employed in wellbeing valuation must be capable of estimating compensating and equivalent measures of welfare change in line with economic theory (as set out in section 2.2.). A key requirement here is to estimate causal effects of the non-market good and income on wellbeing.

Criterion D: Interpretation

The values derived from wellbeing valuation must be interpreted correctly in terms of their normative meaning as well as any technical caveats. This will allow for meaningful comparisons of the values against values from preference-based methods and for a meaningful interpretation of the results from evaluation frameworks that use the values, such as CBA.

With these two additional criteria we can agree with Luechinger and Raschky, (2009. p.622) that if these requirements are met life satisfaction measures and the general wellbeing valuation methodology can be used to value non-market goods. We will be able to derive theoretically-consistent measures of welfare change with a robust interpretation for use in policy evaluation. Criteria (A), (B) and (C) ensure that the values derived from wellbeing valuation are robust and theoretically-consistent, and criterion (D) ensures that the right interpretation is made.

This thesis assesses each of these criteria, but the main focus and contribution is towards criteria (C) and (D). The discussions related to criteria (A) and (B) are based on a review and assessment of the previous literature in this field. I have dealt with issues related to construct validity (A) in depth in Chapter 3. The summary from that discussion is that there are arguments and evidence both in favour and against the construct validity of life satisfaction, the primary measure of SWB used in wellbeing valuation. I have provided a strong defence for life satisfaction and would argue that the counter evidence is certainly not strong enough to dismiss the role of life satisfaction and other evaluative measures of SWB in wellbeing valuation. And indeed, as discussed, if there were a preference for a different measure of SWB, the

new theoretical approach set out in this thesis could equally be applied to that measure. I will, therefore, set aside issues related to construct validity for the rest of this thesis and use life satisfaction measures in my exposition of the wellbeing valuation approach. The next sections address criteria (B), (C) and (D).

4.3. Scaling of wellbeing scores (Criterion (B))

4.3.1. Ordinality versus cardinality

There is some discussion in the wellbeing literature on the cardinality of SWB scores, which is to say whether a given change in SWB scores, say a one index point increase, represents the same psychological impact along the whole length of the scale. In other words, is the change in life satisfaction from 2 to 3 equivalent in psychological or emotional terms as a change from 6 to 7?

Psychologists and sociologists have tended to be happy to assume cardinality and use methods like ordinary least squares (OLS) regression when analysing SWB data. Economists, on the other hand, have been more tentative and many papers have used ordinal models such as ordered probits (e.g. Blanchflower and Oswald, 2000; Clark and Oswald, 1994; van Praag et al., 2000; Mcbride, 2001; Tsurumi and Managi, 2016; Aoshima et al., 2018; Mendoza et al., 2019). This issue is not so important for wellbeing valuation for two reasons. First, marginal rates of substitution can be measured from ordered response models (Luechinger and Raschky, 2009). Second, anyway running life satisfaction models with ordinal and cardinal models produces near identical results both in terms of the ranking of life satisfaction determinants in order of effect size and in terms of the actual magnitude of the coefficients (Luechinger and Raschky, 2009). For these reasons most wellbeing valuation studies have used OLS models under the assumption of cardinality and I shall follow this trend here.

4.3.2. Interpersonal comparability

Here we are interested in whether similar SWB scores across individuals reflect similar life circumstances and levels of welfare. More formally, interpersonal comparability allows us to determine relationships such as $SWB_i > SWB_j$ for two or more different individuals. Following Robbins, it became the norm or fashion in economics to eschew the notion of interpersonal comparability of utility or welfare (Hammond, 1991; Ferrer-i-Carbonell, 2002), which left economists limited to identifying Pareto efficient outcomes or improvements for the purpose of policy analysis.

The issue of interpersonal comparability of utility in terms of the preference satisfaction account of welfare has been discussed at length by Hammond (1991) and Harsanyi (1955). In terms of wellbeing valuation (and wellbeing analysis more generally) the issue of interpersonal comparability is nicely described by Gilbert (2007). Gilbert (2007. p.47, 50, 52) states that SWB may become interpersonally *incomparable* due to the uniqueness of our previous experiences, which leads to what he calls the '*language-squishing hypothesis*' or the '*experience-stretching hypothesis*'.

Language-squishing is where impoverished experiences or histories force people to rate very highly experiences that to other (more fortunate) people would only represent very mediocre experiences. In this case *i*'s 9 out of 10 only represents a 4 out of 10 for *j*. An impoverished experiential background may also lead to experience-stretching, whereby *i*'s 9 out of 10 has the same psychological magnitude as *j*'s 9 out of 10, but *i* reports a 9 for the event of merely eating a piece of cake, whereas *j* reports a 9 for having won a nobel prize. It is a case of *i* being happy because he does not know what he is missing out on.

Which of these hypotheses is correct? Gilbert does not go on to say, and instead he makes the important conclusion that "all claims of happiness are claims from someone's *point of view* ..... whose unique collection of past experiences serves as a context, a lens, a background for her evaluation of her current experience" (2007. p.52, 53).

Gilbert (2007) suggests that we will probably never really know whether SWB and happiness ratings are comparable across two different people, but actually when we use large datasets, as we do in wellbeing valuation and in wellbeing analysis more generally, the issue of interpersonal comparability becomes less problematic anyway (Frey et al., 2009; Luechinger and Raschky, 2009; Di Tella and MacCulloch, 2006; Gilbert, 2007) and as we discuss below, there is growing evidence that life satisfaction and SWB ratings are comparable across people. There are also those who take a more theoretical or normative approach. An early example is Edgeworth who claimed that wellbeing in the form of (hedonic) pleasures is commensurable across different types of pleasure and across people (Bruni and Sugden, 2007). Ng (1997) is a more recent example of this type of argument.

In terms of empirical arguments, following Di Tella and MacCulloch (2006), if we frame the issue of interpersonal comparability in terms of differences in conversion factors/rates from psychological or emotional (wellbeing) states to numeric values across individuals (which is consistent with Gilbert's (2007) hypotheses), then we can see that because wellbeing valuation and wellbeing analysis more generally use group-level data - comparing the SWB of groups of individuals under different conditions – individual differences and personal peculiarities will tend to "counterbalance one another" (Frey et al., 2009. p.12; Di Tella and MacCulloch, 2006). Hence "the underlying assumption of a large part of happiness research in economics is that when people are measured in groups, the combination of their happiness scores does reveal useful information with which to make comparisons about social welfare" (Di Tella and MacCulloch, 2006. p.31-32). In other words, "we can be confident that if we ask enough people the same question, the average answer will be a roughly accurate index of the average experience" (Gilbert, 2007. p.70).

We note also that there is plenty of evidence in favour of interpersonal comparisons in self-reported wellbeing measures. Kahneman (2000) finds considerable convergence in affect ratings, especially pain scores across individuals in medical procedures. As discussed already, there is substantial agreement in wellbeing scores between self and third party reports – people are able to recognise the satisfaction levels of others (Diener et al., 1999; van Praag et al., 2003). There are correlations between self-reported satisfaction responses and (i) physiological measures (Davidson, 2004;

Kahneman, 2000) and (ii) objective circumstances (Easterlin, 2004). People from same language communities "have a common understanding of how to translate internal feelings into a number scale" (van Praag, 2003. p.34) and Van Praag (1991) finds that people translate verbal labels such as 'very good' or 'bad' on to roughly the same numerical scales. As van Praag et al. (2003. p.5) state "although it is very probable that what makes individuals happy or sad differs greatly amongst different cultures, it does seem as if there is a common human 'language' of satisfaction…" Generally, happiness researchers might see the growing data on SWB "filling the gap" for interpersonally comparable welfare data that economists longed for (Duncan, 2010. p.170).

The above discussion suggests that issues concerning the scaling of wellbeing scores are not grave enough to force us to dispose of the wellbeing valuation approach. An indeed in a comparative sense, it could be argued there is more evidence to suggest that SWB measures are more interpersonally comparable than are preference measures of welfare. The literature on the latter has traditionally been couched in theoretical terms such as Harsanyi's theory of extended preference (Hausman and McPherson, 2006), whereas there is a growing literature testing interpersonal comparability in empirical terms in the wellbeing literature. At best SWB measures such as life satisfaction are fully interpersonally comparable, and at worst they are at least as interpersonally comparable as are preferences. As far as the issue of interpersonal comparability goes, therefore, wellbeing valuation performs just as well, and if not better, than preference-based valuation methods. And since preference valuation approaches have been used extensively in policy analysis this would mean that wellbeing valuation as a practical approach for policy evaluation cannot be dismissed purely on the account of interpersonal comparability issues as it is a problem inherent to both approaches to valuation.

4.4. Technical validity (Criterion (C))

There are two aims in this section. I will first derive a theoretical approach to wellbeing valuation that is consistent with economic theory. And second, I will then assess how the current wellbeing valuation methods fare in respect to this (ideal)

theoretical approach. The contribution to the wellbeing valuation literature of this section is that this is the first time that a thorough and precise theoretical approach for wellbeing valuation is developed. And because of this I can also provide the first comprehensive critique of the current methodologies used in the wellbeing valuation literature. This will provide a full explanation of the main biases in the current studies. All of this has generally been missing from the wellbeing valuation literature to date.

This foundational work on the theory of wellbeing valuation will then provide the basis for a new methodology – *The Three-Step Wellbeing Valuation Approach* - for estimating welfare change using the wellbeing valuation method, the focus of Chapter 5 of this thesis.

4.4.1. Theory of measuring welfare change through subjective wellbeing data

As discussed in Chapter 3 in the wellbeing valuation approach the aim is to directly estimate the MRS between the non-market good (Q) and money (M) with an "observable" measure of welfare. We do this by estimating the utility function using SWB data. Let us start, using compensating surplus as an example, and define CS for Q using equation (6) (note that to account for the fact that there may be an indirect effect of Q on welfare through M the term in the right hand side of equation (6) has been slightly modified by adding a superscript 1 to the income variable (M)):

(6.1) $\quad U(Q^0, M^0) = U(Q^1, M^1 - CS)$

I will look at a positive welfare impact of $Q^0 \rightarrow Q^1$, which could be due to a positive change in the quantity or quality of Q. What is crucial to state here is that the welfare impact of the change in the non-market good ($Q^0 \rightarrow Q^1$) should clearly account for *all* of the possible impacts on welfare (Champ et al., 2003). Many non-market goods will foster both direct and indirect welfare impacts. For example, an environmental programme that protects a large forest area would create direct enjoyment for people using the area as well as impacts on other aspects of life that are instrumentally important for welfare, such as any health benefits that people may derive due to, say, improved air quality in the local area. Economic theory captures both the direct and

indirect effects in valuation (Champ et al., 2003) and generally speaking so do stated and revealed preference methods in practice (although stated preference can be used to ascertain the value associated with a single indirect benefit, such as health).

Equation (6.1) can be estimated empirically by substituting SWB equation (8) into (6.1):

(11) $SWB(Q^0, X^0, M^0) = SWB(Q^1, X^1, M^1 - CS)$

Equation (11) acknowledges that there may be indirect welfare impacts of the change in Q through M and X (eg, the policy may affect income and the vector X may contain factors like health), demonstrated by the respective changes $M^0 \rightarrow M^1$ and $X^0 \rightarrow X^1$.

Solving for CS by using first derivatives we get:

(12) $CS = (M^1 - M^0) + \frac{SWB'_X \cdot X'_Q (Q^1 - Q^0)}{SWB'_M} + \frac{SWB'_Q (Q^1 - Q^0)}{SWB'_M}$

In words this states that:

CS = (impact of Q on M) + (the MRS between income and the indirect effect of Q on SWB via X) + (the MRS between income and the direct effect of Q on SWB).

Naturally we must also acknowledge that M may also impact on SWB indirectly in (12) and so SWB'_M should represent the *total derivative* for income. In fact, the CS for a change in the non-market good (ΔQ) in equation (12) can be reformulated in terms of total derivatives for Q and M:

(13) $CS = -\frac{\mathrm{d}\, SWB}{\mathrm{d}\, Q} \cdot \Delta Q \Big/ \frac{\mathrm{d}\, SWB}{\mathrm{d}\, M}$

This simply represents the MRS between Q and M accounting for all of the impacts that Q and M have on SWB. In equation (13) the MRS represents the amount of

money taken away or received that will leave the agent in his initial welfare position following a change in Q from the status quo. It is an exact measure of compensating welfare change accounting (as it should – see Champ et al., 2003) for all of the impacts on welfare.

Equations (11) through to (13) set out the theoretically correct approach to measuring welfare change with SWB data. I demonstrated this using a compensating measure of welfare change, but equally it is possible to set out the approach under the format of equivalent welfare change measures as well.

It would be beneficial to derive some new terminology for the welfare change measures derived using WV in order to distinguish them from welfare change measures estimated in preference valuation methods. A number of different terms have been used in the WV literature. Many papers just tend to revert to WTP and WTA definitions for the values derived in WV (e.g., Luechinger, 2009; Kountiuris and Remoundou, 2011; Ambrey and Fleming, 2011; Menz and Welsch, 2012; Levinson, 2009; Frey et al, 2004; Ferreira and Moro, 2009; Ambrey and Fleming, 2014), which is not helpful for WV since these terms should only be employed with preference valuation methods (I discuss this in more detail below). Other terminology that has been used includes *income compensation values* (e.g. work by Frijters), which is problematic as WV is not restricted to estimating compensation values, and *income equivalence values* Carroll et al. (2009). A more accurate definition, which I will use in this thesis, I believe is as follows:

Compensating wellbeing value (CWV). *This is the amount of money, to be hypothetically deducted or provided, that will leave the agent in his/her <u>initial</u> SWB position <u>following</u> a change in the good.*

Equivalent wellbeing value (EWV). *This is the amount of money, to be hypothetically deducted or provided, that will leave the agent in his/her <u>subsequent</u> SWB position in <u>absence</u> of a change in the good.*

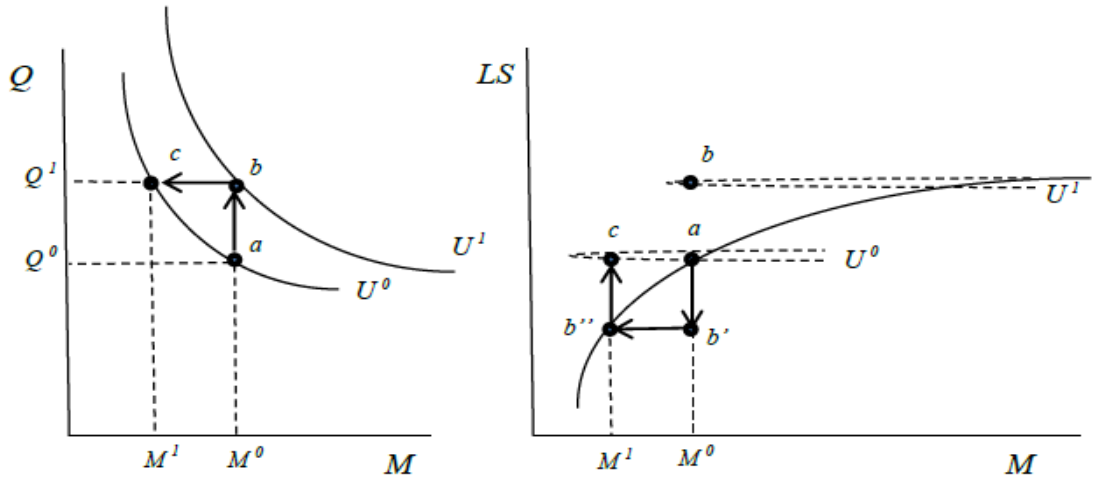
4.4.1.1. Measuring Compensating wellbeing value and Equivalent wellbeing value

Let us define an improvement as a non-market good and a deterioration as a non-market bad. A 'good' is something which leads to a welfare gain and a 'bad' is something that leads to a welfare loss. For 'goods' and 'bads' we can estimate both the CWV and the EWV, but here we shall just focus on CWV. This is because we can make any 'good' a 'bad' by restricting provision or making people forego it and so the discussion on CWV is generalisable to the case of EWV.

Non-market 'goods'

Panel (i) of Figure 3 shows the CWV of the 'good' *Q* using indifference curves that track SWB rather than utility. In other words the indifference curves (*U*) represent the level sets of an SWB function. Ceteris paribus the provision of (or improvement in) the 'good' from Q^0 to Q^1 moves the individual from a starting point of *a* to *b*. The CWV is $(M^0 - M^1)$, the amount of money that returns him back to the initial level of SWB (U^0) at point *c* (where *U* = SWB).

Figure 3. Graphical representation of the wellbeing valuation approach (for non-market goods).



Panel (i): Indifference curves **Panel (ii): Level sets of the SWB function**

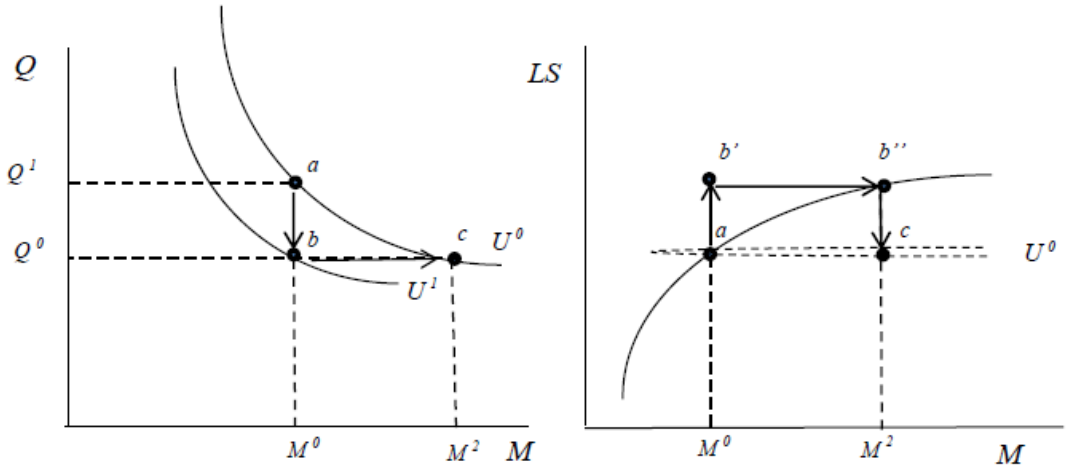
This is replicated in panel (ii), using life satisfaction (*LS*) as the SWB measure and setting it as a function of income, where there is a diminishing marginal utility of income as is standardly assumed in the WV literature and models.

In order to estimate the CWV for a 'good' we must move *down* the *LS* function; the individual starts at *a* and with the 'good' would move to point *b* but to measure this in terms of CWV we have to move the individual in the *opposite direction* from $a \rightarrow b' \rightarrow b'' \rightarrow c$ and by doing so $(M^0 - M^1)$ in panels (i) and (ii) are equal. At point *c* the individual has the 'good' but $(M^0 - M^1)$ less money and is on his same original level of life satisfaction at U^0.

Non-market 'bads'

Figure 4 shows the same process for a 'bad' – a ceteris paribus negative change in Q, which is shown by the change from Q^1 to Q^0. Here $(M^0 - M^2)$ shows the CWV for a non-market 'bad'.

Figure 4. Graphical representation of the wellbeing valuation approach (for non-market bads).



Panel (i): Indifference curves **Panel (ii): Level sets of the SWB function**

In order to estimate the CWV for a 'bad' we must move *up* the *LS* function; the individual starts at *a* and with the negative welfare impact from the change from Q^1 to Q^0 (representing the non-market 'bad') would move to point *b* but to measure this in terms of CWV we have to move the individual in the *opposite direction* in panel (ii) from $a \rightarrow b' \rightarrow b'' \rightarrow c$ and by doing so $(M^0 - M^2)$ in panels (i) and (ii) are equal. At point *c* the individual has suffered from the 'bad' but has more money $(M^0 - M^2)$ to compensate and is therefore on his same original level of life satisfaction at U^0.

It seems counterintuitive in panel (ii) of Figures 3 and 4 that we move in the counterintuitive direction along the LS function when estimating welfare change measures in WV. This is the case in order to account for the shape of the life satisfaction function with respect to income. If the relationship between SWB and income were linear then the CWV values would be the same regardless of which direction we were to move along the life satisfaction function, but because of the diminishing marginal utility of income (where the partial derivative of life satisfaction with respect to income is higher (steeper) at lower levels of income), the direction of travel along the non-linear life satisfaction curve does make a significant difference to the value estimates. This will especially be the case for large changes in SWB and there may be very little difference at smaller or marginal changes in SWB due to the non-market good/'bad'.

Setting out the processes of estimating welfare change measures in WV like this is important because it allows us to make a precise definition of the measure of welfare change and to measure it accurately. In the current WV literature it is usually not clear which exact measure of welfare change has been estimated (papers like Ferreira and Moro (2009) and Welsch and Kuhling (2009) are, however, exceptions to this general trend).

Looking at the mechanics behind valuation in WV we can see that there are three effects going on when we value non-market goods (Q) in the WV approach. First, the value depends on the magnitude of the impact of Q on SWB or life satisfaction (shown as $a \rightarrow b$ in panel (i)), second the value also depends on the strength of the relationship between life satisfaction and income (assumed constant in the Figures above), and third the value depends on the direction we move around the *LS* function in panel (ii).

A disparity in CWV values for a good versus a bad (which is famously demonstrated in the WTA-WTP disparity) can show up in one of two ways. First, the psychologically larger effect of a loss compared to an equivalent gain in Q would show up in a larger absolute impact of Q on SWB for losses, in other words, $|b - a|$ would be larger for the 'bad' which would render $|M^0 - M^2| > |M^0 - M^1|$.

The second way (which further magnifies this effect) is due to the diminishing marginal utility of income. We can see in Figures 3 and 4 that even for something that did not have a larger psychological effect when comparing losses to equivalent gains, we would still see a higher valuation for WTA related measures. The distance $|a - b|$ is drawn at about the same magnitude in both Figures, but $|M^0 - M^2|$ is clearly visually larger than $|M^0 - M^1|$ because from a given SWB starting point, giving money to individual is less impactful on *LS* than taking it away when the impact of income is estimated in logarithmic format in the SWB model. In sum, for a given Q, differences between EWV and CWV will emerge in this framework due to the curvature of the income function. For welfare gains, EWV > CWV and for welfare losses, CWV > EWV.

Table 2 provides the framework for estimating CWV and EWV in wellbeing valuation, where log of income is used in the income model, as in equation (10). The equations are set out for binary Q variables, but for continuous Q variables it would also be possible to use formats to reflect non-linear impacts on SWB.

Table 2. CWV and EWV in wellbeing valuation

	Compensating measure (CWV)	**Equivalent measure (EWV)**
Welfare gain	$CWV = M^0 - exp^{\left[\ln(M^0) - \frac{g'_Q}{f'_M}\right]}$	$EWV = exp^{\left[\frac{g'_Q}{f'_M} + \ln(M^0)\right]} - M^0$
Welfare loss	$CWV = exp^{\left[\frac{-g'_Q}{f'_M} + \ln(M^0)\right]} - M^0$	$EWV = M^0 - exp^{\left[\ln(M^0) + \frac{g'_Q}{f'_M}\right]}$

Notes: M^0 is initial income; f'_M is the effect of income on SWB; and g'_Q is the effect of the non-market good on SWB.

It is important to note that under this framework CWV for welfare gains and EWV for welfare losses are constrained at the level of an individual's income, whereas EWV for welfare gains and CWV for welfare losses have their limits at infinity as we would expect and as would be the case with WTP and WTA values respectively. To see this, for example, take the EWV for a welfare loss. Here g'_Q is negative and for Q with

very large negative impacts, such that $g'_Q \to -\infty$ it can be shown that EWV is constrained at the original level of income (M^0):

(21) $$EWV = M^0 - exp^{\left[\ln(M^0) + \frac{g'_Q}{f'_M}\right]} = M^0 - e^{-\infty} = M^0$$

4.4.2. Technical conditions of the wellbeing valuation approach

There are four key technical conditions that I develop here and which must be satisfied in order to estimate equation (13) correctly and to ensure technical validity (Criterion C).

CONDITION 1: FULL IMPACT

The full (direct + indirect) effects of the non-market good and of income on SWB must be accounted for and measured

If this condition is not satisfied we cannot derive the full welfare change attributable to the non-market good (Champ et al., 2003).

CONDITION 2: CAUSAL ESTIMATORS

The estimated effects of the non-market good and of income on SWB must be unbiased

This is a clearly imperative condition for wellbeing valuation. It means that the total derivatives in equation (13) must have a full causal interpretation. This is an implicit assumption underlying the definition of a total derivative and of course it is not possible to derive welfare change measures without knowing how the non-market good and money impact causally on welfare.

Note that this condition is separate and independent to Condition 1. Condition 2 states that the statistical estimates must be unbiased. In regression analyses this could be an unbiased partial impact. In addition to this, Condition 1 requires that the estimate be a full rather than partial effect. Together, Conditions 1 and 2 therefore stipulate that the estimates represent full and unbiased causal effects.

CONDITION 3: SAMPLE MATCHING

The total causal derivatives in equation (13) must come from the same population group so that they are comparable

The theory of valuation in economics is an individual-level theory, which defines welfare change in terms of impacts on the individual – that is, how the non-market good and income impact the individual's level of utility. The implication for wellbeing valuation, where statistical analysis is run using group level data (because it is not possible to estimate SWB functions for a single individual), is that the total derivatives for Q and M clearly have to come from the same group or be representative estimates for the same group.

This is to avoid the problem of estimating the impact of a non-market good for one group and estimating the amount of money required to have the same effect for a different group. Clearly if the second group values income in a different way to the first group we will have incorrect estimates of value in wellbeing valuation.

CONDITION 4: CLEAR INTERPRETATION

The total causal derivative for the non-market good ($^{\mathrm{d}\,SWB}/_{\mathrm{d}\,Q}$) should have a clear interpretation where there are heterogeneous treatment effects of Q, so that results are useful for policy

If there are differential effects of Q on SWB across the population it is critical to understand to whom the total derivative and welfare change estimates are applicable. To understand this we need to know whether there are heterogeneous treatment effects of Q on SWB due to selection on gains factors. This is where certain individuals select into 'using' the non-market good because they disproportionately benefit from it.

The rationale and details behind these four conditions will be discussed at length below.

4.4.3. Assessing the current wellbeing valuation methodology

4.4.3.1. Current wellbeing valuation methodology

The wellbeing valuation approach is an emerging method with approximately 100 publications since its inception in 2002. The literature has tended to grow by looking at the value of different outcomes and services using the method rather than focus on technical developments. By and large the WV literature to date has predominantly used fairly simple multivariate regression analysis following the approach as discussed in section 3.3. Equation (8) is estimated empirically using the following type of single-equation model:

(9) $$SWB_i = \alpha + \beta_1 Q_i + \beta_2 \ln(M_i) + \beta_3 X_i + \varepsilon_i$$

where income is in logarithmic format to account for the diminishing marginal utility of income and where SWB is usually life satisfaction. As discussed above these models have usually been run assuming cardinality using OLS regression. The technical critiques of the current wellbeing valuation methodologies provided here are applicable for any measure of SWB that may be used in equation (9).

Partial derivatives from the single equation model in (9) are used to estimate the value (here CS) of Q as follows by solving for CS (eg, Frey et al., 2009) (I have dropped the constant term and error terms that feature in both sides of the equality):

(14) $$SWB_i(\beta_1 Q_i^0 + \beta_2 \ln(M_i^0) + \beta_3 X_i^0) = SWB_i(\beta_1 Q_i^1 + \beta_2 \ln(M_i^1 - CS) + \beta_3 X_i^1)$$

(15) $$CS = M^0 - exp^{\left[\ln(M^0) - \frac{\beta_1 (Q^1 - Q^0)}{\beta_2}\right]}$$

The problem is that generally speaking the approaches used to date do not adhere to the four technical criteria/conditions set out above and hence lead to biased estimates of the value of Q. I will focus on each of these issues in turn next.

4.4.3.2. The current wellbeing valuation methodology vis-à-vis the technical criteria

4.4.3.2.1. Full impact (Condition 1)

If we adjust equation (12) to account for a logarithmic format for income, we see that equation (14) does not estimate the correct measure of welfare change (here CS):

$$(16) \qquad CS = M^0 - exp\left[\ln(M^0) - \frac{\frac{\mathrm{d}\,SWB}{\mathrm{d}\,Q}(Q^1 - Q^0)}{\frac{\mathrm{d}\,SWB}{\mathrm{d}\,M}}\right] \neq M^0 - exp\left[\ln(M^0) - \frac{\beta_1(Q^1 - Q^0)}{\beta_2}\right]$$

This is because generally ${}^{\mathrm{d}\,SWB}/_{\mathrm{d}\,Q} \neq \beta_1$, and ${}^{\mathrm{d}\,SWB}/_{\mathrm{d}\,M} \neq \beta_2$ as single-equation models with a number of control variables cannot simultaneously estimate two total derivatives. In order to derive total derivatives for M and Q in a single-equation structure all other covariates must be measured pre-treatment (ie, before changes in M and Q) and this will be near-impossible to guarantee especially when using observational data sets such as those used in WV. Also, of course, we can only estimate *one* total derivative in a single-equation model because M and Q cannot both precede each other simultaneously. The upshot is that these parametric restrictions in single-equation models mean that some important mediators of the effects of M and Q on SWB are likely to be controlled for in X, and hence we cannot account for some of the indirect effects. One key mediator variable will be health, for which data is often collected at the time of the survey meaning that the health response will be capturing effects of M and Q on health.

One possible solution is to move away from the single-equation framework and estimate structural equation models, where relationships between M and Q and other covariates are modelled explicitly in auxiliary models and this has been an approach taken by a few papers. Indeed the indirect effects issue has been well-documented in the WV literature (Stutzer and Frey, 2004b; Adler and Posner, 2008) and there have been a number of attempts to include some of the indirect effects of the non-market good in the valuation (for example Welsch and Kühling, 2009; Carroll et al., 2009; Groot and van den Brink, 2006; Welsch, 2002; Ferrer-i-Carbonell and van Praag, 2002; Rehdanz and Maddison, 2005; Welsch, 2008a; Welsch, 2008b). Predominantly these studies have been concerned with the indirect impact of the non-market good on SWB via income only. Broadly the two types of methodology employed have been either to drop the income variable from the regression model entirely (eg, Carroll et

al., 2009; Welsch, 2002; Groot and van den Brink, 2006; Welsch, 2008a), or to run an auxilary model whereby the relationship between income and the non-market good is estimated explicitly (eg, Welsch, 2008b; Ferrer-i-Carbonell and van Praag, 2002; Groot et al., 2004). For example, this entails running the following models:

(8) $SWB(Q, M, X)$

and

(17) $M(Q, T)$

where T is a vector of other determinants of income. The value of Q is then based on the product of partial derivatives from (8) and (17): $[SWB'_Q + (SWB'_M \cdot M'_Q)]$, which respectively pick up the direct impact of Q on SWB and the indirect impact on SWB through M.

Ferrer-i-Carbonell and van Praag (2002) develop on this auxilary model approach to assess the indirect impacts of different health conditions on SWB via six domain satisfactions, such as job satisfaction, health satisfaction and leisure satisfaction. In other words, health is deemed to impact on these domains which in turn impact on overall life satisfaction and a value is attached to the sum of the direct effects of health and these indirect impacts of health.

These are clearly steps in the right direction because they recognise the issue of indirect effects, but they are problematic for a number of reasons. First, dropping an important variable like income from an SWB regression will further exacerbate any bias that exists in the coefficient on Q. Second, as Ferrer-i-Carbonell and van Praag (2002) state the non-market good could impact on SWB through more than one channel. Third, structural equation model approaches like the one set out in equations (8) and (17) will be using and mixing two (or more) biased estimates of the impacts of Q unless Q is exogenous in both models, which is highly unlikely. And finally, these methods focus on the non-market good and do not attempt to derive the total derivative for income, which is also essential to WV.

4.4.3.2.2. Causal estimators (Condition 2)

It is rare that we can estimate unbiased causal effects with observational data and in the WV literature (and also the wider SWB literature more generally) it is well-documented that bias can arise from endogeneity, simultaneity and measurement error (Pischke, 2010; Frijters et al., 2011; Ferreira and Moro, 2009; Frey et al. 2004; Saris, 2000; Ambrey and Fleming, 2011; Luechinger, 2009). OLS is the predominant estimator used in WV, but it is likely to produce biased causal estimates for Q and M because it relies on a selection on observables assumption which for the most part will not hold. There have been three general types of approaches used in the WV and general SWB literature to address this.

1. Fixed effects models with panel data. This method uses within-person variation to control for time-invariant unobservable variables, which in the context of an SWB model could be something like the individual's underlying preferences or personality traits. The problem with fixed effects models are threefold: (i) the approach cannot control for time-varying unobservable factors; (ii) factors that have little variation over time within individuals, which may causally impact on wellbeing, such as employment and marital status may wrongly be found to have no statistically significant effect. This means that some important non-market outcomes may be assumed to have a zero value although they do impact on SWB; (iii) fixed effects models do not eradicate the problem of measurement error and in fact, fixed effects can exacerbate problems here by increasing the ratio of measurement error to actual variation in variables that are measured with error (Deaton, 1993). Thus, it is unlikely that fixed effects methods will provide unbiased causal estimates for Q and M.

2. Exogenous variables models. A number of papers have used theoretically exogenous or seemingly exogenous variables for income and the non-market good. For example, in their valuation of climate variability Alem and Colmer (2013) use exogenous shocks in weather conditions. Kuehnle and Wunder (2014) claim that the variable they use for the impacts of daylight time saving transitions is exogenous, or at least conditionally so. Ambrey and Fleming (2014) use income windfalls as exogenous shocks for the income variable - this is similar to Lachowska's (2017) study of the impact of tax rebates on life satisfaction (although Lachowska did not use these results to value any non-market goods). These exogenous variables are inputted

directly into the (single-equation) SWB function rather than being employed as instrumental variables.

There are a number of problems with these papers and methods. First in actuality exogeneity tests in these studies tended to show that the variables were not truly exogenous as they were correlated with some other factors. Second, exogenous changes in income due to windfalls is on the face of it a seemingly plausible method. In wellbeing valuation Ambrey and Fleming (2014) use windfalls in income and in more general wellbeing analysis Gardner and Oswald (2007) use lottery wins. In both papers the income variable concerned is inputted directly into the wellbeing function (a mental health function in Gardner and Oswald (2007)). There is a risk here that the income variable is still in a sense 'biased' because although it could be argued that the variable is exogenous, windfalls such as lottery wins and inheritance capture not just a change in income for the individual but also an impact on their emotions (euphoria at winning the lottery or sadness at having a relative pass by) and so the coefficient on income in such models will not necessarily represent the causal effect of income alone. Gardner and Oswald's (2007) solution to this is to compare lottery winners of different win sizes (where the full win size was small to medium), rather than comparing lottery winners to non-winners. This method is one that I replicate below in my own analysis. Third, for this type of method to provide a solution to the causality problems, the study would need to employ exogenous variables for *both* income and the non-market good, which none of these studies have done. Fourth, as I will discuss in more detail below (in reference to instrumental variable methods), even if exogenous variables were used for both income and the non-market good unbiased estimates of welfare change could not be estimated from single-equation models like the ones used in these studies if income and the non-market good are correlated. If the two variables are correlated then including both of them in a single model is likely to run into the indirect effects problem for one of the variables.

3. Instrumental variable (IV) models. A number of papers instrument for income (e.g. Helliwell and Huang, 2005; Ferreira and Moro, 2009; Marsh and Bertranou, 2012; Chandoevwit and Thampanishvong, 2016; Howley, 2016; Huang et al., 2018; Mendoza et al., 2019; Mahasuweerachai and Pangjai, 2018), or for the non-market good (Danzer and Danzer, 2011; Aoshima et al., 2018) and some for both income and

the non-market good in 2SLS (e.g. Luechinger, 2009; Tsurumi and Managi, 2015). But this does not provide a full solution for a number of reasons and before discussing these it is important to set out the main assumptions that underpin IV methods.

IV uses exogenous or conditionally exogenous variation in one or more variables to 'force' exogenous variation in the main variable of interest (ie, Q or M). A common way of using IVs is in two stage least squares (2SLS) as follows:

(17) $D = \tau + \rho Z + \mu$ (First stage)

(18) $SWB = \gamma + \alpha \widehat{D} + \varepsilon$ (Second stage)

Where D is the variable of interest (ie, the treatment), which could be Q in this example; Z is the IV (assumed to be a binary variable here); and α is the causal effect of D on SWB. The key assumptions are,

(i) Independence of the instrument and exclusion restriction: $(SWB_0, SWB_1, D_0, D_1) \perp Z$

(ii) Monotinicity (no defiers): $D_1 \geq D_0$

(iii) First stage variation: $\rho > 0$

(iv) Homogenous effects of D: α is constant for all units i

Item (ii) is an identifying assumption that needs to be made as we cannot observe complier type in the data, so I will assume that this is appropriate as is customarily done in the econometrics literature. Item (iii) can be tested in the first stage regression. Therefore, the main assumptions of interest are (i) and (iv). Assumption (i) states that the instrument is exogenous and that it only impacts on the outcome (SWB) through D. These are untestable assumptions, but I note that the exogeneity element can be satisfied through use of other covariates in the models, which then makes the assumption one of conditional independence and exogeneity. Assumption (iv) claims that the impact of D on SWB is the same throughout the sample.

I note first that the theoretical arguments behind income instruments tend not to be fully validated in the WV literature. Commonly used instruments for income in WV

include spouse's income, spouse's employment status, house ownership, predicted industry wage levels, local area wage levels, date of survey interview, age of the respondent, and social class (Luttmer, 2005; Ferreira and Moro, 2009; Luechinger, 2009; Pischke, 2010; Chandoevwit and Thampanishvong, 2016; Howley, 2016; Mendoza et al., 2019). These instruments are unlikely to be truly independent of the potential treatment (here income) and wellbeing – respectively D and SWB from equations (17) and (18) - because none are true exogenous shocks in Z.

Other studies in the wider wellbeing literature have used as instruments sight of payslips (Powdthavee, 2010) and father's years of education (Knight et al., 2009). The problem with payslips as an IV is that it is unlikely to adhere to assumption (iii) and it is hard to defend parents' education as an independent instrument that adheres to the exclusion restriction, because parents' education is likely to impact on the child's wellbeing through more than just the income of the child in adulthood.

Generally, the literature has found that compared to estimates from OLS, using instruments for income significantly increases the size of the coefficient on the income variable in wellbeing regressions by a scale of up to around 10-12 times the OLS estimates (Levinson, 2012; Luttmer, 2005; Powdthavee, 2010; Mahasuweerachai and Pangjai, 2018).

A promising instrument that has been used for income is lottery wins. After controlling for number of times one plays the lottery the instrument should be independent of D and SWB. Lottery wins will also clearly have an effect on income (assumption (iii)) and if the sample of lottery players is used we can assume that the exclusion restriction also holds. The potential issue around exclusion with the lottery wins IV is that in addition to increasing income, it is fair to argue that the lottery win itself would impact directly on the happiness and wellbeing of the individual. But here we can compare big prize lottery winners to smaller prize winners, rather than lottery winners to non-winners, such that both groups (ie, Z=1 and Z=0) experience euphoria at winning the lottery such that within the sample of lottery players there is no separate effect (separate to the income effect) of the lottery win on SWB.

However, lottery wins have not been employed in the WV literature to estimate monetary values and as I shall argue below the independence assumption may not have been fully met in the lottery wins literature because lottery playing frequency has not been controlled for in the first stage.

There are also further issues with 2SLS frameworks such as the one set out in (17) and (18) for the specific task of WV. First, even with perfect instruments for income and the non-market good 2SLS is problematic because the single-equation framework in the second stage does not allow us to derive total derivatives for the non-market good and income. Instrumenting for income, for the non-market good, or for both income and the non-market good gets us a better handle on causality but we cannot estimate total derivatives unless all of the other controls in the second stage of 2SLS are measured pre-treatment *and* income and the non-market good are orthogonal, which is probably unlikely; if M and Q are correlated, then one of them has to be measured before the other in order to avoid the problem of indirect effects, which means that it is impossible to estimate both total derivatives (for income and the non-market good) in the same model (equation (18)). A case in point that exemplifies this problem is Brown's (2015) analysis of health values. Both health and income are instrumented but they are heavily correlated with each other – better health leads to higher income and vice-versa and so in the second stage of 2SLS we are unable to derive the total derivatives of SWB with respect to income and health.

The second issue relates to the fact that assumption (iv) is problematic in 2SLS. I will discuss this issue in detail in the next section when I address the topic of sample matching.

Before closing this section we can make some hypotheses about the direction of bias introduced by the problems related to indirect effects and causality of the income and the non-market good variables. A-priori statements about the direction or magnitude of bias related to the non-market good (Q) are hard to make, but we can argue that the income variable is likely to be biased downwards in OLS regressions of the type in equation (9) for a number of reasons. First, it is well known that income is measured with error which creates a downwards bias in OLS. Second, the indirect effects of income are likely to be positive (eg, through positive effects on health) and hence

standard OLS SWB models will not produce estimates of the full effect of income on SWB. Third, since earning more money comes with additional stresses, work commitments and time lost for other meaningful pursuits such as spending time with the family, simply looking at salary related income or wealth increases (which nearly all SWB papers do) will dilute the effect of income on wellbeing in comparison to studies that use exogenous changes in income. For example, putting to one side the issue of the euphoria of winning a lottery, a medium-sized lottery win that effectively translates into a 50% increase in income for the year will have a much different qualitative effect on SWB in comparison to the same salary rise that is due to a promotion which entails more responsibility and harder work.

A downward biased income coefficient in OLS (β_2) will ceteris paribus lead to an upward bias in values estimated using WV because β_2 is the denominator in the value calculations in WV (see equations (14) or (15)). And this is supported by the evidence as discussed above. (Clark and Oswald, 2002; Powdthavee, 2008; Frey et al., 2009; Ferreira and Moro, 2009; Levinson, 2012). Although there are a couple of exceptions to the rule, (e.g. van den Berg and Ferrer-i-Carbonell (2007) find that results from WV and values from contingent valuation were very similar for informal care-giving and Cohen (2008) finds similar results for valuations of reductions in crime rates using the two approaches (however, it should be noted that Cohen made a number of significant simplifying assumptions about how results from CV crime studies can be aggregated)), WV values have generally been found to be magnitudes higher than values derived from RP and SP methods (Levinson, 2012; Luechinger, 2009). Frey et al. (2004) set out the robust estimation of the causal effect of income on SWB as a priority area of research for WV.

4.4.3.2.3. Sample matching and clear interpretation (Conditions 3 and 4)

I will address these issues together as they can surface from a common problem: sample matching and treatment effects interpretation become problematic issues when we acknowledge heterogenous treatment effects. The majority of WV papers focus on a binary Q, for example, being employed, being healthy, living in a safe or polluted area etc, and I shall focus the discussion here on binary variables for Q. If the impacts of Q on SWB differ across different population groups then it is essential that (i) our

estimates of the impacts of M and Q are representative of the same sub-population and that (ii) the estimated effect of the non-market good has a meaningful interpretation for policy - in other words, we would like to know whether our estimated value of Q is based on the average effect of Q on SWB across the population or on the average effect for the treated, the non-treated and so on.

In addition to the problems already discussed, this creates further issues for the use of both OLS and 2SLS estimators in WV. OLS provides poorly-defined treatment estimators for Q, that lie somewhere between the average treatment effect for the treated (ATT) and the average treatment effect for the non-treated (ATNT), and how close the estimator lies to the ATT or the ATNT will depend on the proportion of treated and non-treated groups in the sample (Humphreys, 2009). This means that (in addition to being biased) values based on OLS estimates also have no concrete implications for policy because we cannot know whether they signify the value generated by those who were treated, or the value that would be associated with an intervention that impacts on people who otherwise would not participate. If $^{\mathrm{d}\,SWB}/_{\mathrm{d}\,Q}$ were estimated as the ATT, the monetary value would represent the *retrospective* value of Q for those that were treated. The ATNT would tell us something about how valuable it would be if a policy (concerning some non-market good) were rolled out to those who were not initially treated. This would represent the *prospective* value of the policy or non-market good. And the ATE would give us a broad estimate of value for anyone picked from the general population.

It is, therefore, vital that in addition to issues related to causality, or internal validity, estimates used in WV have a clear treatment effect interpretation for the purposes of inputting into policy. This is the separate issue of external validity. An important outcome for the discussion here relates to the interpretation of WV values. I discuss interpretation issues in detail in Chapter 5, but will add some further commentary here in light of the treatment effects discussion. The ability to derive values based on different treatment effects for the non-market good in WV puts it in a unique position in respect to values associated with non-users. The valuation literature highlights two types of value:

i. **Use value** relates to actual use, planned use or possible use of the good.

ii. **Non-use value** relates to the attachment of value to a good although there is no actual, planned or possible use. There are three types of non-use value: (i) existence value, (ii) altruistic value, and (iii) bequest value.

It is possible for users to hold both use and non-use values, whereas non-users will by definition only hold a non-use value for the non-market good. Revealed preference measures are only able to derive use values. Stated preference methods can be used to derive both use and non-use values. In WV if the value is estimated from the ATT for Q it will represent the users of Q. The ATNT is a special case unique to WV. The ATNT for Q will represent the impact of Q on SWB for people who do not use Q, ie, the non-users. A value derived from the ATNT for Q, therefore, represents the use value of Q for non-users, if they were to use Q. This is a value estimator that is unique to WV and which has some important implications for policy. It is unique because it cannot be elicited in preference methods. It may seem similar to the case of when a stated preference survey asks respondents for the value regarding some future non-market good (eg, the value of proposed improved public amenities at a national park). In this case all respondents are non-users because the amenities have not been developed/improved yet, but it is not possible to elicit use values from non-users even in this case because in such surveys non-users will state a value based on their expected non-use of the services, hence non-users will simply state their predicted non-use values. This is different to the use value for non-users should they come to use the non-market services, which is what the WV method can derive.

Use values for non-users is not a trivial oxymoron – it is relevant where governments and other organisations may seek to change the behaviour of individuals such that they are encouraged to consume the non-market good (eg, an environmental awareness programme that encourages non-users to use the national park and its amenities to develop a better understanding and connection with nature and the environment). A standard use value would not suffice to estimate the value of this programme here since users are different to initial non-users who are encouraged to go. Reverting back to the treatment effects literature we can assume users to 'select'

into using the national park based on some level of expected gains to their welfare. This means that in all likelihood they will more strongly prefer the national park than the initial non-users do even if the latter are encouraged to go to the park. The ATT for the park will be higher than the ATNT for the park, which is another way of saying that use values for the park will be higher for users than for non-users who subsequently make use of the park. Only WV can estimate potential use values for initial non-users.

In terms of sample matching heterogenous effects have major implications for 2SLS estimates. Generally speaking assumption (iv) (α is constant for all units i) is not true in 2SLS. 2SLS derives estimates for a localised sub-sample of the population known as compliers to the instrument. This has been termed the local average treatment effect (LATE) by Angrist and colleagues (e.g. Angrist and Pischke, 2009). The LATE is the effect of some variable (e.g. the non-market good, Q) for people whose behaviour complies to the instrument.

The LATE creates problems in WV because it is not generally possible to determine who the compliers are because simply observing cases where $Z = 1$ and $D = 1$ simultaneously cannot rule out non-compliers, because some individuals that we observe in this group may still get treatment ($D = 1$) even if $Z = 0$. This means that we cannot be sure about to whom the impact estimates apply from 2SLS. Neither OLS nor 2SLS, therefore, provides impact estimates that can be clearly attributed to some section of the population, or the population as a whole, which makes it impossible to sample match groups in the value calculation. In the case of 2SLS where both Q and M are instrumented (with the aim of deriving causal estimates) impacts of Q and M will pertain to two unobservable groups that are likely to be different and this may be very misleading in WV. It results in biased estimates of the value of Q because, for example, the impact of Q on SWB for the group that complies to the income instrument may actually be very different to the effect estimated by the Q instrument for a different complier group and there is no way to test this in 2SLS.

In the next chapter I develop a new methodology for wellbeing valuation that addresses the problems inherent to the current methods in relation to technical

conditions (1) to (4). After I have developed this new methodology I will discuss how results from wellbeing valuation should be interpreted (Criterion (D)). Then in Chapter 6 I use the new wellbeing valuation approach to value employment-related outcomes as a case study.

Chapter 5

5. Three-Step Wellbeing Valuation

5.1. Background

Above I have set out the four key validity criteria for wellbeing valuation. I have already dealt with Criteria (A) and (B). This chapter sets out a new approach to wellbeing valuation that is capable of deriving total causal derivatives in a way that allows for sample matching and value interpretation for policy purposes. Therefore, the new approach addresses all four conditions that are required for technical validity (Criterion (C)). The fourth validity criterion (D) - related to overall interpretation of the results - will be dealt with towards the end of this chapter after I have set out the framework of the new approach.

The new approach to wellbeing valuation reverts back to the conceptual model set out in Figure 2. There I showed that the value of a non-market good (Q) can be estimated from the MRS between Q and money (M) using equation (10),

(10) Value (Q) = $-(\beta_Q \cdot Q) \Big/ \beta_M$

As discussed above the problems associated with the current wellbeing valuation approach are generally all due to the current approach of using a single equation outcome model with OLS or two staged least squares. In light of these issues, when empirically estimating wellbeing models of the type in equation (8), they are clearly better explained and understood as a set of simultaneous equations in which SWB and the explanatory variables may be jointly determined and may interact with each other. The general approach to estimating simultaneous equation models (SEM) is full-model maximum likelihood estimation (MLE) (Kline, 2005). Estimation through MLE requires a-priori knowledge of the relationships between all variables in the system and the nature of the error terms. However, SEM modelling through MLE does not provide a solution to the technical problems that affect WV. This is because without exogenous variation in the explanatory variables we are unable to attribute

causality in SEM – we still rely on a selection on observables story for identification. Contrary to popular belief SEM does not provide a solution to the causal question in statistics (Kline, 2005). Therefore, moving away from single-equation models and SEM, I develop a new approach here for wellbeing valuation.

The starting point and the key to addressing these problems and to ensuring that wellbeing valuation has technical validity is to separate the estimation process such that the full wellbeing model is estimated in *two separate steps.* The first step models the relationship between income and SWB and the second step models the relationship between the non-market good and SWB.

The main intuition behind this is that we can deal with the issue of indirect effects and estimate total derivatives for each variable, which is not possible under a single-equation set-up. Separating the estimation procedure will also provide a better mechanism for estimating and interpreting heterogeneous effects of Q and it provides greater flexibility to hone each model to derive unbiased estimators. The method will have to be cautious to the issue of sample matching, but a multi-model approach does not in theory make sample matching any more difficult than it is in a single-equation approach.

From the results of these two models the MRS between Q and M can be derived in the final stage of the process. I will show that this three step process can deliver estimates of monetary value that are consistent with economic theory and that are more robust and superior to values derived from the traditional wellbeing valuation methods. I call this method the *Three-Step Wellbeing Valuation* (3S-WV) approach.

5.2. Theoretical framework of the Three-Step Wellbeing Valuation approach

The 3S-WV approach is comprised of the following three stages:

STEP 1: INCOME MODEL

$$(19) \quad SWB_i = f(\ln(M_i))$$

STEP 2: THE NON-MARKET GOOD MODEL

(20) $SWB_i = g(Q_i)$

STEP 3: MONETARY VALUE ESTIMATION

Calculate $MRS_{Q,M}$ from the income and non-market goods models

We can re-interpret the four technical criteria for the specifics of the 3S-WV model. Conditions 1 and 2 require that the SWB impacts of Q (g'_Q) and M (f'_M) are causal total derivatives such that $g'_Q = {}^{d\,SWB}/_{d\,Q}$; and $f'_M = {}^{d\,SWB}/_{d\,M}$. An implication of this is that clearly only non-market goods/services that have a statistically significant impact on SWB can be valued in wellbeing valuation, as an insignificant impact would signify that Q does not have a causal effect on SWB.

Condition 3 requires that individuals (i) from equations (19) and (20) are representative of each other or of the same population.

Condition 4 requires that the SWB impact of Q (g'_Q) has a well-defined treatment effect.

Next I will discuss how 3S-WV addresses these four main technical criteria at a theoretical level. Having set out the theoretical foundations and rationale I will then proceed to discuss issues related to interpretation in 3S-WV. In the final part of this chapter I will focus on estimation of the income model in Step 1 of 3S-WV. Then in Chapter 6 the 3S-WV method is used to value employment-related outcomes.

5.2.1. Causal total derivatives in 3S-WV (Technical conditions 1 and 2)

The key first step is to estimate unbiased causal estimates for the SWB impact of Q and M. One benefit of the division in to two steps of the wellbeing model is that it allows us to employ a mix of any statistical methods that can provide unbiased causal estimates. Following the statistics and econometrics literature the gold standard here

would be to estimate g'_Q and f'_M from two separate studies where treatment (Q and M) are randomised. Assuming that the standard assumptions are met, randomised trials (RCTs)[5] provide unbiased causal estimates with well-defined treatment effects – the ATE and ATT. Further, the non-parametric difference in means estimated from an RCT represents the total derivative of the treatment because no other mediating variables need to be controlled for. Even if some baseline factor is controlled for in an RCT it cannot be a mediating factor because it is measured at baseline. To ensure sample matching we can run two RCTs on two non-overlapping random samples of the same population. If we were to suspect that SWB outcomes are not independent across i in these trials bootstrap standard errors should be used for inference.

Clearly experimental evidence provides the best approach for WV since it provides unbiased causal estimates that represent the full effect of the treatment on SWB, encompassing both the direct and indirect effects. However, the current WV methodology cannot accommodate information from two experimental studies because it has exclusively been based on single-equation models with observational data. The traditional single-equation approach can only be used to assess the results of random assignment for one variable (either Q or M), unless an experiment were to randomly assign both variables together across the same sample, which would be problematic. Only a multi-model approach, like the one set out in 3S-WV, can incorporate results from two different trials or studies.

So a two-step estimation process for the wellbeing model allows us to use optimal methods for estimating causal total derivatives for income and the non-market good. In practice, however, in policymaking random assignment may not always be possible and it is unlikely that we will be able to randomise income in large samples in order to estimate the income model in Step 1 due to financial and ethical constraints. This is problematic because of the central role that income plays in the WV approach.

[5] I use the term 'RCT' to include any study where treatment has been randomised, such as field experiments.

However, in 3S-WV we can employ other statistical methods that will still produce unbiased causal estimates. These are methods that must be able to deal with selection on observable and unobservable factors as well as being able to account for selection on gains to treatment, which creates heterogeneity in impacts on SWB and will have implications for sample matching and the interpretation of values for policy. Under the right conditions methods such as difference-in-difference methods, synthetic control, and regression discontinuity design (RDD) would all provide sufficiently robust results (Angrist et al., 1996) for the income and non-market good models. Instrumental variables can also be used but with the appropriate fixes to allow us to extrapolate the results to well-defined sample populations rather than the complier sub-group. One such technique is the *control function* method, which I shall discuss in detail below.

Again the two-step process is key here. This is because methods that can produce casually robust estimates from observational data (eg, difference-in-difference, synthetic control and RDD methods) can only look at one treatment at a time and as has already been discussed 2SLS with two instrumented variables cannot be used in WV, which means that estimates for the causal impact of Q and M can only be derived separately from different models (data and assumptions permitting). 3S-WV allows for this and a key point to note about 3S-WV is that because the income and non-market good models have been separated it accommodates a variety of statistical methods - *any mix of experimental and non-experimental techniques can be used to estimate the three steps, provided that the modelling criteria are adhered to*.

It should also be noted that if a selection on observables assumption holds then we can use methods such as matching techniques and regression adjustment models. Under these selection assumptions (which will rarely hold), simple matching and propensity score matching techniques will provide unbiased estimates of the total causal derivatives of M and Q, with well-defined treatment effects, such as the ATT. As for regression adjustment methods, these are regression models which account for heterogeneous impacts through interactive terms and again under these selection assumptions it would be possible to estimate the total causal derivatives of M and Q, with well-defined treatment effects provided that only pre-treatment controls are

included in the model. However, as discussed above the assumptions underlying selection on observables methods are unlikely to hold true for wellbeing models and hence experimental methods or non-experimental methods that are robust to selection on unobservable factors, such as IV, are preferred in 3S-WV.

5.2.2. Sample matching and interpreting treatment effects in 3S-WV (Technical conditions 3 and 4)

I have touched on these issues in the discussion in the previous section (5.2.1). They are related more to the methodology involved in implementing the 3S-WV model as they relate to how the results from specific methods are to be interpreted and so I cover the issues of sample matching and interpreting treatment effects in section 5.3. on the 3S-WV methodology.

5.3. Interpreting and understanding values derived from wellbeing valuation (Criterion (D))

This section is set within the context of 3S-WV and as such assumes that the values discussed have been derived robustly from 3S-WV. I cover a list of separate issues concerning interpretation of the wellbeing valuation approach and its results:

5.3.1. Comparing wellbeing values to preference-based values

As discussed in Chapter 3 wellbeing valuation represents the most direct form of valuation in terms of alignment with the general theory of valuation. It derives values for non-market goods without recourse to data on preferences. This has implications for how values derived from wellbeing valuation should be interpreted. In Chapter 3 I argued that values from WV should not be seen as direct complements or comparators to WTP/WTA values. This is contrary to most of the WV literature to date (e.g. Luechinger, 2009; Kountiuris and Remoundou, 2011; Ambrey and Fleming, 2011; Menz and Welsch, 2012; Levinson, 2009; Frey et al, 2004; Ferreira and Moro, 2009). There are some exceptions to this trend – for example, Frijters et al. (2011) and Carroll et al. (2009) respectively use the terminology *income compensation values* and *income equivalence values* – but the general attitude seems to be to view wellbeing values as being qualitatively identical to preference values. This is not

surprising given that a lot of economists are happy to equate SWB measures to utility in economics.

Evidence to support my approach of making a clear differentiation between SWB and preference here comes from a handful of studies that have compared SWB values with values derived from preference-based methods for the same good and have found that in general SWB values differ quite significantly from preference-based values. This has been found in the cases of valuing urban regeneration (Dolan and Metcalfe, 2008) and environmental goods (Levinson, 2012; Luechinger, 2009). Interestingly, Dolan and Fujiwara (2012) find that for the case of valuing free adult learning courses SWB values align quite closely with stated preference values when the stated preference question asks people their WTP for a course that *explicitly* leads to an improvement in life satisfaction. This suggests, therefore, - as per Adler's argument - that preferences entail something different (or in addition) to SWB, although we note that some of the difference in reported values may be driven by econometric problems associated with the wellbeing valuation approach, which are discussed at length in this thesis.

5.3.2. What do wellbeing values mean?

Wellbeing values are linked to changes in people's SWB. The values themselves are different in nature to values derived using preference-based valuation methods and the specific interpretation of a given value depends entirely on the interpretation of the coefficient on the non-market good in the SWB model. There are two types of non-market good that I shall discuss here (the discussion assumes that the models and valuations highlighted here have been derived in a robust fashion using the 3S-WV approach).

(i) Non-market goods measured through binary outcomes

Many of the non-market goods that have been analysed in the wellbeing valuation literature are binary in terms of their possible outcomes. For example, good health, employment, suffering from a disease or illness, living in a safe area, being a victim of crime, achieving an educational qualification and so on.

The coefficient on such a variable in the non-market good model will represent the SWB impact of the outcome that has been coded as "1". For example, if the non-market good is health and the health variable takes on a value of "1" if the individual suffers from the illness and "0" otherwise, the SWB coefficient will represent the impact of the illness on wellbeing (in this case likely to be a negative coefficient).

The actual interpretation of the value (cost) associated with this health condition will depend on two factors in wellbeing valuation: (i) severity, and (ii) duration. Severity refers to how severe the illness is and duration refers to the length of time of suffering which will change due to how quickly people can adapt to the health condition. All binary non-market good variables will depend on these two factors, whereby for positive outcomes we would refer to benefits rather than severity. So for example, for employment – which has a positive effect on SWB – the value would depend on the 'goodness' of the job and the duration of the job to date.

Since wellbeing valuation uses aggregated data severity (benefits) and duration will be determined by the average levels in the sample of the analysis. For example, for a given health condition the coefficient on that variable in the non-market good model shows the impact of the condition measured at the sample average level of severity and adaptation. Some people will have severe symptoms whilst others have very trivial ones. And some people will have been living with the condition for a long time which may mean that they have adapted to it more and hence the negative impact on SWB for them will be lower.

Clearly it would be possible in the 3S-WV analysis to focus on people with particular levels of severity and adaptation to the illness by segmenting the sample and focusing on people that meet these conditions. In general, though, wellbeing values should be interpreted as the value derived at the sample average level of severity and adaptation. For some applications this may be too general and hence problematic if we do not know what the average level of severity or adaptation is from the data. Often a sense of adaptation can be garnered from an assessment of the length of time people have been in that state, but severity is impossible to know unless the survey contains data on severity (e.g. through a self-reported ranking on a Likert scale). However, in many cases the sample average impact and value are sufficient in WV. In which case the

interpretation on severity/benefits and duration/adaptation should be noted in the results.

(ii) Non-market goods measured through continuous outcomes

Some non-market goods are measured on a continuous or non-binary scale in the wellbeing valuation literature. These include outcomes like pollution levels or CO2 emissions. Many of the non-market goods assessed as binary outcomes in the literature could also be assessed using a continuous outcome variable using the levels of severity.

Using a scale for the non-market good reduces or eradicates uncertainty about the interpretation of severity (or benefit), but still leaves open the issue of duration and adaptation.

For example, if we were to now assume that the health condition discussed above were measured on a severity scale where 0 = 'does not have the condition' and 100 = 'worst possible severity of the condition' then we can move away from sample average levels of severity to focus on the value (costs) associated with each level of severity, or a (one unit) change in severity. This may provide a more meaningful interpretation of the value (cost) associated with that health condition. Two things to note here are that this approach only really solves for the issue around interpretation of severity if the categories on the scale are narrow enough. If, for example ,the health condition were measured on a three-point scale from 1 to 3 then actual experienced severity levels may differ quite substantially even among those people that report the same level on the three-point scale and therefore we would run into the problem of producing values (costs) for sample average severity levels again. The second issue to note is that this approach will require assumptions around the functional form of the relationship between unit changes in the non-market good and SWB (which was not required in the binary non-market good approach).

The issues of duration and adaptation still apply with continuous non-market good variables. This is because people will have experienced the non-market good (at different levels of severity or benefit) for different periods of time. If there is adaptation involved then those who have experienced it for longer will experience less

impact on their wellbeing. Therefore, even with continuous non-market good variables the interpretation on duration is still the same: it is the impact of the non-market good at a given level of severity/benefit for the sample average level of adaptation.

5.3.3. Time-horizons

A key question about interpretation posed by Metcalfe (2009) relates to the time frame against which we can measure wellbeing values. Wellbeing valuation studies to date have tended to use large annual survey data whereby people respond to the life satisfaction question once per year (either in a repeated cross-section or panel design). Do these responses signify the SWB for that point in time, or for that whole year, or do they include much more than that such as future expectations? There is unfortunately nothing we can ascertain from the question itself as it is vague on this point. For example, in two of the key wellbeing data sets in the UK the life satisfaction question is posed as follows:

- How dissatisfied or satisfied are you with your life overall? (Understanding Society).
- Overall, how satisfied are you with your life nowadays? (Annual Population Survey).

Neither question has a time reference point and thus could be interpreted in many different ways. The consensus seems to be that wellbeing values represent annual values (for the past year) (e.g. Luechinger and Raschky, 2009; Oswald and Powdthavee, 2008; Powdthavee, 2008; Welsch, 2008b; Helliwell and Huang, 2005; Levinson, 2009) probably mainly for the reason that the data are taken each year or that some of the questions which are used in the wellbeing analysis ask about the past 12 months (e.g. employment status over the past 12 months).

Frijters et al. (2011) is an example of a study that digs a little deeper to move away from assuming simple annualised values. They use quarterly life event data to map out a longer-term impact of the non-market good to include anticipation and adaptation effects. They claim that their wellbeing valuations - based on life

satisfaction - therefore represent a multi-year value for the effect of the event over a number of years, rather than a one-off annual value. However, caution must be applied to these findings. Whilst Frijters et al. (2011) have accounted for anticipation and adaptation effects they still use a time-independent life satisfaction question which means that we cannot say for certain that the values are multi-year annualised values. The problem of interpretation of duration still exists in their paper.

Life satisfaction scores can include expectations about an individual's future wellbeing. For example, if a significant event such as an illness, an accident, or divorce by coincidence happened a few weeks before the survey the individual's life satisfaction response is probably driven mainly by how they feel about the future given the life-changing event and may not take into account much about what had happened over the past year. On the flipside a life satisfaction response is likely to be more reflective of the past 12 months if at the start of the period the individual had experienced some significant event which has impacted on their lives for the past year in many ways, which makes the focus of their evaluation (at the time of the survey) the past year or so. The timeframe for the life satisfaction question is, therefore, likely to be driven to some extent by the timing of significant events before the survey. It could also be driven by upcoming future events on the horizon. A 17 year old's life satisfaction response during, say, early summer is likely to reflect to a large degree the anticipation of going to university from October, and not just recent events or the events of the past 12 months.

The upshot is that - using the example of employment - the statistical association between being employed and life satisfaction will not necessarily reflect the impact of employment on SWB over the past year. It could be entirely driven by people's recent experiences of the job and/or thoughts about the future. A wellbeing value derived from this impact estimate, therefore, may not reflect the annual value of employment to the individual.

In sum, wellbeing values could reflect (i) the value of Q over the past 12 months (or since the last survey); (ii) the value of Q for a sub-period in the past 12 months (or since the last survey); (iii) the value of Q now and in the future. We can make a

number of hypotheses about what wellbeing values may be picking up, but in reality this issue cannot be resolved without recourse to a change in the survey instrument, whereby the time period for the life satisfaction question is specified. For example, to derive annual values we could ask respondents to rate their overall life satisfaction over the past 12 months. For now and for the purposes of this thesis I will assume, as many previous papers have done, that since life satisfaction ratings are usually taken annually (and they are in the data sets that I use here) and that the determinants of SWB used in statistical modelling apply to the past 12 months that wellbeing values represent annual values with the caveat that future research is required to understand this better.

5.3.4. Wellbeing valuation and cost-benefit analysis

5.3.4.1. Aggregation of values

An important conclusion from all this is that since we have argued that SWB values are not WTP/WTA amounts, CBA under the WV methodology must rely on the social welfare function approach rather than the Kaldor-Hicks compensation test approach. This means that wellbeing values need to be distributionally weighted in CBA if individual-specific income levels are used in the calculation of values and indeed the weight itself can be estimated from SWB data. Layard et al. (2008) find that the elasticity of the marginal utility of income estimated from SWB data tends to be higher than the unity values often derived from consumer demand analysis (Blundell et al., 1994; Evans et al., 2005; Evans and Sezer, 2002) and from revealed social values embodied in the taxation system (Cowell and Gardiner, 1999), which implies a higher weight for lower income groups. CBA under WV may therefore be more redistributive or progressive than preference-based CBA.

5.3.4.2. Total economic value (TEV) and wellbeing valuation

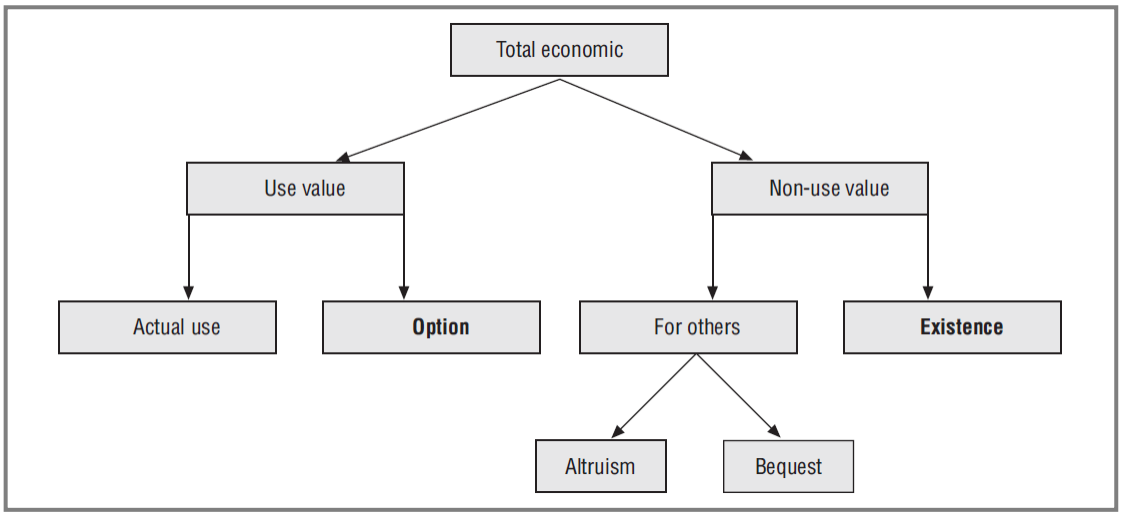
TEV provides an all-encompassing measure of the economic value of any non-market good. It was originally developed in the field of environmental valuation, but is a generic framework that can be applied to any good, although some definitions of value may not be applicable to some goods. WV can be and has been used with non-market goods that encompass an array of values under TEV (e.g. environment) and so

it is important to discuss the extent to which WV can measure TEV. TEV is divided in to *use* and *non-use* value.

Use value relates to the benefits derived from actual use, planned use or possible use of the good.

Non-use value relates to the attachment of value to a good although there is no actual, planned or possible use. There are three types of non-use value: (i) existence value, (ii) altruistic value, and (iii) bequest value. " Existence value refers to the WTP to keep a good in existence in a context where the individual expressing the value has no actual or planned use for his/herself *or for anyone else*. Motivations here could vary and might include having a feeling of concern for the asset itself (*e.g.* a threatened species) or a "stewardship" motive whereby the "valuer" feels some responsibility for the asset. Altruistic value might arise when the individual is concerned that the good in question should be available to others in the current generation. A bequest value is similar but the concern is that the next and future generations should have the option to make use of the good." (OECD, 2006. p.86).

Figure 5. The total economic value (TEV) framework



Source: OECD (2006)

An issue that has been picked up in the wellbeing valuation literature concerns the ability of wellbeing valuation to estimate non-use value. In the revealed preference

literature it is a commonly held belief that non-use values cannot be estimated because revealed preference relies on there being some behavioural 'trace' or 'trail' on the individual's market behaviour and so generally revealed preference methods are not used to ascertain non-use values. Stated preference methods are generally seen as the best (and only) method for estimating non-use value (OECD, 2006).

This mindset has translated over to wellbeing valuation and there seems to be a general consensus that wellbeing valuation cannot shed any light on non-use value (Frey et al., 2004; Levinson, 2009) for similar reasons to those stated in the revealed preference literature that there needs to be some 'trace' or impact on SWB which non-use type outcomes might not have. It is clear from Figure 2 and equations (11), (12) and (13) that in wellbeing valuation, we do need the non-market good to demonstrate some kind of impact on SWB. This will allow us to estimate the derivative of SWB with respect to the non-market good. Thus, whereas in revealed preference, the non-market good must impact on market prices, in wellbeing valuation it must ultimately impact on SWB in some way. Where people use some non-market good or service this should show up in their SWB ratings (if the good/service is important enough) and so use value can clearly be assessed in the wellbeing valuation method.

We should note, however, that contrary to the general belief in the literature the wellbeing valuation approach actually *does not* preclude the possibility of measuring non-use value. The issue regarding non-use value measurement in wellbeing valuation is a data-related issue rather than a technical point. If the following three conditions hold then the non-use value of some non-market good (Q_1) can be estimated in wellbeing valuation:

i. There exists variation in Q_1 either across time or across individuals.
ii. The variation in Q_1 is picked up in the data.
iii. People are aware of this variation in Q_1.

If these three factors are true then Q_1 will impact on SWB regardless of whether this is use or non-use in nature. For example, if Q_1 is the number of rhinos in existence, then if people care about the existence of rhinos and there is variation in Q_1 which we

can pick up in data and introduce in the SWB function then there is no technical or theoretical reason why non-use value would not be picked up. A different issue is whether we would find a statistically significant impact on SWB for Q_1 in this case. Thus, non-use value can in theory be picked up in the wellbeing valuation approach and so the question is more about whether in reality the statistical analysis would do so.

There is some empirical evidence to support this claim that I make; a number of studies have shown that large-scale human or natural disasters impact on the SWB scores of people who were not involved in any way. For example, Metcalfe et al. (2011) found that the 9/11 terrorist attacks in the US had large negative causal impacts on the wellbeing of people in the UK. And Rehdanz et al. (2015) find that the impacts of the Fukushima nuclear power plant disaster can be traced to people in other areas of Japan too. Whilst some of this impact could be from people fearing that it could happen to them, it would be plausible to think that some of the negative impact is due to people feeling sorry and concerned about those that have been affected, which is non-use in nature (altruistic value).

Non-use values or costs can therefore be ascertained in wellbeing valuation where data allow, but this is likely to be rare outside of a few cases (natural disasters and wars are some example areas where non-use related values could be estimated in the data). If, for example, we were interested in the non-use value associated with a cultural institution such as a national museum this could only be estimated in wellbeing valuation if the institution were to suddenly cease operating or if the services that provide the non-use value (such as research activities that are undertaken by the museum) are stopped. These types of cases are unlikely to happen with any frequency which makes estimating non-use value difficult in some cases in wellbeing valuation. In sum, wellbeing valuation can derive non-use value in some areas where there is significant change, but generally speaking stated preference methods are still more comprehensive in their coverage of non-use value than the wellbeing valuation method.

5.3.4.3. Discounting future impacts in wellbeing valuation

A key question related to the use of wellbeing values in CBA is the issue of discounting future wellbeing impacts. CBA uses the social discount rate, which is made up of pure time preference and catastrophe risk plus the value attached to economic or GDP growth by future generations. The latter term is not applicable to wellbeing valuation since wellbeing values only encompass welfare impacts and do not include income or economic impacts. So the question then is whether people discount future SWB gains by the rate of pure time preference and catastrophe risk, but on this topic the wellbeing literature has been pretty silent. I am unaware of any evidence to suggest that people place a (present) time preference on SWB like they do for money and consumption and intuitively there does not seem to be a strong rationale why they would. The work on SWB that uses a discount rate is normally just based on assumptions rather than on any empirical evidence (e.g. Blanchflower and Oswald, 2004; Frijters et al., 2011). Hence, in line with Blanchflower and Oswald (2004) I assume that future wellbeing values *should not be discounted,* although further research can shed more light on this topic going forward.

5.3.4.4. Ex-ante wellbeing valuation versus ex-post wellbeing valuation

As currently practiced wellbeing valuation is an *ex-post method,* whereby values are estimated from the impacts that non-market goods have had on people, which relies on people having experienced the non-market good.

It would, however, be possible to undertake wellbeing valuation in an *ex-ante* mode as well by asking people to *project* the SWB impacts of a future policy, which would provide estimates of β_Q in Figure 2. For example, this could entail the following type of question,

> *"Imagine that under programme X the government will provide more of Y. How satisfied would you be with your life overall if this were to happen?"*

The individual's current level of life satisfaction could be subtracted from the projected level of life satisfaction. This would provide an estimate of the impact of the non-market good on life satisfaction (β_Q). This, in effect, is equivalent to the non-

market good model and after these results are estimated for an ex-ante change the remaining processes of 3S-WV can be carried out to derive an ex-ante value for Q.

The ex-ante wellbeing valuation approach, although never to have featured in any discussion or analysis in the wellbeing valuation literature to date, is potentially important for two reasons. First, it allows us to derive values for non-market goods where data does not exist. This could be, for example, for a new type of project or programme. Up to now valuation of future non-market goods was restricted entirely to stated preference methods. Second, ex-ante wellbeing valuation would provide a potential framework for assessing non-use values by asking people to project their SWB scores under the scenario where the good with non-use value is destroyed or its provision is ceased. For example, the ex-ante life satisfaction question could ask about the impact on life satisfaction due to the loss of wildlife or a cultural institution such as a museum. For non-users this would extract the non-use value of these non-market goods (if there were projected changes in life satisfaction scores).

The ex-ante method, which could also be labelled the *hypothetical wellbeing valuation method* has many theoretical benefits, but in practice there are a number of important issues. First, hypothetical wellbeing valuation relinquishes one of the major advantages of wellbeing valuation – the ability to estimate values based on people's actual experiences.

Second (and relatedly), because impacts are not based on actual experiences then a number of problems emerge. The non-market good model – whether it is estimated from pre-administered survey data or from projected life satisfaction impacts – needs to derive causal estimates for Q, that is that β_Q must be unbiased. Projected impacts can suffer from a large number of biases as can statistically-estimated impacts (as discussed in Chapter 4). As people are asked to estimate the impact of a change in some outcome on their life satisfaction they may struggle if the non-market good/service is not something they are familiar with. And they may be encouraged to provide socially-desirable answers. As discussed people may mis-report their levels of wellbeing depending on the context and to whom they are reporting their ratings. If asked about the impact of some environmental cause or event on life satisfaction

people may overstate the impact on life satisfaction (eg, say that the impact of a loss in wildlife or damage to the environment on their life satisfaction is much greater than it actually is) in order to portray oneself as an environmentally-friendly or ethical person. This will clearly lead to over-stated estimates in hypothetical wellbeing valuation.

The example survey question given above is a within-person survey design. It would also be possible to run a between-person survey design, whereby one group is asked for their SWB scores now and another is asked for their SWB scores if the policy were to take place. The people selected into the different surveys should be done so on a random basis so that the two groups are identical on average. This between-person method is the approach usually taken in the psychological sciences to remove the effect of salience effects and focussing illusions. Whatever survey design is used the problems related to causal inference discussed above will need to be acknowledged. This should be an area for future research.

5.4. Estimation methodology for the Three-Step Wellbeing Valuation approach

The key to the 3S-WV approach is to estimate the income model and the non-market good model in a robust way that adheres to the four technical criteria for wellbeing valuation. From this, step three (estimating a monetary value) can be undertaken. The methodology related to this final third step has been sufficiently discussed above and therefore I focus on methods for estimating the income model and the non-market good model here.

There is one methodological aspect in 3S-WV that is a constant requirement across any study. This is the estimation of the income model. The income model is therefore the crucial component of 3S-WV and I will start with this model and derive an optimal method for its estimation. I will show how the methodology and results from this model can be applied to other 3S-WV studies that use UK data. The methodology can also be replicated using other data from different countries.

The non-market good model is more varied because a given wellbeing valuation study could look at any non-market good. It is, therefore, not possible to produce a single best-practice model. Instead in Chapter 6 I will use a case study looking at employment outcomes and show one possible modelling approach that would adhere to the four technical criteria of wellbeing valuation.

5.4.1. The income model

The income model (equation (18) of 3S-WV) is problematic because of the near impossibility of running large experiments with income and the difficulties associated with previous methods (all have used IVs) for assessing the causal effect of income on SWB, which have been discussed above. As discussed 2SLS methods are problematic for wellbeing valuation and statistical methods that employ non-exogenous income variables are vulnerable to a range of problems in the form of measurement error bias, endogeneity bias and the issue of indirect effects (people who earn more money generally have to work harder to earn it and these indirect dis-benefits or costs (eg, work-related stress) are also captured in the income variable).

In this section I will derive an unbiased estimate of the total derivative of income with respect to SWB for a well-defined sample group - the general UK population. This estimator will be broad enough to actually be used as an 'off the shelf' estimate of the income model in other 3S-WV studies that focus on the same population and life satisfaction variable. Producing the methodological framework and results for a generalizable income model in 3S-WV - which can be widely applied in terms of the framework or the actual results in other 3S-WV studies - is one of the central contributions of this thesis.

5.4.1.1. The relationship between income and wellbeing

There has been a significant amount of work on income and SWB – indeed it is one of the most studied research questions in the wellbeing economics literature. Estimation of an income model needs to take this literature into account.

The role of income in wellbeing came to prominence with the early work by Easterlin (1974) who found that over time wellbeing does not continue to rise with income or

GDP (the Easterlin paradox). However, work since then has tended to refute the idea of an Easterlin paradox. As we shall see in the literature reviewed below income has consistently been found to have a statistically significant relationship with various measures of SWB using a number of data sets from across the world. And work by Stevenson and Wolfers (2008) that focuses directly on the issue of the Easterlin paradox finds strong evidence from large and newer world data that the paradox does not exist and that absolute levels of income matter for people's wellbeing at all levels of income. This thesis takes as a given the assumption that income matters for wellbeing. This is, of course, a fundamental assumption in economics and the theory of monetary valuation (for example the non-satiation assumption relies on the proposition that income improves welfare).

The main questions, therefore, relate to how we should model the relationship between income and wellbeing in an empirical sense. Firstly, the evidence is in favour of assuming a non-linear relationship between income and wellbeing, which accounts for the well-documented diminishing marginal utility of income. Certainly nearly all empirical models in the SWB literature in economics make this assumption (Layard et al., 2008) and I will follow best-practice here.

Second, as already discussed in detail, exogenous changes in income should be used to evidence the relationship between income and SWB. Here I will use exogenous changes in income due to lottery wins to estimate the causal effect of income on SWB. I argue that lottery wins are likely to be the best source of exogenous income changes that we will be able to find in non-experimental data because by law lottery wins have to be randomly assigned across the pool of lottery players.

5.4.1.2. Estimating the causal effect of income on SWB from lottery wins data

A small literature has used lottery wins in the past to identify causal effects of income on wellbeing and health. Apouey and Clark (2010) and Gardner and Oswald (2007) use lottery wins from the British Household Panel Survey (BHPS) as an explanatory variable and they find positive impacts on health and wellbeing. Lindahl (2009) uses data on Swedish lottery winners in 2SLS and finds positive impacts on health and also in Sweden (Lindqvist et al., 2018) use a primary data set of Swedish lottery players and found that compared to matched controls, "large-prize winners experience

sustained increases in overall life satisfaction that persist for over a decade and show no evidence of dissipating with time" (p.12). As Apouey and Clark (2010) say, in the health and wellbeing literature "Lottery wins are an arguably under-exploited source of exogenous variation in income". Outside of the health and wellbeing literature lotteries have been used in research related to labour market and economic decisions (see Apouey and Clark (2010) for examples).

The conclusion is that lottery wins are seen as an acceptable source of exogenous income changes (from a technical perspective), but their use has been limited in the wellbeing literature. This is potentially due to a lack of data. I focus on the studies by Apouey and Clark (2010), Gardner and Oswald (2007) and Lindahl (2009) as they are the most relevant here.

5.4.1.2.1. Problems

The fundamental problem with all three studies is that data in the BHPS and from Sweden only provide information on the size of annual lottery wins. We do not know how often people play and so annual lottery wins are not strictly exogenous: people who play more are more likely to win more money and this is problematic as those who play more are also likely to have different levels of potential income and wellbeing to start off with. This means that in a standard regression setting (as per Apouey and Clark, 2010; Gardner and Oswald, 2007), lottery win income is not exogenous and in relation to 2SLS (Lindahl, 2005) lottery win income will be endogenous in the first stage and so in this case neither regression nor 2SLS provide a solution. And this is demonstrated by Lindahl (2005) and Apouey and Clark (2010), who show that annual lottery wins in both datasets are correlated with a host of socioeconomic variables and this is why all of the papers hold these variables constant in an attempt to ensure exogeneity in the lottery prize variable. In other words, the papers rely on a conditional exogeneity assumption for lottery wins.

However, the fact that a range of socioeconomic variables are found to be determinants of lottery win size means that there are also likely to be a host of other *unobservable* confounding variables. Hence, only controlling for some of the observable characteristics that determine winnings is unlikely to produce unbiased causal estimates for income.

The second problem with previous 2SLS lottery wins studies (for the purposes of wellbeing valuation) is that the localised complier estimates from 2SLS are too vague for use in wellbeing valuation. In other words, it is impossible to do sample matching with 2SLS outputs as discussed above. This problem is not dealt with sufficiently by OLS regression either since OLS produces vague and uninterpretable treatment effects that lie somewhere in between the ATT and the ATNT (Humphreys, 2009).

Finally, one other problem with OLS is that we cannot eradicate the problem of measurement error in the lottery wins variable, which would lead to a downward bias in the coefficient size when wellbeing is simply regressed onto lottery win amounts.

5.4.1.2.2. Solutions

The favoured approach here is to employ an IV framework since this eradicates the issue of measurement error, which we know is problematic for income variables, whilst also dealing with the other issues highlighted above such as endogeneity. However, an approach is required whereby we are able to derive a causal effect with a clear interpretation regarding to whom the estimates apply, and which addresses the issue around endogeneity of the lottery wins variable. This cannot be the 2SLS estimator.

The approach taken here is therefore different to the rest of the lottery wins literature. I use lottery wins data from the BHPS with the *control function* approach. The control function is an alternative method for estimation using IV that can be used instead of 2SLS or the Wald estimator. The control function approach allows us to derive estimates of the sample average partial effect (APE) for income, whilst also dealing with the issue of measurement error in the same way that 2SLS does. The control function estimator represents the effect of income on life satisfaction that we would expect for anyone in the sample, which is a much clearer and broader treatment effect than that obtained from 2SLS. This simplifies the task of sample matching in wellbeing valuation and makes the control function approach ideal for 3S-WV.

To address the issue of endogeneity I hypothesise that the amount of previous lottery wins will capture lottery playing preferences and hence current playing frequency

more accurately than observable socioeconomic factors - on the assumption that people who played a lot in the past will always tend to play a lot, unless they win very large amounts, but large winners are excluded from the analysis. I find that controlling for previous wins leaves all other observable background variables statistically insignificant in determining annual lottery win size (see Table 3) - evidence that controlling for previous wins will ensure exogeneity in the lottery wins instrument. Previous lottery win amounts have a positive statistically significant effect on current lottery win amounts as would be predicted from my hypothesis.

Table 3. Determinants of annual lottery wins size

Independent variables	Coefficient	S.E.
low education	136.903	117.398
age	-2.066	-3.457
male	129.526	112.586
poor Health	-154.732	-200.634
unemployed	-98.941	-446.597
no. of children	81.733	70.75
lagged income	-0.001	-0.002
previous lottery wins	0.07***	0.014
constant	249.086	228.99
Observations	5,269	

Notes: Dependent variable: annual lottery win amounts. Variable descriptions in Table 4.
**** significance at 0.01 level, ** significance at 0.05 level, * significance at 0.1 level.*

Below I will discuss how the 3S-WV approach with the control function addresses the technical criteria set out in Chapter 4.

5.4.1.3. Data

Data comes from the BHPS, which is a nationally representative sample of British households, containing over 10,000 adults, conducted every year since 1991. Life satisfaction (measured on a scale of 1 – 7) was added in 1997 and so we analyse the period 1997- 2009, excluding 2001 which did not include life satisfaction questions. The BHPS asks respondents whether they have won money on lotteries or football pools and how much they have won in total during the year. In the UK there are a large number of lottery players (Provencher et al., 2012) and these swamp the football pool players in the BHPS dataset (Gardner and Oswald, 2007). I will therefore refer to

this group simply as lottery winners as Gardner and Oswald (2007) do. Table 4 shows the descriptions for all variables used in the analysis.

Table 4. Variable descriptions

Variables	Descriptions
Life satisfaction	Life satisfaction score, coded on a seven-point scale so that 1 = very dissatisfied, 7 = completely satisfied
Job satisfaction	Job satisfaction score, coded on a seven-point scale so that 1 = very dissatisfied, 7 = completely satisfied
Leisure satisfaction	Leisure time satisfaction score, coded on a seven-point scale so that 1 = very dissatisfied, 7 = completely satisfied
Health satisfaction	Health satisfaction score, coded on a seven-point scale so that 1 = very dissatisfied, 7 = completely satisfied
Social life satisfaction	Satisfaction with social life score, coded on a seven-point scale so that 1 = very dissatisfied, 7 = completely satisfied
GP visits	Number of GP visits
Household income	Annual equivalised gross household income
Household size	Number of people living in the home
House ownership	= 1 if respondents owns their home
Unemployed	= 1 if not employed or self-employed
Spouse employed	= 1 if spouse is employed or self-employed
Redundant unemployed	= 1 if respondent was made redundant (and is still unemployed)
Retired	= 1 if retired
Job hours	Hours worked per week
Male	= 1 if male
Age	Age of respondent
Low education	=1 if left education after minimum compulsory
Poor health	= 1 if respondent assesses own health as 'poor' or ' very poor'
Carer	= 1 if respondent provides care of others

Previous lottery wins	Sum of previous lottery wins (£)
Lottery win	= 1 if respondent won between £100 - £50,000 in lotteries over the year
No. of children	Number of children under age 16 in the household
Married	= 1 if married)
Divorced	= 1 if divorced
Widowed	= 1if widowed
Separated	= 1 if separated
Never married	= 1 if never married
Winter interview	= 1 if survey was taken in winter
Living in safe area	= 1 if respondent does not live in an area where they perceive vandalism and crime to be a problem.
Debt burden	= 1 if repayment of debt and associated interest is a 'heavy burden' or 'somewhat of a burden'

5.4.1.4. Methodology

The control function uses some of the basic set up from 2SLS, but explicitly accounts for impact heterogeneity in the model so that sample average (rather than complier average) effects can be measured. Following Heckman and Vytlacil (1998) I run a correlated random coefficient (CRC) model using lottery wins as an IV (Z) for household income and controlling for previous lottery wins. For previous wins, I sum annual lottery wins over all years in which the respondent was present in the data up to and including $t-1$. The model is set up as follows (dropping the time and individual subscripts for simplicity):

(22) $$LS = \pi + \beta_1 \ln(M) + \beta_2 X + \varepsilon$$

(23) $$\beta_1 = \alpha_1 + \vartheta_1$$

(24) $$\ln(M) = \pi + \gamma Z + \vartheta_2$$

so that,

(25) $$LS = \pi + \alpha_1 \ln(M) + \beta_2 X + \vartheta_1 \cdot \ln(M) + \varepsilon$$

Here the impact of income on life satisfaction is made up of a constant term and an individually unique term (ϑ_1). This is the unobserved heterogeneity and in essence, the term $\vartheta_1 \cdot \ln(M)$ in (25) removes the complier effect so that $E(\beta_1) = \alpha$ = the average effect of income for the sample. Equation (24) is equivalent to the first stage in 2SLS as it shows the relationship between the instrument (lottery wins) and income. Since M is endogenous in (22), ε and ϑ_2 are correlated, and under the assumption of heterogenous treatment effects ϑ_1 and ϑ_2 are also correlated. Therefore, ϑ_1 and ε in (25) are estimable from the error term from equation (24): $E(\vartheta_1|X, M) = \theta_1\vartheta_2,\ E(\varepsilon|X, M) = \rho_1\vartheta_2$. Equation (25) then becomes:

(26) $$LS = \pi + \alpha_1 \ln(M) + \beta_2 X + \theta_1\hat{\vartheta}_2 \cdot \ln(M) + \rho_1\hat{\vartheta}_2$$

where $\hat{\vartheta}_2$ is the predicted error term from (24).

The assumptions underlying the control function are somewhat more restrictive than those for 2SLS. In addition to the standard assumptions for valid instruments, we assume that $E(\varepsilon|\vartheta_2)$ and $E(\vartheta_1|\vartheta_2)$ (respectively unobserved self-selection and unobserved selection on gains) are linear functions as is standard in the control function approach. Also, we note that the composite error term in (25) ($\vartheta_1 \cdot \ln(M) + \varepsilon$) has a non-zero heteroscedastic mean and so heteroscedasticity-robust standard errors are used. It is noted that we do not require the monotonicity assumption in this set-up because we can assume that we have one-sided non-compliance to the instrument – in other words, it is reasonable to assume that the subject pool comprises of compliers and always-takers for the lottery wins instrument. Never-takers would be people that do not cash in on winning lottery tickets, which seems unlikely.

Under these assumptions α_1 in (26) represents the causal effect of a log-point change in household income on life satisfaction for the average person in the sample.

In equation (24) I use the following conditional independence assumption:

(27) $$(LS_0, LS_1, D_0, D_1) \perp Z \mid \text{previous wins}$$

where the "treatment" (D) is an increase in household income and Z is lottery wins. This implies that (conditional on previous win amounts) lottery wins cannot be correlated with other determinants of household income (exogeneity) and that lottery wins can only affect life satisfaction through the impact on income (exclusion restriction). Proof for the exogeneity assumption under this set up is demonstrated in Table 3. Now, it could be argued that the exclusion restriction could fail here as lottery winners may also be happier because of euphoria experienced at winning the lottery. Therefore, here I compare lottery winners of different amounts as in Gardner and Oswald (2007) and Imbens et al. (2001). So $Z = 0$ for people with (small) annual wins of under £100 and $Z = 1$ for people with medium sized annual wins of £100 to £50,000. Wins are restricted to a maximum of £50,000 since sample sizes get very low after this point, which makes extrapolation shaky. Here both groups are winners and will feel some happiness due to having won. Is there still a problem that larger winners (the $Z = 1$ group) may feel more euphoria than smaller winners (the $Z = 0$ group)? This is will be undoubtedly true, but it suggests that the level of euphoria experienced at winning the lottery is correlated with win size, which suggests that it is the money prize that causes happiness; precisely the effect we are interested in for the instrument. Second, the euphoria felt from the act of winning itself may only be temporary anyway and not picked up in the life satisfaction responses at the time of the survey. Hence, I argue that the exclusion restriction assumption is satisfied for lottery wins under this set up.

By comparing the sample of small to medium-sized lottery winners the control function will derive the causal effect of income for the average lottery player in the UK. This has implications for the task of sample matching, which is discussed in more detail below.

The control function approach is preferred here to other methods in the literature that attempt to extrapolate localised IV effects (LATE) to population average effects. Examples of such studies are Aronow and Sovey (2010), Follmann (2000) and Angrist and Fernandes-Val (2010). The basic premise of these methods is to explain heterogenous impacts through differences in observable characteristics across the sample, an assumption also used by regression adjustment techniques. Sub-group

differences in LATEs are estimated (eg, for a certain age group) by restricting the sample in 2SLS by this characteristic and then this is extrapolated to other sample groups based on the breakdown of age and other characteristics.

These methods are problematic because it is assumed that differences in impact size can be explained solely by observable characteristics and because proper extrapolation requires knowledge of the characteristics of non-compliers' (always-takers and never-takers) who, like compliers cannot be observed in the data. The control function approach does not rely on these tenuous assumptions and explicitly models and controls for the heterogeneity instead.

5.4.1.4.1. The control function and the conditions underlying 3S-WV

I focus on the parts relevant to the income model here. The first three conditions are relevant for the income model.

5.4.1.4.2. Estimating the full causal (direct + indirect) effects of the non-market good and income on SWB (Condition 1 and Condition 2)

In respect to income in the control function model provided that the conditional independence assumption holds along with the other assumptions that are also made in the first stage of 2SLS, then α_1 in equation (23) will represent the causal effect of income on life satisfaction. The only control variable we use in the control function is previous lottery win amounts in order to ensure exogeneity of the lottery wins variable. In other words, the only other variable in the control function regression model is a pre-treatment variable and so income can still have indirect effects on wellbeing. Hence α_1 represents the total derivative of wellbeing with respect to income ($^{\mathrm{d}\,SWB}/_{\mathrm{d}\,M}$) and therefore, we can go one step further and claim that α_1 in equation (23) will represent the full causal effect of income on life satisfaction.

5.4.1.4.3. Sample matching (Condition 3)

The control function provides a clear treatment effect interpretation for the total derivative and hence makes the task of sample matching possible. α_1 in equation (23) represents the full causal effect of income on life satisfaction for the average lottery player in Britain. For the purposes of sample matching this means that we would have to either estimate the impact of the non-market good on life satisfaction (g'_Q from

equation (20)) for the average lottery player in Britain or find a way of converting the causal estimate (α_1) for lottery players to a general population effect. The latter is preferable and since the evidence suggests that a large proportion of the UK population (over 70%) play lotteries (Telegraph, 2014)[6] we will assume here that the causal effect of income for the sample of lottery players from the control function is equivalent to the causal effect of income for the general population in Britain.

In other words, the control function approach allows us to derive an estimate of the average effect of income on life satisfaction for anyone in the British population. This is a broad treatment effect which can be used with results from a generalised non-market good model or the control function could be re-estimated for different population groups (eg, different age and socio-economic groups) to more closely match the sample profile of a specific non-market good model. Given the generalizability of the results from the control function it would also be possible to use the results as an 'off the shelf' estimate for the income model in any 3S-WV study pertaining to the UK that uses the same life satisfaction variable: we can use the results from (26) with any non-market good model that can derive UK population average estimates for g'_Q in equation (20). In the results section I will present a table of value estimates for different impact sizes of the non-market good.

5.4.1.5. Control function results

Tables 5 and 6 present the results of the control function for income. The first stage (shown in Table 5) is equation (24). I find that winning the lottery has a highly significant positive effect on household income after controlling for previous win amounts.

[6] https://www.telegraph.co.uk/finance/personalfinance/11228989/National-Lottery-20-amazing-facts-from-20-years-of-creating-millionaires.html

Table 5. Control function: first stage regression

Dependent variable: log(household income)

Independent variables	Coefficient	S.E.
lottery win	0.102***	0.015
previous lottery wins	6.82e-06***	0.000
constant	9.999***	0.007
observations	10,461	

Notes: Model estimated using equation (24). Variable descriptions in Table 4.

**** significance at 0.01 level, ** significance at 0.05 level, * significance at 0.1 level.*

I find that income then has a statistically significant effect on life satisfaction in the second step of the control function as shown in Table 6.

Table 6. Control Function: the causal effect of income on life satisfaction

Dependent variable: life satisfaction

Independent variables	Coefficient	S.E.
log (household income)	1.103***	0.252
previous lottery wins	-0.00001***	0.000
$\hat{\vartheta}_2$	-1.108***	0.260
$\hat{\vartheta}_2 \cdot \ln(M)$	0.011*	0.006
constant	-5.777**	2.530
Observations	10,328	

Notes: Model estimated using equation (26). Variable descriptions in Table 4.

Predicted error term ($\hat{\vartheta}_2$) estimated from first stage regression (Table 5).

**** significance at 0.01 level, ** significance at 0.05 level, * significance at 0.1 level.*

Heteroscedasticity-robust standard errors.

The sample average effect of a log-point change in household income on life satisfaction is 1.1, which is also highly significant. This represents the causal effect of income on life satisfaction for any lottery player chosen at random in the BHPS, which we can assume to represent the average effect for the UK population. No post-treatment variables are included in the model and hence this is the total derivative of household income on life satisfaction:

(28) $$\frac{\mathrm{d}\, SWB}{\mathrm{d}\, M} = 1.1$$

A comparison of this coefficient against an income coefficient derived using OLS can be made by reference to Table 11 in Chapter 6. The OLS models in that table, which use log of household income from the same BHPS data, show a coefficient size of around 0.1, significant at the 1% level (the mean for the two OLS models is 0.1035). This indicates that in comparison to an income coefficient estimated using a non-exogenous income variable (household income) the results from the control function using a lottery wins instrument for income are around 10 times higher. This is in line with findings from previous studies in the literature that have employed instrumental variables for income which on average find a 10-12 fold increase in the size of the income coefficient when using an instrument (see section 4.4.3.2.2.).

We note that the interactive term ($\hat{\vartheta}_2 \cdot \ln(M)$) is significant at the 10% level, showing some evidence for heterogeneous impacts of income. Also, $\hat{\vartheta}_2$ is significant which is proof that the income variable is endogenous in the life satisfaction equation and is likely that standard OLS would generate biased estimates of the causal effect of income. The coefficient is negative implying that in cases where income is not exogenously determined we will see downward bias in the income coefficient in regression models.

Table 7 offers a quick-reference chart of values (compensating surplus) for hypothetical impact sizes based on the causal effect of log of income of 1.1 and an average income of £23,000. This gives an idea of the values associated with different coefficient sizes for non-market goods or 'bads'. The values are based on life satisfaction models where life satisfaction is measured on a seven-point scale.

Table 7. Monetary values for hypothetical wellbeing impacts

Hypothetical impact size for Q	CS for welfare gain	Hypothetical impact size for Q	CS for welfare loss
0.0001	£2	-0.0001	£2
0.0005	£10	-0.0005	£10
0.001	£21	-0.001	£21
0.005	£104	-0.005	£105
0.01	£208	-0.01	£210
0.05	£1,022	-0.05	£1,070

0.1	£1,999	-0.1	£2,189
0.25	£4,676	-0.25	£5,869
0.5	£8,401	-0.5	£13,235
0.75	£11,369	-0.75	£22,482
1	£13,733	-1	£34,087
1.5	£17,118	-1.5	£66,939
2	£19,267	-2	£118,695

Notes: CS and ES values for different impact sizes on life satisfaction based on seven-point scale. Calculations made based on SWB impact of log (household income) of 1.1 and average income level of £23,000.

The issue of unconstrained CS values in welfare losses has been discussed above and is demonstrated empirically in Table 7. The CS for large negative impacts on welfare (life satisfaction) quickly exceeds the income constraints in this particular example (£23,000). The reasons that explain why the CS for a welfare loss is greater than the CS for the equivalent welfare gain (which is detectable for non-trivial levels of welfare change in Table 7) have been discussed in section 4.4.1.1.

5.5. Summary

This chapter is the key chapter in this thesis. I start by developing a new technical framework for assessing the validity and rigour of WV studies, culminating in a set of four key conditions and criteria. I demonstrated that the current WV methodology does not adhere to these criteria in a number of ways and that because of this the results are likely to be biased.

Based on these new criteria I developed a new approach to WV, the Three-Step Wellbeing Valuation (3S-WV) method and estimated the key parameters in the model. In the next chapter I apply the new 3S-WV approach to employment outcomes and compare how 3S-WV performs relative to the current WV methodology.

Chapter 6

6. Application of the Three-Step Wellbeing Valuation method to employment outcomes

6.1. Introduction

Chapter 5 set out the technical conditions for estimating monetary values using the wellbeing valuation approach together with a discussion on interpreting wellbeing values. The main contribution of the section was the development of a new approach to wellbeing valuation, known as 3S-WV which provided a solution to the main technical criteria. I produced a generalizable income model for 3S-WV. It was noted that a similarly generalizable model for the non-market good was not possible given the variety in the types of non-market goods assessed and the availability and format of the relevant data.

In Chapter 6 I apply the 3S-WV methodology to the case of valuing employment outcomes. I will look at the values associated with the non-pecuniary aspects of employment, which will form the non-market good in this case study. The study will, therefore, provide a practical example of using 3S-WV. I choose employment because many government departments and labour ministries, such as the UK Department for Work and Pensions have struggled to quantify and evidence the benefits of employment beyond the wage salary. As discussed in Chapter 1 the non-financial benefits of employment should be considered in valuation and CBA and so the thesis makes an important contribution in this respect.

The labour market has long been an area of great interest and enquiry for economists and following in this tradition, a key research question for economists now working with SWB data concerns the impact of employment status and different aspects of the job itself on people's wellbeing. In this chapter I seek to identify the causal effect of unemployment on life satisfaction for different population groups. This will allow me

to value employment status from the perspective of different people in society broken down by factors such as age, gender, educational status and so on.

I contribute to the literature by using a natural experiment exploiting data on job redundancies to estimate the casual effect of unemployment on life satisfaction. I develop a model that can be used to provide for the first time the full causal effect (total derivative) with a clear treatment effect interpretation and a clear sample for sample matching. This ensures that the particular non-market good model (for employment) developed here aligns with the four technical criteria of 3S-WV in order to deliver unbiased estimates of monetary value in wellbeing valuation.

In section 6.2. I start with a review of the literature on employment and wellbeing before developing the statistical model in section 6.3.

6.2. Literature review

There has been a large amount of literature devoted to the topic of employment and wellbeing within the social sciences, but as Kassenboehmer and Haisken-DeNew (2009) explain much of the previous literature is limited in terms of its ability to attribute causality. The main focus of this literature review is on studies that use methodologies that can provide causal estimates for employment, but I shall also touch on general themes in the employment and wellbeing literature.

6.2.1. Theoretical literature

Work could impact on wellbeing in a number of potential ways - both positively and negatively. We can think of there being pecuniary and non-pecuniary impacts. In terms of pecuniary effects, in work people receive a wage income and possibly other monetary rewards, such as employer contributions to a pension scheme. On the flip-side there may also be some unavoidable monetary costs, such as travel costs and childcare costs (Greenberg and Knight, 2007).

In addition, employment may have positive impacts on an individual's wellbeing over and above the effects of the pecuniary benefits. These are the non-pecuniary benefits of employment and the theoretical framework that has traditionally underpinned much of the work in this area derives from Jahoda's (1982) *Latent Deprivation Theory*, which proposes that employment provides access to five categories of experience that are important to health and wellbeing. These are i) *structured time use;* ii) *activity;* iii) *social contact;* iv) *collective purpose* and v) *status.* Similar theoretical models include Warr's *Vitamin Model* for employment (Warr, 1987; Warr, 1994).

People may also incur some non-pecuniary costs in employment as they have to substitute leisure and/or home production time for time at work. Leisure and home production may have a positive value to the individual and so there could be a loss to the individual from forgoing this time (Greenberg, 1997; Greenberg and Robins, 2008). There may also be dis-amenities associated with the job, such as stress and fatigue (Waddell and Burton, 2006).

6.2.2. Studies that use endogenous employment variables

Based on these theoretical frameworks, there has been a proliferation of empirical work looking at the variety of ways in which work can impact on wellbeing. A major review of the evidence was conducted by Waddell and Burton in 2006 and subsequently meta-reviews of the literature on different aspects of employment and wellbeing have been conducted by for example Erdogan et al. (2012) and Joyce et al. (2010). Like Kassenboehmer and Haisken-DeNew (2009) Waddell and Burton (2006) also recognise the dearth of studies that are able to adequately attribute causality.

In Waddell and Burton's (2006) review of approximately 400 studies they are clearly conscious of not attributing causality too readily and much of the language they use refers to 'relationships' and 'associations'.

A wide range of methods have been used with endogenous employment variables. These are methods that do not use random assignment in some form or attempt to

apply methods such as IV to solve for the endogeneity problem. There are two types of methods:

(i) Quasi-experimental methods that apply statistical 'fixes' to control for confounding factors. Here methods that have been used include, regression analysis using cross-sectional and panel data (for example Clark and Oswald, 1994; Tella et al., 2003), where in the latter fixed effects may be used to control for time-invariant factors (for example Gerlach and Stephan, 1996; Korpi, 1997; Winkelmann and Winkelmann, 1998); and structural equation modelling (for example Barnett and Brennan, 1995; de Jonge et al., 2001; de Jonge and Schaufeli, 1998; ter Doest and de Jonge, 2006).

(ii) Non-experimental methods that make no attempt to control for differences and confounding factors. This includes t-test analysis; follow up studies which track people over time as they change employment status (see Murphy and Athanasou, 1999 for a meta-analysis) and qualitative analysis such as semi-structured interviews (van den Berg et al., 2015).

The main messages and findings from these types of studies are as follows. There is an overall finding that employment is associated with wellbeing measured as happiness, job satisfaction and life satisfaction (Andersson, 2008; Berger, 2009; Binder and Coad, 2012; Booth and Ours, 2012; Bowling et al., 2010. p.201; Brereton et al., 2008; Clark, 2010; Cuyper et al., 2008; Khattab and Fenton, 2009; Maennig and Wilhelm, 2012; Ryan et al., 2010; Zeng et al., 2012).

The happiness impact seems to depend on the type of job and the context of the job, with the self-employed generally happier than the full-time employed (Benz and Frey, 2008; Binder and Coad, 2012; Brereton et al., 2008), and the full-time employed happier than part-time employees (Booth and Ours, 2012). Non-participation in the workforce, either through unemployment or unpaid family work are commonly associated with lower life satisfaction compared to those who are working (Berger, 2009; Clark et al., 2010; Maennig and Wilhelm, 2012). Informal jobs have been found to have no worse effect on life satisfaction than employment by formal contracts (Aistov et al., 2012). Generally there is no difference in life satisfaction between

private and public sector employees, although private sector employees' subjective wellbeing is more sensitive to fluctuations in unemployment rates than that of public sector workers (Luechinger et al., 2010). Men are more negatively affected by regional unemployment rates than women (Clark et al., 2010).

A number of studies have analysed the association between work-life balance and wellbeing (Booth and van Ours, 2009; D'Addio et al., 2007; Gröpel and Kuhl, 2009; Pouwels et al., 2008). Flexibility of working hours is associated with higher levels of happiness (Atkinson and Hall, 2011; Torka and Schyns, 2007). In contrast mandatory overtime has been found to be associated with lower levels of happiness (Golden and Wiens-Tuers, 2006). Booth and van Ours (2009) find that part-time women are more satisfied with working hours than full-time women, and that women's life satisfaction is increased if their partners work full-time. Relative salary has been found to affect satisfaction, with satisfaction levels dependent on the rank of an individual's wage within a comparison group (Clark et al., 2009; Wyld, 2011).

A large number of studies have analysed the characteristics which increase life satisfaction and happiness in the workplace. These include engagement and security (Pouwels et al., 2008; Siu et al., 2007), perceived control (Håkansson et al., 2011; Khattab and Fenton, 2009), and trust (Helliwell and Huang, 2010; 2011). Opportunities for promotion, supervisory roles and union membership are also associated with higher life satisfaction (Campione, 2008; Chaiprasit and Santidhiraku, 2011; Cuyper et al., 2009). A large number of such studies focus on job satisfaction (Altinoz et al., 2012; Bernhard-Oettel et al., 2008; Bilgin and Demirer, 2012; Brown et al., 2008; Brown et al., 2008; Kristensen and Johansson, 2008; Lévy-Garboua et al., 2007; Origo and Pagani, 2009), with a subset of studies linking job satisfaction to job performance and productivity (Bowling, 2007; Cornelissen et al., 2011; Jones, 2006; Judge et al., 2010).

These studies are generally problematic from the point of view of inferring causality. None of the quasi-experimental methods discussed here can fully account for endogeneity biases such as selection bias, reverse causality and measurement error. These biases have been discussed at length elsewhere in this thesis. Some of the studies that have used panel data methods such as fixed effects have been highly cited

in the literature, but they are still susceptible to the problem of reverse causality and selection on unobservable time-variant factors such as motivation and ability. And we must not forget that fixed effects modelling can lead to attenuation bias in that by taking out between-individual variation, it increases the ratio of measurement error to actual variation in variables that are measured with error (Deaton, 1993). This can lead to a downward bias in the coefficients on employment variables in fixed effects regression models. Most of the discussion regarding problems of inferring causality that was presented in the section on income above is relevant to the issue of the causal effect of employment.

Non-experimental studies are also highly problematic for inferring causality. Simple t-test analysis allows for all confounding factors to influence the difference in mean estimates for SWB. Before and after trend studies do not control for history effects and other potential biases such as the impact of regression to the mean. Qualitative approaches are problematic because people are unable to predict what the counterfactual would have been like, people may provide socially desirable answers (it will seem ungrateful if an employed person says he is unhappy, especially during times of economic hardship), and cognitive dissonance may play a part in that people re-align their beliefs to fit their behaviour and could say they are happier with a job than without a job, or how else would they explain the fact that they go to work every day (for example see Gangl, 2010; McGill, 2000).

6.2.3. Studies that use exogenous employment variables or that exploit methods that permit better inferences of causality

Much of the past literature is not directly relevant to the study here since we cannot directly infer causality. It is fair to say that much of the literature on employment and wellbeing is at a conceptual or theoretical level (e.g., Bartley, 1994; Shortt, 1996) or uses non-exogenous changes in employment status (Kassenboehmer and Haisken-DeNew, 2009).

However, there do exist a handful of studies that seek to use methods that allow for better inferences about causality. Randomisation is clearly very difficult to use when

we are interested in employment status per se as it would involve randomly assigning jobs to people. Most randomised trials or experiments in this area have focussed on the random assignment of job training advice and employment programmes (e.g., Hendra et al., 2011; Caplan et al., 1996), or work-related conditions such as flexible working (Joyce et al., 2010), rather than employment per se. These types of trials do not estimate the effect of employment on SWB, but could feasibly do so through an instrumental variable methodology provided that the assignment of the employment training itself does not impact on SWB through channels other than employment status, but this would be a highly questionable assumption.

Instead, an alternative strategy is to find naturally occurring exogenous changes in employment status. This drove some of the early work on factory closures and health, which would find large factories that were about to go out of business and measure health status of the employees before and after compared to control groups. They tended to find large negative effects of job loss on objective measures of health, such as medical usage rates and mortality (Beale and Nethercott, 1987; Iversen et al., 1989; Studnicka et al., 1991; Burgard et al., 2005; Keefe et al., 2002). These studies, however, have been criticised for small sample sizes, the lack of generalisability of the results and for not using appropriate control groups (Morris and Cook, 1991). The study designs used in factory closures, despite these criticisms, are of theoretical interest to this current study although the earlier work on factory closures was focussed on physical health outcomes rather than wellbeing.

As far as this author is aware, Kassenboehmer and Haisken-DeNew (2009) is the only study to date that employs a large nationally representative dataset and aims to use exogenous changes in employment status to identify the impact of unemployment on SWB. They use 15 years of the GSOEP and look at the reasons respondents reported for leaving their previous job. In the survey people can choose: *'quit for personal reasons', 'transferred by firm', 'transferred on own account', 'reaching retirement age', 'wanting to look for another job', 'personal reasons', 'time-limited work contract', 'quit on one's own', 'giving up working', 'fired by employer', 'on leave or sabbatical', 'company closing'* and *'other reasons'.* The authors use the 'company closing' response to proxy for exogenous moves in to unemployment. They run OLS and logit models (where in the latter the life satisfaction responses are set as a binary

variable indicating high or low wellbeing). Models are run with and without fixed effects and with a standard set of controls, including marital status, education, household size, health and income. They also control for being unemployed and entry into unemployment due to being sacked.

Schmitz (2011) in many ways replicates the Kassenboehmer and Haisken-DeNew (2009) study by using the same unemployment variables and GSOEP dataset, but instead looking at unemployment and health outcomes rather than wellbeing.

This chapter focuses on causality and the creation of a model that aligns with the main technical conditions of 3S-WV. It, therefore, relates more closely to the much smaller literature on employment and wellbeing which attempts to exploit data on exogenous changes in employment status. I employ the same basic structure and assumptions as the Kassenboehmer and Haisken-DeNew (2009) study and the factory closure studies to run an employment-related non-market good model which adheres to the four technical conditions of 3S-WV. In sum, this study derives the total causal effect of employment status on SWB with a clear treatment effect, which allows for sample matching and an unambiguous interpretation of the monetary value in wellbeing valuation.

6.3. Methodology

6.3.1. Strategy

The study presented here develops from the Kassenboehmer and Haisken-DeNew (2009) paper in a number of important ways using redundancy data from the BHPS for the UK rather than the German GSOEP. First, I test whether redundancies are exogenous by looking at balance tests across those made redundant and those that keep their jobs. Kassenboehmer and Haisken-DeNew (2009) did not test this and hence their paper relied on assumed exogeneity of the company closure variable. But there is evidence within the paper that company closures as measured in the GSOEP are endogenous to some degree, because Kassenboehmer and Haisken-DeNew show very large changes in the size of the unemployment coefficient when fixed effects are

included in the model and on many occasions 'company closing' becomes insignificant. If 'company closing' were truly exogenous it should be robust to the inclusion of other control variables including fixed effects.

Second, Kassenboehmer and Haisken-DeNew include a large set of standard control variables for SWB models. It is not clear why this is necessary if the company closure variable is truly exogenous, but aside from that it makes it highly likely that some indirect effects, such as the impact of employment on SWB through health, are controlled for when other control variables are included in the model. The model developed here fully acknowledges the indirect effects of employment on SWB. This is because exogeniety of the redundancy variable permits a comparison of the non-parametric difference in means in SWB scores across the employed and unemployed groups. In this framework we can capture any *indirect effects* that unemployment may have on SWB (for example, unemployment shocks may lead to reductions in health and increased likelihood of divorce or separation which are themselves determinants of SWB). This is the first study that assesses the *full effect* of unemployment on wellbeing capturing all of the possible indirect effects. Therefore the results will represent the total (causal) derivative of life satisfaction with respect to unemployment and this can be used directly in 3S-WV.

Third, Kassenboehmer and Haisken-DeNew (2009) and other highly-cited papers such as Winkelmann and Winkelmann (1998) have tended to use conditional logit models where a marginal effect for unemployment cannot be estimated, and even where marginal effects can be computed (e.g., Korpi's (1997) OLS fixed effects model), a clearly defined treatment effect does not exist because the regression estimates are conditional on variables taken from the whole sample. If we use the term 'treatment' for the state of unemployment, then the study presented in this chapter will derive the sample average treatment effect (ATE)[7] of unemployment with clear implications for labour market policy and for how we interpret the results for 3S-WV.

[7] Although unemployment has a negative connotation I will use the term 'treatment' to align with the treatments effects literature.

Fourth, I analyse potential *moderating* factors in the unemployment-SWB relationship. The main analysis (as described above) will estimate the full impact of unemployment on life satisfaction, but I am also interested in breaking down the sample average effect to assess whether the magnitude of unemployment impacts (and hence the value of employment) differs by demographic and socio-economic group as defined by gender, marital status, age, and educational attainment. Since we have an exogenous change in employment status we can assess the impact of these moderating factors on life satisfaction in a causal framework.

These differences can be seen as some of the unique contributions associated with this chapter of the thesis.

6.3.2. Data

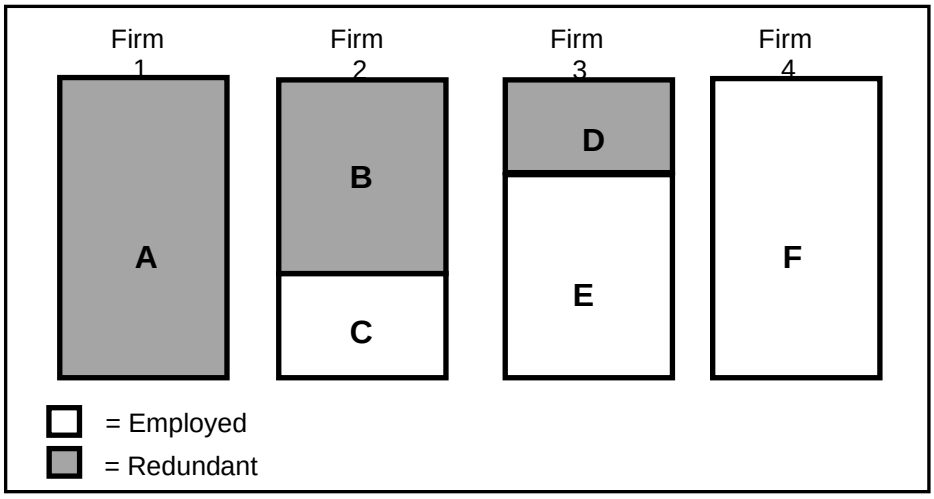
The analysis uses the British Household Panel Survey (BHPS) and the life satisfaction question: "*How dissatisfied or satisfied are you with your life overall?"* Life satisfaction was added in 1997 and so I analyse the period 1997- 2009, excluding 2001 which did not include life satisfaction.

In the BHPS, respondents are asked about their previous three job moves during the sampling year. They can select one of 13 different options for terminating their previous employment: *'promoted', 'left for better job', 'made redundant', 'dismissed or sacked', 'temporary job ended, 'took retirement', 'stopped health reasons', 'left to have baby', 'children/home care', 'care of other person', 'moved area', 'started college/university'* and *'other reason'*. In the data I look at only the most recent job move and define someone to be 'redundant unemployed' if they were made redundant from their previous job and are still unemployed now (when the survey was undertaken). As far as this author is aware this is the first time that the redundancy data in the BHPS has been used in analysis of unemployment and SWB.

6.3.2.1. Redundancy and exogenous unemployment

In the BHPS we may observe redundancies for people from a wide range of previous jobs and industries. Figure 6 describes the typical types of groups we might observe in the data at any point in time. The shaded group are the people who have been made redundant and continue to be unemployed at the time of the survey and the non-shaded group are people that kept their jobs and did not quit for other reasons.

Figure 6: Potential patterns of redundancy across different types of organisations



In the example we have four different firms or organisations, where Firm 1 stops trading and makes everyone redundant, Firm 4 runs business as usual and makes no redundancies and Firms 2 and 3 make different levels of redundancy. If we think of unemployment being the 'treatment' then the treated group consists of A+B+D and the control group consists of C+E+F. It may be the case that the characteristics of group (A+B+D) differ from those of group (C+E+F) before redundancy, which would result in biased causal estimates if these characteristics also impact on wellbeing.

It could be argued that redundancy is more likely for some groups in companies like Firm 2 and Firm 3. For example, less productive and less motivated people, those with caring duties or other commitments and those in poor health may be more likely to be made redundant. Furthermore, there could be reverse causality in the sense that

the less 'happy' or less ‘satisfied’ may be more likely to be made redundant. In this case groups C and E will be distinct from groups B and D.

There may also be differences across Firm 1 and Firm 4 if risk-averse people tend to select into the more established Firm 4 (think of Firm 1 being the high-risk blue-chip and Firm 4 being the public sector) or if more productive people selected into Firm 4 (hence it survived). This would mean that groups A and F would also be considerably different from each other making direct comparison between the treatment group (A+B+D) and control group (C+E+F) impossible. In effect we might have no area of common support between the two groups.

The way that the question about previous jobs is asked in the BHPS could help to solve for this because in the question itself people can state that they finished their last job because of health reasons or caring duties or because they were sacked. Therefore, if the job termination question is answered properly this would help to remove the carers and the less healthy, less productive and unmotivated people from the redundancy group, which would help to increase the exogeneity of the redundancy variable. It is of note that the structure of the job termination question in the BHPS may be more helpful to the present enquiry (in terms of creating an exogenous unemployment variable) than the equivalent question in the GSOEP, since in the GSOEP respondents cannot say that they quit for health or caring reasons.

We have hypothesised that if redundancy were not random groups (A+B+D) and (C+E+F) may differ in pre-redundancy variables like productivity, motivation, health, life satisfaction and risk aversion. Although some of these variables are not observable, Table 8 shows the results from balance tests for redundancy and for unemployment. The balance tests use the following regressions to test whether pre-unemployment factors are associated with becoming unemployed:

(29) $RU_{it} = f(F_{it})$

(30) $U_{it} = g(F_{it})$

The models are run using logit models. RU_{it} is a dummy variable which equals 1 if individual i has been made 'redundant unemployed' in time t (and was employed at time $t-1$); and equals 0 if the individual is employed continuously at times $t-1$ and t. U_{it} is a dummy variable which equals 1 if individual i has become unemployed (for any reason) at time t (and was employed at time $t-1$), and equals 0 if the individual is employed continuously at times t-1 and t.

F_{it} is a vector of pre-redundancy/pre-unemployment factors covering 11 different pre-redundancy/unemployment characteristics: life satisfaction; job satisfaction; satisfaction with leisure time; health satisfaction; satisfaction with social life; caring duties; household size; age; average number of hours worked; level of education; medical services usage. These factors were chosen as all of them would intuitively and plausibly be associated with the likelihood of falling into unemployment.

The hypothesis is that redundant unemployed (RU_{it}) is exogenous and in which case the coefficients on F_{it} in equation (29) should be statistically insignificant. This would show that these 11 pre-redundancy variables do not differ between those who go on to become unemployed and those that stay employed at baseline (ie, before redundancy). It is expected that the coefficients on F_{it} in equation (30) will be statistically significant for some variables at least as the general unemployment variable will be endogenous.

This kind of test replicates a typical *sample balance* test that is usually conducted with randomised trials to check whether the randomisation has been carried out effectively.

The descriptions and descriptive statistics of these variables can be found in Table 2.

The results, presented in Table 8, are very encouraging. First, looking at Panel (i) none of the pre-redundancy variables are statistically significant. Average number of hours worked has a p-value of 9%, but it is not significant at the 5% level. This states that there are no statistical differences in these variables between people that become redundant and those that stay in employment before redundancy. Factors that we would usually be concerned with, such as health, caring duties, age, job satisfaction

and education levels, are statistically independent of the redundancy decision. I note that the main outcome variable, life satisfaction, does not differ between the redundant unemployed and employed groups in the period before redundancy.

The analysis covers a wide range of important pre-redundancy factors and it would suggest that since these factors are not significantly different other factors measured at baseline (if we were to observe them) are also unlikely to be significantly different. The results we find for the balance tests for the redundant unemployed are typical of what one would expect from a setting where the treatment (employment status) is exogenous and hence these results suggest that we can be confident that redundancy (as defined and measured in the BHPS) is for all intents and purposes exogenously determined and hence analysing differences in wellbeing between the redundant unemployed group and the employed group is likely to derive estimates with a robust causal interpretation.

Second, a number of these factors are statistically significant in equation (30) for people who become unemployed. Lower job satisfaction, lower age and lower levels of education are all associated with a greater probability of falling into unemployment as we would expect intuitively. Importantly job satisfaction and education levels were lower for people that go on to become unemployed. This is as we would expect from a non-exogenous unemployment variable. It suggests that if we were to collect further pre-unemployment data we would probably find a large number of factors that differ between people who become unemployed and people who stay in employment at baseline.

Since these factors are also important determinants of life satisfaction then this strongly confirms that a simply measured unemployment variable is likely to be endogenous in a life satisfaction model, meaning that we have to apply considerable caution to the literature that uses non-exogenous employment variables.

Table 8: Balance tests for redundancy and unemployment

Independent variable	Panel (i) Redundant unemployed		Panel (ii) Unemployed	
	Coefficient	Standard error	Coefficient	Standard error
lag life satisfaction	-0.061	0.125	-0.047	0.07
lag job satisfaction	-0.047	0.076	-0.086**	0.042
lag leisure satisfaction	-0.122	0.108	-0.044	0.06
lag health satisfaction	-0.001	0.094	-0.086*	0.052
lag social life satisfaction	0.057	0.114	0.041	0.063
lag caring duties	0.421	0.532	0.413	0.311
hhsize	0.089	0.085	0.056	0.048
lag age	0.008	0.01	-0.028***	0.006
lag job hours	-0.017*	0.01	-0.009	0.006
lag low education	0.307	0.232	0.753***	0.138
lag GP visits	-0.040	0.111	0.040	0.061
Constant	-2.964***	0.935	-1.023*	0.532
Observations	3,492		3,676	
Pseudo R-squared	0.016		0.043	

*Notes: Dependent variable in Panel (i): redundant unemployed. Dependent variable in Panel (ii): unemployed. Variable descriptions in Table 4. *** Significance at 0.01 level; ** Significance at 0.05 level; * Significance at 0.1 level. Variables are measured in the period (year) preceding redundancy or unemployment when both groups are still employed.*

6.3.3. Econometric methods

Given the strong evidence of exogeneity of redundancy (see Table 8) it is possible to estimate the causal effect on life satisfaction by a simple difference-in-means estimator (t-test) between the employed and redundant unemployed groups. However, here I prefer to use ordinary least squares (OLS), which (in a reduced form model) provides the exact same results as the difference-in-means estimator, because it allows us to include other controls in the model to run a further check on the exogeneity of the redundant unemployed variable. If redundancy is truly exogenous then OLS estimates of the effect of unemployment on life satisfaction should not be affected by the inclusion of other control variables in the model provided that they are not measured post-redundancy.

In order to assess the internal validity of the results I run two different models in stepwise fashion:

(31) $$LS_{it} = \alpha + \beta_1 RU_{it} + \beta_2 \ln(M_{it}) + \beta_3 Emp\_status_{it} + \varepsilon_{it}$$

(32) $$LS_{it} = \alpha + \beta_1 RU_{it} + \beta_2 \ln(M_{it}) + \beta_3 Emp\_status_{it} + \beta_4 X_{it} + \varepsilon_{it}$$

where LS_{it} is life satisfaction at time t for individual i, RU_{it} is a dummy variable which equals 1 if the individual is 'redundant unemployed' and equals 0 otherwise. In Emp_status_{it} I control for all other employment status categories ('self-employed'; 'student'; 'retired'; 'maternity leave'; 'long term sick and disabled'; 'government training schemes'; and 'other job status'). This is done in order to make the employed the reference group for the analysis. M_{it} is household income and X_{it} is a vector of other (pre-redundancy) covariates. In other words, I start with the reduced form model (31) and add to it a set of other determinants of life satisfaction in equation (32). If the redundant unemployed variable is exogenous we would expect there to be little change in the size of the coefficients on RU_{it} across the two models.

Income is included for two reasons: (i) because I want to focus on the non-pecuniary impacts of unemployment over and above the loss in wage income and (ii) redundancy is likely to come with a financial package and so for a period of time people may be unemployed but their household income may not have changed. In this scenario being unemployed may not have such a negative impact on wellbeing and so it will be important to control for income in all of the models.

To assess the moderating factors in the relationship between unemployment and life satisfaction I run an additional set of models where the 'redundant unemployed' variable is interacted with some socio-demographic factors:

(33) $$LS_{it} = \alpha + \beta_1 RU_{it} + \beta_2 \ln(M_{it}) + \beta_3 Emp\_status_{it} + \beta_4 MOD_{it} \cdot RU_{it} + \beta_5 X_{it} + \varepsilon_{it}$$

where mod_{it} is the moderating variable and β_4 tells us whether there are heterogenous impacts from unemployment. Since RU is exogenous then the estimates

of the impacts of the moderating factors (β_4) will have a causal interpretation as well. The four categories that I assess in mod_{it} are:

i. Gender
ii. Age (under 30/30-50/over 50)
iii. Educational attainment (degree/no degree)
iv. Marital status (married/not married)

These categories were chosen out of interest as they test a number of hypotheses and intuitions. If the traditional male-female responsibilities still exist then we would expect to see men suffer more from unemployment as expectations regarding work would be higher for them (Warr, 1994; Warr, 1987; Waddell and Burton, 2006). Over the life course people will face different levels of responsibility and hence age interactions should show this (Waddell and Burton, 2006). Note here that the age interaction would cover parental responsibilities and so parental status has not been included as a separate interactive factor in mod_{it}. Unemployment may have a smaller negative impact for more highly educated people because they may feel that they can find another job again quickly. On the other hand, unemployment may be less worse for the less educated if they generally tend to be employed in less enjoyable jobs or if the stigma of unemployment is low due to higher prevalence of unemployment in their reference/peer groups (Waddell and Burton, 2006; Warr, 1987). Finally, unemployment may be less detrimental to wellbeing if the unemployed person is married and hence can receive support (Waddell and Burton, 2006). Note that all models still control for household income meaning that any interactive effect that we find to be significant will represent a heterogenous effect of unemployment over and above any impact on household finances.

Table 9 shows the descriptive statistics for the variables used in the models.

Table 9: Descriptive statistics

Variable	Description	Mean	Std. Dev.	Min	Max
Life satisfaction	Satisfaction with life on a scale of 1-7	5.16	1.21	1	7

Health satisfaction	Satisfaction with own health on a scale of 1-7	5.17	1.43	1	7
Social satisfaction	Satisfaction with social life on a scale of 1-7	4.92	1.43	1	7
Spouse satisfaction	Satisfaction with spouse/partner on a scale of 1-7	4.16	3.48	1	7
Leisure satisfaction	Satisfaction with use of leisure time on a scale of 1-7	4.74	1.46	1	7
Redundant unemployed	= 1 if made redundant from previous job and still unemployed; 0 = otherwise	0.02	0.14	0	1
Unemployed	=1 if unemployed; 0 = employed or self-employed	0.11	0.31	0	1
Log (household income)	Logarithm of annual household income	9.98	6.68	1.61	14.06
Age	Age of respondent	33.53	11.80	16	78
Married	= 1 if married; 0 = otherwise	0.41	0.49	0	1
Male	=1 if male; 0 = otherwise	0.50	0.50	0	1
Carer	= 1 if cares for someone; 0 = otherwise	0.04	0.19	0	1
Non-degree	= 1 if less than university-level education; 0 = otherwise	0.50	0.50	0	1
Wales	= 1 if lives in Wales; 0 = otherwise	0.13	0.37	0	1
Scotland	= 1 if lives in Scotland; 0 = otherwise	0.16	0.37	0	1
N. Ireland	= 1 if lives in Northern Ireland; 0 = otherwise	0.08	0.27	0	1
Live in safe area	= 1 if feels that living in a safe area; 0 = otherwise	0.82	0.38	0	1
Debt burden	= 1 if feels burdened with debt; 0 = otherwise	0.40	0.49	0	1
House owned	= 1 if house owned; 0 = otherwise	0.70	0.46	0	1
Number of children	Number of children in the household	0.60	0.96	0	8
Wave	Year of the interview	N/A	N/A	N/A	N/A

6.4. Results

6.4.1. The causal effect of unemployment on life satisfaction

The results from the two models are set out in Table 10. Column (I) shows the reduced form model (31) with employment status variables and income as the only independent variables. It shows that the non-pecuniary aspect of being unemployed reduces life satisfaction by -0.352 points (95% confidence interval: -0.468 to -0.236) on a scale of 1-7 compared to being employed. Column (II) presents equation (30) with a set of pre-redundancy control variables added to the model. In this model I provide further evidence of the exogeneity of the redundant unemployed variable by showing that the coefficient for RU does not change significantly from the result in the reduced form model (-0.352). The important point here is not to include post-redundancy variables, such as current health status, in (30). This would in effect control for the mediating factors associated with unemployment and would impact on the coefficient on RU. Therefore in column (II) the aim was to include a set of lagged variables for the main determinants of life satisfaction, such as health. However, due to the nature of the BHPS dataset – which is not a balanced panel (many individuals do not have complete year-on-year records in the panel and will hence fall out of the regression model) – there was a large loss in sample size when using lagged variables. The sample size fell from over 26,000 observations in the reduced form model to around only 5,000 observations when adding lagged variables to the reduced form model. This significantly changed the composition of the sample which would make the results from models (31) and (32) incomparable.

Since redundancy has already happened when the survey is administered (because the redundancy variable comes from historic data about moves from previous jobs), then any variable measured in the survey at the time has the potential of being a post-redundancy variable. Therefore, in model (32) I restrict the vector X_{it} to a set of variables that would not change as a consequence of the first year of being made unemployed. These are gender, age, education level, region and number of children. These variables are entered for the current time period t rather than for the period $t-1$ in equation (32).

It should be noted that the fact that I have not been able to control for *all* other determinants of life satisfaction should not be seen as a serious deficiency of the methodology since all I am trying to do in model (30) (column (II)) is provide some further support that the redundant unemployed variable is to all intents and purposes exogenous to life satisfaction as Table 8 has already strongly demonstrated.

Table 10: The causal effect of unemployment on life satisfaction

Independent variable	**(I)**		**(II)**	
	Coefficient	**Standard error**	**Coefficient**	**Standard error**
Redundant unemployed	-0.352***	0.059	-0.337***	0.059
Log (household income)	0.153***	0.011	0.173***	0.012
Self-employed	0.118***	0.031	0.144***	0.032
Retired	0.385***	0.038	0.137***	0.049
Student	0.011	0.04	-0.002	0.041
Maternity leave	0.397***	0.087	0.379***	0.089
Sick leave	-1.078***	0.083	-1.076***	0.084
Government training	-0.092	0.142	-0.149	0.145
Other job status	-0.165*	0.09	-0.188**	0.092
Male			-0.031**	0.015
Age			-0.048***	0.004
Age-squared			0.001***	0
Non-degree			-0.020	0.016
Wales			-0.026	0.023
Scotland			-0.063***	0.021
N. Ireland			0.134***	0.029
Number of children			0.034***	0.009
Constant	3.634***	0.111	4.269***	0.126
Observations	26,849		26,394	
R-squared	0.030		0.036	

Notes: Dependent variable: life satisfaction. Model in column (I) estimated using equation (31). Model in column (II) estimated using equation (32). Variable descriptions in Table 9.
**** significance at 0.01 level; ** significance at 0.05 level; * significance at 0.10 level. Heteroscedasticity-robust standard errors. Employed is the reference case for the employment status variables.*

The coefficient on RU remains stable and robust to the inclusion of the other control variables in equation (32). Adding the set of control variables to the reduced form model in column (I) marginally shifts the coefficient on 'redundant unemployed' from -0.352 to -0.337 which comfortably lies within the 95% confidence interval for the coefficient on 'redundant unemployed' from the reduced form model. There is, therefore, no statistical difference between the coefficients on RU across the two models even after including a range of control variables. This is as we would expect if 'redundant unemployed' is exogenous to life satisfaction as I have argued in this chapter.

An interesting question that is sometimes posed in studies like this is whether statistical methods that impose a selection on observables assumption with endogenous variables, like regression analysis or matching techniques, can replicate the results from a model with exogenous variables. And if they cannot this is generally seen to provide further support for the models that use the exogenous variable. However, although such an analysis is clearly intuitively appealing it may not necessarily shed much important light on the analysis here. This is because when working with observational (ie, non-experimental) data we invariably will include some control variables in our models that are measured post-treatment (or post-unemployment in our case). This means that we could have conflicting biases. For instance, in an SWB model the unemployment variable may be biased upwards because people who are likely to fall into unemployment are the ones that would be unhappy anyway. But at the same time we may control for post-unemployment health status, which would then serve to reduce the unemployment coefficient, offsetting the endogeneity bias to some extent. By pure chance there will be occasions when the offsetting effects balance out to leave a coefficient estimated under a selection on observables assumption, using say regression analysis, to mimic the robust causal effect we derive from a study where the variable of interest is exogenous, but clearly it would be wrong in such a circumstance to conclude that conditioning on a set of control variables ensures exogeneity in the variable of interest. Under multivariate analysis it is hard to keep track of all of these possible conflicting effects making it difficult to compare the methods.

Nevertheless, for completeness I present here estimates derived from OLS with a standard unemployment variable and a full set of control variables. Column (I) in Table 11 shows the results of a reduced form model using the standard (endogenous) unemployed variable, rather than the 'redundant unemployed' variable. Column (II) presents the OLS results for a typical life satisfaction model which controls for a wide range of factors. The model is representative of the types of models used frequently in the wellbeing literature.

Table 11. Estimating the effect of unemployment on life satisfaction using OLS regression

Independent variable	**(I)**		**(II)**	
	Coefficient	**Standard error**	**Coefficient**	**Standard error**
Unemployed	-0.589***	0.033	-0.517***	0.033
Log (household income)	0.125***	0.011	0.082***	0.011
Self-employed	0.068**	0.031	0.073**	0.031
Retired	0.315***	0.038	0.102**	0.048
Student	-0.065	0.04	-0.030	0.04
Maternity leave	0.334***	0.088	0.391***	0.089
Sick leave	-1.152***	0.083	-0.807***	0.084
Government training	-0.166	0.141	-0.107	0.141
Other job status	-0.229**	0.09	-0.176*	0.091
Male			-0.104***	0.015
Age			-0.068***	0.005
Age-squared			0.001***	0.000
Non-degree			0.014	0.015
Wales			-0.021	0.023
Scotland			-0.031	0.02
N. Ireland			0.163***	0.029
Number of children			-0.032***	0.01
Wave			-0.004*	0.002
Health (GP visits)			-0.154***	0.007
Married			0.112***	0.022
Divorced			-0.363***	0.043

Widowed			-0.377***	0.081
Separated			-0.559***	0.063
Never married			-0.282***	0.023
Live in safe area			0.197***	0.02
Constant	3.967***	0.109	6.015***	0.136
Observations	26,849		26,195	
R-squared	0.045		0.092	

*Notes: Dependent variable: life satisfaction. Model in column (I) estimated using equation (31) and substituting the unemployed variable in place of redundant unemployed. Model in column (II) estimated using equation (32) and substituting the unemployed variable in place of redundant unemployed. Variable descriptions in Table 9. *** significance at 0.01 level; ** significance at 0.05 level; * significance at 0.1 level. Heteroscedasticity-robust standard errors. Employed is the reference case for the employment status variables.*

The coefficient on unemployment in column (I) is *upwardly biased* in comparison to the results from models using the 'redundant unemployed' variable. The coefficient estimate of -0.589 is statistically significant with a 95% confidence interval of -0.654 to -0.524. Once all of the main determinants of life satisfaction have been added to the model the coefficient on unemployment falls to -0.517, with a 95% confidence interval of -0.581 to -0.453.

Note that the reduced form model is not robust to the inclusion of control variables when an endogenous unemployment variable is used. This downward shift in the size of the coefficient on unemployment is as we would expect for an endogenous variable that a-priori is expected to be upward biased. Controlling for a range of other determinants of life satisfaction reduces the negative effect of unemployment on life satisfaction as there are likely to be a host of factors that jointly make people more likely fall out of employment and at the same time less satisfied with their lives as well.

However, even controlling for a wide range of confounding factors that are used in the empirical wellbeing literature does not reduce the size of the coefficient on unemployment to level found in the model using the 'redundant unemployed' variable.

The two models that use an endogenous unemployment variable produce very different estimates of the impact of unemployment on life satisfaction compared to the redundant unemployed model. The 95% confidence intervals for the two models with unemployment do not contain the value of the coefficient on 'redundant unemployed' from Table 11 (-0.352). In this instance the OLS models, therefore, correctly predict the direction of impact for unemployment (negative), but they produce upwardly biased estimates of the causal effect of unemployment on life satisfaction, as we would expect for an endogenous unemployment variable.

In sum, the results from Tables 8, 10 and 11 are stark and important. They all point to the same conclusion: that the evidence suggests that the 'redundant unemployed' variable is exogenous to life satisfaction in the BHPS data. Table 8 showed no significant differences for a number of important variables between the 'redundant unemployed' and 'employed' pre-redundancy. Table 10 showed that in regression analysis the redundant unemployed variable is robust to the inclusion of a range of other control variables. And Table 11 shows that the standard unemployed variable is not robust to the inclusion of other control variables in regression analysis. These results are very encouraging, strongly suggesting that the estimated effect of unemployment on life satisfaction for those made redundant is highly likely to have a causal interpretation.

The preferred model here is the reduced form model in column (I) (equation (29)) as the coefficient on 'redundant unemployed' is equivalent to the non-parametric difference-in-means estimator, which here is averaged over different income-level cells since income is held constant. The reduced form model has a number of desirable properties for the purposes of wellbeing valuation.

First, it eliminates concerns relating to indirect effects. As far as this author is aware this study presents the first reduced form model for unemployment and wellbeing. The results represent the total causal derivative of life satisfaction with respect to unemployment.

Second, moving away from models that require conditioning on other variables allows us to derive treatment effects with a clear interpretation. Where conditioning on other

covariates is necessary, they will be measured from the treated and non-treated populations and as Humphreys (2009) shows, this results in poorly-defined estimators that lie somewhere between the average treatment effect for the treated (ATT) and the average treatment effect for the non-treated (ATNT), where in this case the ATT would be the retrospective effect of unemployment on people who become unemployed and the ATNT would be the prospective effect of unemployment if someone in employment were to be made unemployed. In effect, results from multivariate analyses that require conditioning on a large set of variables cannot tell us whether the estimated impact was for those who fell into unemployment or is what it would be like for someone in employment now to lose their job.

On the other hand, the difference-in-means estimator in the reduced form model (31) tells us the expected effect of unemployment on life satisfaction (over and above the loss in income) for any employed individual chosen at random from the sample (akin to the sample average treatment effect (ATE)[8]): for any employed person in our sample we would expect unemployment to lead to a 0.35 point reduction in life satisfaction and since the BHPS is a nationally representative sample this is a very generalisable result. This has a much clearer interpretation and meaning for policy-making purposes as it allows us to estimate the loss in wellbeing that can be prevented by keeping someone in work. For the purposes of wellbeing valuation the clarity of the estimator allows for sample matching and allows us to make a full interpretation of the wellbeing values.

This estimate should be interpreted as the impact for the first year in unemployment and it is the impact of both entry into unemployment and being unemployed since the 'redundant unemployed' are people that have remained unemployed. There may be adaptation (Frederick and Loewenstein, 1999; Kahneman and Sugden, 2005; Loewenstein and Ubel, 2008) to unemployment in subsequent years, which would reduce the impact on life satisfaction. I ran a second version of equation (29) using a one-year lag for the 'redundant unemployed' variable to test for adaptation effects, but the sample size became too small to provide meaningful results.

[8] Note that it also represents the average treatment effect for the treated, but we shall focus on the ATE interpretation throughout this paper.

6.4.2. Moderating factors in the effects of unemployment on life satisfaction

Column (I) shows the results of the preferred reduced form model (Base model) with 'redundant unemployed' and the other columns incorporate interactive terms separately to account for heterogeneous effects. The effects estimated in Table 12 represent the causal effect of unemployment for groups categorised by gender, age, previous education, and marital status.

Table 12. Moderating effects of unemployment on life satisfaction

Independent variable	Base	Gender	Age	Education	Marital status
Redundant unemployed	-0.352***	-0.315***	-0.194*	-0.342***	-0.529***
	(0.059)	(0.090)	(0.104)	(0.062)	(0.086)
Log (household income)	0.153***	0.153***	0.153***	0.152***	0.152***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Redundant unemployed*male		-0.062			
		(0.118)			
Redundant unemployed*30-49 age			-0.356**		
			(0.139)		
Redundant unemployed*50+ age			-0.075		
			(0.151)		
Redundant unemployed*degree				-0.049	
				(0.203)	
Redundant unemployed*married					0.357***
					(0.116)
Constant	3.634***	3.636***	3.630***	3.637***	3.642***
	(0.111)	(0.111)	(0.111)	(0.112)	(0.111)
N	26,849	26,849	26,849	26,540	26,849
R-squared	0.030	0.029	0.030	0.030	0.030

*Notes: Dependent variable: life satisfaction. Model estimated using equation (33). All other independent variables from Table 10 included as control variables. The following labour force status categories are dropped from the analysis: 'retired', 'maternity leave', 'student', 'long term sick', 'disabled', 'family care', 'government training scheme'. Variable descriptions in Table 9. *** significance at 0.01 level; ** significance at 0.05 level; * significance at 0.1 level. Standard errors in parentheses. Heteroscedasticity-robust standard errors.*

There were no differences in the impact of unemployment on life satisfaction by gender or previous educational attainment. In other words, males and females experience unemployment equally badly as do people with degrees compared to those without university education.

There are, however, some marked differences by age and marital status. People in the 30 to 50 year old age group are impacted on much greater than other age groups. For 30 to 50 year olds, when financial responsibilities are probably at their highest, unemployment has a negative causal effect of -0.55 life satisfaction points. For people under 30 or over 50 years of age the negative impact of unemployment on life satisfaction is -0.194 points.

The results also show that the negative impact of unemployment is softened considerably for married people. The causal effect of unemployment on life satisfaction for married people is -0.172. For the non-married, which includes people who are widowed, separated, divorced and co-habiting, the causal effect of unemployment on life satisfaction is much higher at -0.529.

6.4.3. Summary of findings

This chapter analyses the causal effect of unemployment on life satisfaction using the BHPS dataset for the UK by exploiting random variation in employment status due to redundancy. I provide strong evidence that the redundancy variable in the BHPS is exogenous.

I find that unemployment has a large detrimental causal effect on life satisfaction over and above the loss in wage income and including all the possible channels through which unemployment impacts on wellbeing, such as through health. Unemployment is worse for people of ages 30 to 50 (probably due to larger relative financial commitments during this period in life) and for people who are not married (possibly because a partner offers financial and emotional support for someone made unemployed).

6.5. Estimating the cost of unemployment using Three-Step Wellbeing Valuation

Section 6.4. of this chapter provided a set of results for the non-market good (employment) model (Step 2 of 3S-WV). Chapter 5 provided results for the income model (Step 1 of 3S-WV). This allows me to now undertake Step 3, the final stage of the 3S-WV process. It is useful to recap the findings from Step 1 and Step 2 first.

6.5.1. Step 1: The income model

The lottery wins data permitted estimation of the full causal effect of income on life satisfaction. Since no post-intervention (post-lottery) variables were controlled for (only lottery wins in previous years was included in the model) the estimate from the control function model represents **the total causal derivative** of life satisfaction with respect to income (This adheres to Technical Conditions 1 and 2).

The control function approach permitted estimation of the total causal derivative of income for the general UK population of lottery players. Since most of the population play lotteries in the UK, it can be assumed that the results from the control function model are generalizable to the whole UK. Since the results of the reduced form unemployment model are also representative of the UK using the same BHPS dataset the results from the income and non-market good models can be used together as the two **samples match** (This adheres to Technical Condition 3).

The results of the income model are as follows:

$$\frac{\mathrm{d}\,SWB}{\mathrm{d}\ln(M)} = 1.1$$

6.5.2. Step 2: The non-market good model

The reduced form of the redundant unemployed model produces the **total causal derivative** of life satisfaction with respect to unemployment. It is the effect of unemployment on life satisfaction over and above the financial consequences and

includes all indirect effects of unemployment such as through the adverse effects on health (This adheres to Technical Conditions 1 and 2).

The total causal derivative of unemployment has a **clear interpretation**, as the negative effect of unemployment on life satisfaction for any employed person in the UK. This permits a full interpretation of the meaning of the wellbeing value estimates (This adheres to Technical Condition 4).

There are a number of results from the non-market good model that can be used in the 3S-WV analysis:

(i) Average impact of unemployment: $\frac{\mathrm{d}\,SWB}{\mathrm{d}\,Q} = -0.352$

(ii) Impact of unemployment (under 30 age group): $\frac{\mathrm{d}\,SWB}{\mathrm{d}\,Q} = -0.194$

(ii) Impact of unemployment (30-50 age group): $\frac{\mathrm{d}\,SWB}{\mathrm{d}\,Q} = -0.55$

(iv) Impact of unemployment (over 50 age group): $\frac{\mathrm{d}\,SWB}{\mathrm{d}\,Q} = -0.194$

(v) Impact of unemployment (married people): $\frac{\mathrm{d}\,SWB}{\mathrm{d}\,Q} = -0.172$

(vi) Impact of unemployment (not married people): $\frac{\mathrm{d}\,SWB}{\mathrm{d}\,Q} = -0.529$

6.5.3. Step 3: Deriving wellbeing values for unemployment

I estimate the compensating surplus (CS) and equivalent surplus (ES) for unemployment, which is a welfare loss. The CS for a change in the non-market good (ΔQ) where the impact of ΔQ is negative for welfare can be presented in terms of total derivatives for Q and M:

(34) $$CS = exp\left[\frac{-\frac{\mathrm{d}\,SWB}{\mathrm{d}\,Q}\cdot\Delta Q}{\frac{\mathrm{d}\,SWB}{\mathrm{d}\,M}} + \ln\left(M^0\right)\right] - M^0$$

Similarly the ES for a change in the non-market good (ΔQ) can be derived as follows:

(35) $$ES = M^0 - exp\left[\ln\left(M^0\right) + \frac{\frac{\mathrm{d}\,SWB}{\mathrm{d}\,Q}\cdot\Delta Q}{\frac{\mathrm{d}\,SWB}{\mathrm{d}\,M}}\right]$$

As the welfare impact of ΔQ is negative the CS will represent the amount of money required to compensate people such that their life satisfaction is left unaffected by (the move into) unemployment. As discussed the value relates to the compensation related to the non-pecuniary aspects of employment. Another way of putting it is that the CS is the amount of extra annual household income that would be required in order to keep a randomly chosen employed person just as satisfied with life if he/she were made unemployed (in addition to accounting for the loss of wage income). This is the cost associated with the first year of unemployment only.

The ES represents the amount of money that would have to be taken off someone in employment such that their life satisfaction would reduce to the level that would pertain had they been made unemployed (in addition to accounting for the loss of wage income). This could be translated in some ways as the value of employment as it is the amount of money people would 'pay' to stay in employment (I have put 'pay' in inverted commas here to make clear that since these are not values based on preferences people may not actually be willing to pay this amount). As with the CS, the ES is the value associated with the first year of unemployment only.

The various estimates for $\frac{\mathrm{d}\,SWB}{\mathrm{d}\,Q}$ from section 6.4. can be used in equation (34) when estimating CS. The figure for $\frac{\mathrm{d}\,SWB}{\mathrm{d}\,M}$ will come from the income model estimated in Chapter 5 such that $\frac{\mathrm{d}\,SWB}{\mathrm{d}\,M}$ = 1.1. M^0 is estimated as the sample average level of annual income, which I set to £26,000 here. The higher the level of M^0, the higher the value of Q as discussed in Chapter 4. Finally, ΔQ = 1 since unemployment is a binary variable.

6.5.4. The non-pecuniary costs of unemployment

Table 13 sets out the estimates of CS and ES for the range of employment outcomes estimated in the non-market good model.

Table 13. Values associated with employment status using 3S-WV

Employment outcome	Causal effect (life satisfaction)	Compensating Surplus	Percentage of annual income	Equivalent Surplus	Percentage of annual income
Unemployed	-0.352	£9,805	37.71%	£7,120	27.39%
Unemployed (under 30)	-0.194	£5,015	19.29%	£4,204	16.17%
Unemployed (30-50)	-0.550	£16,867	64.87%	£10,230	39.35%
Unemployed (over 50)	-0.194	£5,015	19.29%	£4,204	16.17%
Unemployed (married)	-0.172	£4,401	16.93%	£3,764	14.48%
Unemployed (not married)	-0.529	£16,056	61.75%	£9,926	38.18%

Notes: Values estimated using results from Table 12. Percentage of annual income calculated as percentage to £26,000.

As discussed in Chapter 4, ES < CS for a given welfare loss due to the curvature of the life satisfaction-to-income function. The ES represents movements *down* the life satisfaction-income function and the curvature of the function (ie, the marginal utility of income) is steeper from *any* point when one moves downwards rather than upwards. Also it should be noted that ES measures for a welfare loss are constrained by income, whilst CS for welfare losses is not.

6.5.4.1. Compensating wellbeing value

The amount of money required to compensate someone for being unemployed for the first year of unemployment such that their life satisfaction stays at the level it would be had they been in employment (CS) is **£9,805** in addition to the loss in wage income. This is equivalent to around 38 per cent of their salary when in employment.

The equivalent annual CS figures for 30-50 year olds (**£16,867**) and for people who are not married (**£16,056**) are higher.

These values are applicable to any person who is employed in the UK.

6.5.4.2. Equivalent wellbeing value

The amount of money that would have to be taken away from someone in employment for a year such that their life satisfaction would fall to the level they would experience in unemployment (ES) is **£ 7,120** in addition to the loss in wage income. This is equivalent to about 27 per cent of their salary.

The equivalent annual ES figures for 30-50 year olds (**£10,230**) and for people who are not married (**£9,926**) are higher.

These values are applicable to any person who is employed in the UK.

6.5.5. Comparing 3S-WV against traditional wellbeing valuation methods

The CS and ES estimates derived in 6.4.1. and 6.4.2. have been estimated using the 3S-WV approach. They come from a methodology that adheres to the four technical criteria of wellbeing valuation – in other words, they are based on robust estimates of the total causal derivatives of the SWB function with respect to the non-market good and income; they adhere to sample matching since all results represent impacts for the general UK population and they have a clear interpretation.

The analysis in this section compares the results derived from 3S-WV with the method that is commonly used in the wellbeing valuation literature.

The most frequently used statistical method in wellbeing valuation is OLS regression with one single life satisfaction function as set out in equation (9) (set out again here).

$$(9) \qquad SWB_i = \alpha + \beta_1 Q_i + \beta_2 \ln(M_i) + \beta_3 X_i + \varepsilon_i$$

Partial derivatives from the single equation model in (9) are used to estimate the value (here CS) of Q as follows:

(15) $$CS = M^0 - exp^{\left[\ln(M^0) - \frac{\beta_1(\Delta Q)}{\beta_2}\right]}$$

Examples of studies that have used this approach are numerous and were discussed in section 3.3.2. .

A typical model for (9) would be the full regression model estimated in column (II) of Table 11. Here β_1 = -0.517 and β_2 = 0.082. With M^0 set to £26,000 as before this leads to implausibly high estimates of value for employment status.

The CS for unemployment is **£14 million** for the first year and the ES for unemployment is right up against the income constraint at **£25,952** for the first year.

This astronomical increase in wellbeing values is being driven by the downward bias in the income coefficient and the upward bias in the unemployment coefficient in the standard approach. Nearly all of the increase, however, is being driven by the difference in the income coefficient in the standard model (if we use the coefficient for unemployment of β_1 = -0.517 in a 3S-WV model where the income model is derived from lottery wins such that β_2 = 1.1 then the CS for unemployment is £15,600, and the ES for unemployment is £9,750, which are upwardly biased estimates of monetary value but the bias is far less significant than the bias observed when the income coefficient also comes from the standard model.

The standard OLS model with an endogenous income variable derives a coefficient for log of household income of 0.082, whereas the control function model with an exogenous income variable derives a coefficient for log of household income of 1.1. The OLS model tells us that a log-point increase in household income which is equivalent to about a 130% increase in income (ie, for a household with an annual income of £50,000 this represents a £65,000 increase in income to a total income of £115,000), is associated with a very trivial increase in life satisfaction of 0.082 index points, which seems much too low intuitively for such a major change in life circumstances. The implausibility of this estimate is even more clear when we see that the same model predicts that the life satisfaction impact of such a large increase in

income is equivalent only to about the impact of a one year increase in age (in absolute terms).

These results are in line with intuition as we would expect certain individuals to be more likely to become unemployed and be less satisfied with life anyway. The bias in the income coefficient is much more severe. The causal estimate derived from lottery winners is more than ten times larger than the OLS estimate. This direction of change is expected given that instrumenting for income generally tends to result in an increase in coefficient size (Pischke, 2010). The size of the income coefficient may have increased using the lottery wins instrument for a number of reasons:

(a) People who would be happy anyway may tend to earn less money;
(b) Income is measured with error;
(c) There are costs associated with earning more money such as more stress and time at work which are not controlled for in OLS regression; and
(d) Many of the indirect effects of income (such as health status) are controlled for in this OLS model.

The upward bias observed in the estimates from the traditional wellbeing valuation method reflects many of the findings from the previous WV literature, which as discussed has generally found wellbeing values to be implausibly too high and magnitudes larger than preference-based valuation methods. The results here suggest that this is being driven by a severe downward bias in the coefficient on income in the traditional wellbeing valuation models.

In addition to this bias I note that there are further problems with the estimates derived from the traditional wellbeing valuation method. The first relates to interpretation. The treatment effect estimates in OLS are vague and as discussed above lie somewhere in between the traditional ATT and ATNT, which prevents us from deriving a clear interpretation of the wellbeing value for policy purposes as I did in 3S-WV. The second issue is sample matching. Since treatment effects are ambiguous in OLS with a selection on observables assumption we cannot be sure whether we are matching similar people when deriving wellbeing values.

In sum, the empirical analysis lends support to the theoretical hypotheses that were generated in Chapter 4 regarding the problems associated with traditional wellbeing valuation methods undertaken with single-equation models. The theoretical analysis suggested that wellbeing values would be over-stated and this was clearly borne out in the analysis. The theory and evidence suggest that the traditional single-equation methods should not be used to value non-market goods in wellbeing valuation.

6.6. Summary

In this chapter I applied the new 3S-WV methodology to value the non-pecuniary benefits of employment. I used redundancy data as a natural experiment to estimate the causal effect of unemployment on life satisfaction. I find that unemployment has a large detrimental causal effect on life satisfaction over and above the loss in wage income and that the negative impact is larger for some groups in society.

Using the 3S-WV approach I find that the value of employment is between £7,120 and £9,850 per person per annum depending on whether we estimate compensating or equivalent surplus. These values seem reasonable and plausible at around a third of national average salary in the UK and are much lower than the values estimated by the current WV methodology. This demonstrates the improvement in estimation and results we get from the new 3S-WV method. I will discuss some of the policy implications as well as possible future avenues of research in the next and final chapter.

Chapter 7

7. Conclusions and policy implications

7.1. Summary of the thesis

Valuing non-market goods, services and outcomes is critical to policy evaluation in the UK and many other countries and increasingly to the social value agenda in the private sector. The HM Treasury Green Book is the core policy evaluation manual for the UK Government and it stipulates that all projects, programmes and policies be evaluated using Cost-Benefit Analysis (CBA). CBA requires the valuation of all outcomes of the policy, including hard to measure non-market outcomes such as education, health, the environment and heritage. Traditionally this has been undertaken using preference-based valuation methods, namely revealed preference and stated preference methods.

However, subjective wellbeing (SWB) data and analysis are topics of growing interest in economics and policy evaluation. One key use for SWB data is in the wellbeing valuation (WV) method to estimate values for non-market goods and services. These values, if estimated robustly, have the potential to be used in CBA in the Green Book and other related policy evaluation frameworks in order to appraise, evaluate and inform policy decisions and investments. Since the data used in WV usually comes from pre-administered national data sets, the method is relatively a lot cheaper and less resource-intensive than the more traditional stated preference method, but at the same time it has a wide application to a large number of sectors and policy areas. WV, therefore, has the potential to transform CBA and how governments and other organisations evaluate their policies.

Since the first application of this method in 2002, there has been a large number of studies that have employed the method to a range of outcome areas to value for example, environment, employment, crime, health, education and natural disasters.

The WV literature to date has been primarily interested in extending the methodology developed in the first studies in 2002 in to other policy areas. There have been no studies devoted entirely to the development of the methodology although some studies have discussed the problems inherent to WV and have offered some relatively minor adjustments in the method. As discussed in this thesis the methodologies currently used in WV are somewhat crude and unfortunately do not derive values according to Hicks' value theory. A number of technical problems lead to bias in many of the estimates to date, rendering the results from the WV methodology not robust enough for use in CBA and policy evaluation. In addition to this, as I have set out in Chapter 3, the interpretation and use of wellbeing values has also been problematic in the literature to date.

The aims of this thesis were to provide for the first time a full assessment of the WV methodology and its current problems and limitations and from that to develop a new and improved theoretical approach and methodology for wellbeing valuation such that it is able to produce robust values in line with economic theory which can be used in CBA and policy evaluation. A final key contribution of the thesis is to provide a more detailed and robust approach to interpreting the values derived from wellbeing valuation.

The core development and contribution of the thesis is the Three Step Wellbeing Valuation approach, which takes a completely different statistical approach to the current WV methodology by separating the estimation process into three different steps, thereby allowing each parameter in the models to be estimated more robustly. The method allows us to derive better estimates of the impacts of income and the non-market good on wellbeing leading to more robust value estimations. The results of the new 3S-WV method in the case study for employment were strong and demonstrated significant benefits over the current WV methods when the results were compared. I generally find that for a given outcome, like pollution or education, values estimated using the new 3S-WV methodology are lower than values estimated using the current WV approach. This is primarily because the impact of income on SWB is significantly higher in the 3S-WV method which means that for any given impact on wellbeing the monetary amount required to have the same effect on the individual is lower as each additional £1 has a higher value. This is a very positive outcome from

the new approach as the key criticism of WV to date has been the tendency for it to produce very high and unrealistic values. The 3S-WV method can be used with any SWB variable and is thus flexible and robust, an improvement on the current methodology.

7.2. Policy implications

I finish with a discussion of the main policy implications of the developments and contributions made in this thesis.

A key advantage of the WV method is its ability to produce values from secondary data sources rather than requiring expensive and time-consuming primary data collection. SP methods always require primary data collection and RP methods also often do so as well. This makes the WV method a highly cost and resource efficient method in comparison to alternative approaches. The key stumbling block to date which I believe has made it difficult and almost impossible to use WV values in policy evaluation has been the size of the values derived in the literature to date using the current methodology, which as discussed in this thesis have generally been seen by most commentators as being unrealistically high.

I have demonstrated that the method developed in this thesis, 3S-WV, provides a robust theoretical approach and is capable of deriving much more plausible and realistic values which are in line with economic theory. The implications of this are significant since it now means that policy makers have access to a highly cost-effective method for valuing goods. A typical stated preference study can cost upwards of £50,000 to conduct. WV values can be estimated at a fraction of that cost at around 10% of the cost of a SP study. A major implication is that policy makers can conduct more valuation studies and will have access to a larger range of values for non-market goods and services to include in CBA and policy evaluation leading to an improvement in the data inputted into these evaluations and the decisions that are made on the basis of it. Currently, it is common in the UK and in many other OECD countries to exclude from the evaluation any policy outcomes that cannot be valued and for there to be a qualitative assessment or discussion of the outcome (see for

example the UK's Department for Transport's WebTAG guidance[9]). This ensures that the issue does not fall out entirely from the assessment but it means that it is unlikely to have any meaningful impact on the policy decision if it is not included in the full CBA assessment.

A second important implication for policy is that the WV approach allows us to value some outcomes that are inherently difficult to value using preference-based methods. This was discussed in detail in Chapter 3. Some outcome areas such as health and the environment are difficult to value using people's preferences, but WV allows us to get around these issues as we do not need to ask anyone their preferences for these types of outcomes, which can be an ethically challenging thing to do. WV, therefore, provides the opportunity now (under this more robust method) to provide values for these areas and to incorporate them fully into the policy evaluation process. For example, the WV method can be used to derive values for different health states and conditions allowing health to be assessed using CBA rather than through the QALY method and cost-effectiveness analysis. This would allow governments to compare health interventions and policy in the same terms as all other policy areas that already use CBA.

A case in point is the specific case study example for employment used in this thesis. It is not possible to use preference-valuation methods to derive a value for the non-pecuniary aspects of employment. Using the new 3S-WV method I have demonstrated how the WV method can be applied to employment outcomes and have derived robust values for the non-pecuniary aspects of employment. To my knowledge this is the first time that such values have been estimated and it allows labour ministries such as the UK Department for Work and Pensions (DWP) to incorporate the wider benefits and value of employment in their CBA assessments. In the UK Government there has been a long history of promoting and advocating the wider non-financial benefits of employment; that employment is important for communities and society and for people's sense of purpose and wellbeing has been a key message across the whole political spectrum. They can now evidence this in their evaluations of different labour market policies and programmes.

[9] https://www.gov.uk/guidance/transport-analysis-guidance-webtag

7.3. Future research

Valuation of non-market goods has relied almost entirely on stated preference and revealed preference methods to date. The new 3S-WV method developed here in this thesis provides a rigorous framework for using SWB data in valuation. There is now an alternative valuation method to sit alongside stated and revealed preference methods. Using the new 3S-WV framework, future research should focus on how we can derive robust estimates of impact for the non-market good/service being valued. It would also be useful to further explore the relationship between income and SWB. This thesis has set out a model for the impact of income on SWB using lottery wins for the UK population. The BHPS is the only large national data set in the UK that contains lottery wins and wellbeing data. The results from the BHPS should be compared to the results we would obtain from other countries using the same econometric method. In due course the lottery wins estimate derived here will need to be re-estimated using more recent data since the BHPS data only go up to 2002. A final interesting area for future research which would impact on the wellbeing valuation methodology would be to get a better understanding of whether the impact on life satisfaction of gains in income differ to losses in income due to loss aversion. This will require exogenous losses in income to be compared to the exogenous gains in income for which we have data from lottery wins. Although this would not impact on the 3S-WV methodology overall, if there are differences then this would mean that we would need to take a different coefficient for income (β_M) in Step 1 of 3S-WV depending on whether we were estimating CS or ES.

References

ACCENT. 2010. Review of Stated Preference and Willingness to Pay Methods. (Prepared for the Competition Commission).

ADLER, M. 2012. *Well-Being and Fair Distribution: Beyond Cost-Benefit Analysis,* Oxford, Oxford University Press.

ADLER, M. & POSNER, E. 2008. Happiness Research and Cost-Benefit Analysis. *Journal of Legal Studies,* 37, 253-292.

AMBREY, C. & FLEMING, C. 2011. Valuing scenic amenity using life satisfaction data. *Ecological Economics,* 72, 106–115.

AMIR, O., ARIELY, D. & CARMON, Z. 2008. The Dissociation Between Monetary Assessment and Predicted Utility. *Marketing Science,* 27**,** 1055–1064.

ANDERSON, E. 1995. *Value in Ethics and Economics,* US, Harvard University Press.

ANGRIST, J. & FERNANDEZ-VAL, I. 2010. ExtrapoLATE-ing: External Validity and Overidentification in the LATE Framework. *NBER Working Paper,* 16566.

ANGRIST, J., IMBENS, G. & RUBIN, D. 1996. Identification of causal effects using instrumental variables. *Journal of American Statistical Association,* 91**,** 444-455.

ANGRIST, J. & PISCHKE, J. 2009. *Mostly Harmless Econometrics,* Princeton, Princeton University Press.

AOSHIMA, I., UCHIDA, K., USHIMARU, A. & SATO, M. 2018. The influence of subjective perceptions on the valuation of green spaces in Japanese urban areas. *Urban Forestry & Urban Greening,* 34**,** 166-174.

APOUEY, B. & CLARK, A. 2010. Winning big but feeling no better? The effect of lottery prizes on physical and mental health. *IZA Discussion Paper*, 4730.

ARIELY, D., LOEWENSTEIN, G. & PRELEC, D. 2003. "Coherent Arbitrariness": Stable demand curves without stable preferences. *The Quarterly Journal of Economics,* 118**,** 73-105.

ARONOW, P. & SOVEY, A. 2010. Beyond LATE: A Simple Method for Recovering Sample Average Treatment Effects. http://polmeth.wustl.edu/media/Paper/aronowlate.pdf.

ARROW, K., SOLOW, R., PORTNEY, P., LEAMER, E., RADNER, R. & SCHUMAN, H. 1993. Report of the NOAA Panel on Contingent Valuation. *NOAA Publication*.

BACKHOUSE, R. 2002. *The Penguin History of Economics,* UK, Penguin.

BARNETT, R. & BRENNAN, R. 1995. The relationship between job experiences and psychological distress: A structural equation approach. *Journal of Organizational Behaviour,* 16**,** 259-276.

BARRINGTON-LEIGH, C. & BEHZADNEJAD, F. 2017. Evaluating the short-term cost of low-level local air pollution: a life satisfaction approach. *Environmental Economics and Policy Studies,* 19**,** 269–298.

BARTLEY, M. 1994. Unemployment and ill health: understanding the relationship. *Journal of Epidemiology and Community Health,* 48**,** 333-337.

BATEMAN, I., BURGESS, D., HUTCHINSON, W. & MATTHEWS, D. 2006. Preference Learning versus Coherent Arbitrariness: NOAA Guidelines or a Learning Design Contingent Valuation. *CSERGE Working Paper*, 06-18.

BATEMAN, I., DAY, B., JONES, A. & JUDE, S. 2009. Reducing gain–loss asymmetry: A virtual reality choice experiment valuing land use change. *Journal of Environmental Economics and Management,* 58**,** 106-118.
BEALE, N. & NETHERCOTT, S. 1987. The health of industrial employees four years after compulsory redundancy. *Journal of the Royal College of General Practitioners,* 37, 390-394.
BECCHETTI, L., CONZO, P. & DI FEBBRARO, M. 2018. The monetary-equivalent effect of voluntary work on mental wellbeing in Europe. *Kyklos,* 71**,** 3-27.
BEJA, E. 2011. Subjective well-being approach to the valuation of income inequality. *MPRA Paper, University Library of Munich, Germany*.
BENJAMIN, D., HEFFETZ, O., KIMBALL, M. & REES-JONES, A. 2012. What do you think would make you happier? What do you think you would choose? *American Economic Review,* 102**,** 2083-2110.
BENTHAM, J. (ed.) 1983. *Deontology, Together with a Table of the Springs of Action and the Article on Utilitarianism,* Oxford, Clarendon Press.
BERRIDGE, K. & KRINGELBACH, M. 2011. Building a neuroscience of pleasure and well-being. *Psychology of Well-Being: Theory, Research and Practice,* 1, 1-3.
BERTRAND, M. & MULLAINATHAN, S. 2001. Do people mean what they say? Implications for subjective survey data. *The American Economic Review,* 91**,** 67-72.
BLANCHFLOWER, D. & OSWALD, A. 2004. Well-being over time in Britain and the USA. *Journal of Public Economics,* 88**,** 1359-1386.
BLEICHRODT, H. 1997. Health utility indices and equity considerations. *Journal of Health Economics,* 16**,** 65-91.
BLUNDELL, R., BROWNING, M. & MEGHIR, C. 1994. Consumer demand and the life-cycle allocation of household expenditures. *Review of Economic Studies,* 61**,** 57-80.
BODENHAUSEN, G. & WYER, J. 1987. *Social cognition and social reality: Information acquisition and use in the laboratory and the real world.* New York, Springer-Verlag.
BOK, S. 2010. *Exploring Happiness. From Aristotle to Brain Science*, Yale University Press.
BOND, T. & LANG, K. 2019. The sad truth about happiness scales. *Journal of Political Economy,* 127, 1629-1640.
BONINGER, D., GLEICHER, F. & STRATHMAN, A. 1994. Counterfactual thinking: From what might have been to what may be. *Journal of Personality and Social Psychology,* 67**,** 297-307.
BRAGA, J. & STARMER, C. 2005. Preference anomalies, preference elicitation and the Discovered Preference Hypothesis. *Environmental and Resource Economics,* 32**,** 55-89.
BRENIG, M. & PROEGER, T. 2016. Putting a price tag on security: Subjective well-being and willingness-to-pay for crime reduction in Europe. *cege Discussion Paper,* 278.
BRICKMAN, P., COATES, D. & JANOFF-BULMAN, R. 1978. Lottery winners and accident victims: Is happiness relative? *Journal of Personality and Social Psychology,* 36**,** 917-927.
BROWN, T. 1984. The Concept of value in resource allocation. *Land Economics,* 60**,** 231-246.

BRUNI, L. & SUGDEN, R. 2007. The road not taken: how psychology was removed from economics, and how it might be brought back. *Economic Journal,* 117**,** 146-173.
BURGARD, S., BRAND, J. & HOUSE, J. 2005. Causation and selection in the relationship of job loss to health in the United States. *Mimeo, University of Michigan.*
CAPLAN, R., VINOKUR, A., PRICE, R. & VAN RYN, M. 1996. Job seeking, reemployment, and mental health: a randomized field experiment in coping with job loss. *Journal of Applied Psychology,* 74**,** 759-769.
CARLSSON, F. 2010. Design of stated preference surveys: is there more to learn from behavioral economics? *Environmental and Resource Economics,* 46**,** 167-177.
CARROLL, N., FRIJTERS, P. & SHIELDS, M. 2009. Quantifying the costs of drought: new evidence from life satisfaction data. *Journal of Population Economics,* 22**,** 445-461.
CARSON, R., FLORES, N. & MEADE, N. 2001. Contingent valuation: controversies and evidence. *Environmental and Resource Economics,* 19**,** 173-210.
CHAMBERLAIN, K. & ZIKA, S. 1992. Stability and change in subjective well-being over short time periods. *Social Indicators Research,* 26**,** 101-117.
CHAMP, P., BOYLE, K. & BROWN, T. 2003. *A primer on nonmarket valuation.* Boston, Kluwer Academic Press.
CHANDOEVWIT, W. & THAMPANISHVONG, K. 2016. Valuing social relationships and improved health condition among the Thai population. *Journal of Happiness Studies,* 17**,** 2167-2189.
CHOI, S., KARIV, S., MUELLER, W. & SILVERMAN, D. 2011. Who is (more) rational? *University of Vienna Working Paper,* 1105.
CLARK, A., FRIJTERS, P. & SHIELDS, M. 2008. Relative income, happiness, and utility: an explanation for the Easterlin Paradox and other puzzles. *Journal of Economic Literature,* 46**,** 95-144.
CLARK, A. & OSWALD, A. 1994. Unhappiness and unemployment. *The Economic Journal,* 104**,** 648-659.
CLARK, A. & OSWALD, A. 2002. A simple statistical method for measuring how life events affect happiness. *International Journal of Epidemiology,* 31**,** 1139-1144.
COHEN, M. 2008. The Effect of crime on life satisfaction. *Journal of Legal Studies,* 37**,** 325-353.
COHEN, S., DOYLE, W., TURNER, R., ALPER, C. & SKONER, D. 2003. Emotional style and susceptibility to the common cold. *Psychosomatic Medicine.,* 65**,** 652–57.
COSTA, P. & MCCRAE, R. 1980. Influence of extraversion and neuroticism on subjective well-being: happy and unhappy people. *Journal of Personality and Social Psychology,* 38**,** 668-678.
COWELL, F. & GARDINER, K. 1999. Welfare weights. . *London School of Economics, STICERD Paper,* 20.
CSIKSZENTMIHALYI, M. 1990. *Flow: the psychology of optimal experience,* New York, Harper and Row.
CUMMINGS, R. & TAYLOR, L. 1999. Unbiased value estimates for environmental goods: a cheap talk design for the contingent valuation method. *American Economic Review,* 89**,** 649-665.

DANZER, A. & DANZER, N. 2011. The long-term effects of the Chernobyl catastrophe on subjective well-being and mental health. *IZA Discussion Paper,* 5906.
DAVIDSON, R. 2004. Well-being and affective style: Neural substrates and biobehavioral correlates. *Philosophical Transactions of the Royal Society,* 359 1395-1411.
DE JONGE, J., DORMANN, C., JANSSEN, P., DOLLARD, M., LANDEWEERD, J. & NIJHUIS, F. 2001. Testing reciprocal relationships between job characteristics and psychological well-being: a cross-lagged structural equation model. *Journal of Occupational and Organizational Psychology,* 74**,** 29-46.
DE JONGE, J. & SCHAUFELI, W. 1998. Job characteristics and employee well-being: a test of Warr's Vitamin Model in health care workers using structural equation modelling. *Journal of Organizational Behaviour,* 19**,** 387-407.
DEATON, A. (ed.) 1993. *Data and Econometric Tools for Development Analysis*: Elsevier.
DEL SAZ-SALAZAR, S., NAVARRETE-TUDELA, A., ALCALÁ-MELLADO, J. & DEL SAZ-SALAZAR, D. 2019. On the use of life satisfaction data for valuing cultural goods: a first attempt and a comparison with the contingent valuation method. *Journal of Happiness Studies,* 20**,** 119-140.
DESVOUSGES, W., JOHNSON, F., DUNFORD, R., BOYLE, K., HUDSON, S. & WILSON, N. 1992. *Measuring Nonuse Damages Using Contingent Valuation: An Experimental Evaluation of Accuracy*. RTI Press.
DI TELLA, R. & MACCULLOCH, R. 2006. Some uses of happiness data in economics. *The Journal of Economic Perspectives,* 20**,** 25-46.
DIENER, E. 1984. Subjective well-being. *Psychological Bulletin,* 95**,** 542-575.
DIENER, E. 1994. Assessing subjective well-being: Progress and opportunities. *Social Indicators Research,* 31**,** 103-157.
DIENER, E., LUCAS, R., SCHIMMACK, U. & HELLIWELL, J. 2009. *Well-Being for Public Policy*. Oxford Scholarship Online.
DIENER, E. & SELIGMAN, M. 2004. Beyond money: toward an economy of well-being. *Psychological Science in the Public Interest,* 5**,** 1-31.
DIENER, E. & SUH, E. 1997. Measuring quality of life: economic, social, and subjective indicators. *Social Indicators Research,* 40**,** 189-216.
DIENER, E., SUH, E., LUCAS, R. & SMITH, H. 1999. Subjective well-being: three decades of progress. *Psychological Bulletin,* 125**,** 276-302.
DOLAN, P. 1998. The measurement of individual utility and social welfare *Journal of Health Economics,* 17**,** 39-52.
DOLAN, P. 2014. *Happiness by Design*. UK, Penguin.
DOLAN, P. & FUJIWARA, D. 2012. Valuing adult learning: comparing wellbeing valuation to contingent valuation. *BIS Research Paper,* 85.
DOLAN, P. & KAHNEMAN, D. 2008. Interpretations of utility and their implications for the valuation of health. *Economic Journal,* 118**,** 215-234.
DOLAN, P., KAVETSOS, G., KREKEL, C., MAVRIDIS, D., METCALFE, R., SENIK, C., SZYMANSKI, S. & ZIEBARTH, N. 2019. Quantifying the intangible impact of the Olympics using subjective well-being data. *Journal of Public Economics,* 177**,** 104043.
DOLAN, P., LAYARD, R. & METCALFE, R. 2011. Measuring subjective wellbeing for public policy: recommendations on measures. . *Centre for Economic*

Performance, London School of Economics and Political Science. Special Paper.
DOLAN, P. & METCALFE, R. 2011. Comparing measures of subjective wellbeing and views about the role they should play in policy. *UK Office for National Statistics Paper.*
DOLAN, P. & WHITE, M. 2007. How can measures of subjective well-being be used to inform public policy? *Perspectives on Psychological Science,* 2**,** 71-85.
DORRIS, M. & GLIMCHER, P. 2004. Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron,* 44**,** 365-78.
DUBORG, R., Jones-Lee M. & Loomes, G. 1997. Imprecise preference and survey design. *Economica,* 64**,** 681-702.
DUNCAN, G. 2010. Should happiness-maximization be the goal of government? *Journal of Happiness Studies,* 11**,** 163-178.
EASTERLIN, R. 1974. *Does economic growth improve the human lot? Some empirical evidence. ,* New York, Academic Press.
EASTERLIN, R. 1995. Will raising the incomes of all increase the happiness of all? *Journal of Economic Behavior & Organization,* 27**,** 35-47.
EASTERLIN, R. 2004. The economics of happiness. *Daedalus,* 133**,** 26-33.
EHRHARDT, J., SARIS, W. & VEENHOVEN, R. 2000. Stability of life satisfaction over time. *Journal of Happiness Studies,* 1**,** 177-205.
EID, M. & DIENER, E. 2004. Global judgments of subjective well-being: situational variability and long-term stability. *Social Indicators Research,* 65**,** 245-277.
EKMAN, P., DAVIDSON, R. & FRIESEN, W. 1990. The Duchenne Smile: emotional expression and brain physiology. *Journal of Personality and Social Psychology,* 58**,** 342-53.
ERDOGAN, B., BAUER, T., TRUXILLO, D. & MANSFIELD, L. 2012. Whistle while you work: a review of the life satisfaction literature. *Journal of Management.* 38**,** 1038-1083.
EVANS, D., KULA, E. & SEZER, H. 2005. Regional welfare weights for the UK: England, Scotland, Wales and Northern Ireland. *Regional Studies,* 39**,** 923-937.
EVANS, D. & SEZER, H. 2002. A time preference measure of the social discount rate for the UK. *Applied Economics,* 34**,** 1925-1934.
FEDDERSON, J., METCALFE, R. & WOODEN, M. 2012. Subjective well-being: weather matters; climate doesn't. *University of Oxford. Department of Economics Working Paper Series*.
FERNANDEZ, C., RAITZER, D. & GINTING, E. 2019. Potential use of the life satisfaction approach to value nonmarket goods and services. *ADB Economics Working Paper Series,* 569.
FERREIRA, S. & MORO, M. 2009. On the use of subjective well-being data for environmental valuation. *University of Stirling Economics Discussion Paper 2009-24*.
FERRER-I-CARBONELL, A. & VAN PRAAG, B. 2002. The subjective costs of health losses due to chronic diseases. An alternative model for monetary appraisal. *Health Economics,* 11**,** 709-722.
FOLLMANN, D. A. 2000. On the effect of treatment among would-be treatment compliers: an analysis of the multiple risk factor intervention trial. *Journal of the American Statistical Association,* 95**,** 1101-1109.

FREDERICK, S. & LOEWENSTEIN, G. 1999. *Hedonic Adaptation.* New York, Russell Sage Foundation.
FREY, B., LUECHINGER, S. & STUTZER, A. 2007. Calculating tragedy: assessing the costs of terrorism. *Journal of Economic Surveys,* 21**,** 1-24.
FREY, B., BENZ, M. & STUTZER, A. 2004. Introducing procedural utility: not only what, but also how matters. *Journal of Institutional and Theoretical Economics,* 160**,** 377-401.
FREY, B., LUECHINGER, S. & STUTZER, A. 2009. Valuing public goods: the life satisfaction approach. *Public Choice,* 138**,** 317–345.
FREY, B., OBERHOLZER-GEE, F. & EICHENBERGER, R. 1996. The old lady visits your backyard: a tale of morals and markets. *Journal of Political Economy,* 104**,** 1297-1313.
FREY, B. & STUTZER, A. 2005. Happiness research: state and prospects. *University of Zurich Working Paper*.
FRIJTERS, P. 2000. Do individuals try to maximize general satisfaction? *Journal of Economic Psychology,* 21**,** 281-304.
FUJITA, F. & DIENER, E. 2005. Life satisfaction set point: stability and change. *Journal of Personality and Social Psychology,* 88**,** 158-164.
FUJIWARA, D. & CAMPBELL, R. 2011. Valuation techniques for social cost-benefit analysis: stated preference, revealed preference and subjective well-being approaches. *Department for Work and Pensions and HM Treasury Discussion Paper.*
FUJIWARA, D., KUDRNA, L. & DOLAN, P. 2014. Quantifying and valuing the well-being impacts of culture and sport. *Department for Culture, Media and Sport Research Paper*.
GANGL, M. 2010. Causal inference in sociological research. *Annual Review of Sociology,* 36**,** 21-47.
GARDNER, J. & OSWALD, A. 2007. Money and mental wellbeing: a longitudinal study of medium-sized lottery wins. *Journal of Health Economics,* 26**,** 49-60.
GENESOVE, D. & MAYER, C. 2001. Loss aversion and seller behavior: evidence from the housing market. *The Quarterly Journal of Economics,* 116**,** 1233-1260.
GERLACH, K. & STEPHAN, G. 1996. A paper on unhappiness and unemployment in Germany. *Economics Letters,* 52**,** 325-330.
GILBERT, D. 2007. *Stumbling on happiness,* London, Harper Perennial.
GIOVANIS, E. & OZDAMAR, O. 2016. Structural equation modelling and the causal effect of permanent income on life satisfaction: the case of air pollution valuation in Switzerland. *Journal of Economic Surveys,* 30**,** 430-459.
GLIMCHER, P. 2010. *Foundations of neuroeconomic analysis.* Oxford, Oxford University Press.
GREENBERG, D. & KNIGHT, G. 2007. Review of the DWP Cost Benefit Framework and how it has been applied. *University of Westminster Research Paper*.
GREENBERG, D. & ROBINS, P. 2008. Incorporating non market time into benefit cost analyses of social programs: an application to the self sufficiency project. *Journal of Public Economics,* 92**,** 766-794.
GREENBERG, D. 1997. The leisure bias in cost-benefit analyses of employment and training programs. *The Journal of Human Resources,* 32**,** 413-439.
GREGORY, R., LICHTENSTEIN, S. & SLOVIC, P. 1993. Valuing environmental resources: a constructive approach. *Journal of Risk & Uncertainty,* 7**,** 177-197.

GRETHER, D. & PLOTT, C. 1979. Economic theory of choice and the preference reversal phenomenon. *The American Economic Review,* 69**,** 623-638.
GROOT, W. & VAN DEN BRINK, H. 2006. The compensating income variation of cardiovascular disease. *Health Economics Letters,* 15**,** 1143-1148.
GUL, F. & PESENDORFER, W. 2008. *The Case for Mindless Economics*: Oxford University Press.
HAMMOND, P. 1991. *Interpersonal comparisons of utility: why and how they are and should be made.* Cambridge, Cambridge University Press.
HANLEY, N. & SHOGREN, J. 2005. Is cost–benefit analysis anomaly-proof? *Environmental and Resource Economics,* 32**,** 13-24.
HARSANYI, J. 1955. Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy,* 63**,** 309-321.
HASTIE, R. & DAWES, R. 2010. *Rational choice in an uncertain world.* London, Sage.
HAUSMAN, D. & MCPHERSON, M. 2006. *Economic analysis, moral philosophy, and public policy.* New York, Cambridge University Press.
HAYBRON, D. 2007. Life satisfaction, ethical reflection, and the science of happiness. *Journal of Happiness Studies,* 8**,** 99-138.
HAYBRON, D. 2010. *The pursuit of unhappiness: the elusive psychology of well-being.* Oxford, Oxford University Press.
HAYBRON, D. 2000. Two philosophical problems in the study of happiness. *Journal of Happiness Studies,* 1**,** 207-225.
HECKMAN, J. & VYTLACIL, E. 1998. Instrumental variables methods for the correlated random coefficient model: estimating the average rate of return to schooling when the return is correlated with schooling. *The Journal of Human Resources,* 33, 974–987.
HEEGER, D. 1992. Normalization of cell responses in cat striate cortex. *Visual Neuroscience,* 9**,** 181-197.
HELLER, D., WATSON, D. & ILIES, R. 2004. The role of person versus situation in life satisfaction: a critical examination. *Psychological Bulletin,* 130**,** 574-600.
HELLIWELL, J. & HUANG, H. 2005. How's the job? Wellbeing and social capital in the workplace. *NBER Working Paper Series*.
HENDRA, R., RAY, K., VEGERIS, S., HEVENSTONE, D. & HUDSON, M. 2011. Employment Retention and Advancement (ERA) demonstration: Delivery, take-up and outcomes of in-work training support for lone parents. *Department for Work and Pensions Research Report,* 727.
HICKS, J. & ALLEN, R. 1934. A reconsideration of the theory of value. Part I. *Economica,* 1**,** 52-76.
HOEHN, J. & RANDALL, A. 1989. Too many proposals pass the benefit cost test. *The American Economic Review,* 79**,** 544-551.
HOWLEY, P. 2016. Valuing the benefits from health care interventions using life satisfaction data. *Health and Econometrics Data Group University of York,* 16/01.
HSEE, C. 2000. *Attribute evaluability and its implications for joint-separate evaluation reversals and beyond*. Cambridge University Press.
HSEE, C. 1998. Less is better: When low-value options are valued more highly than high-value options. *Journal of Behavioral Decision Making,* 11**,** 107-121.
HSEE, C. & HASTIE, R. 2006. Decision and experience: why don't we choose what makes us happy? *Trends in Cognitive Sciences,* 10**,** 31-37.

HUANG, L., FRIJTERS, P., DALZIEL, K. & CLARKE, P. 2018. Life satisfaction, QALYs, and the monetary value of health. *Social Science & Medicine,* 211**,** 131-136.
HUMPHREYS, M. 2009. Bounds on least squares estimates of causal effects in the presence of heterogeneous assignment probabilities. Columbia University. https://pdfs.semanticscholar.org/41ee/651ac1d5ecc9fa294d536acdccecc1fb4d05.pdf.
HUPPERT, F. 2006. *Positive emotions and cognition: developmental, neuroscience and health perspectives.* Philadelphia, Psychology Press.
HURKA, T. 1993. *Perfectionism.* Oxford, Clarendon Press.
HURWICZ, L. & UZAWA, H. 1971. *On the problem of integrability of demand functions.* New York, Harcourt Brace.
HUTCHINSON, G., CHILTON, S. & DAVIS, J. 1995. Measuring non-use value of environmental goods using the contingent valuation method: Problems of information and cognition and the design methods application of cognitive questionnaire design methods. *Journal of Agricultural Economics,* 46, 97-112.
IMBENS, G., RUBIN, D. & SACERDOTE, B. 2001. Estimating the effect of unearned income on labor earnings, savings, and consumption: Evidence from a survey of lottery players. *The American Economic Review,* 91**,** 778-794.
IRWIN, J., SLOVIC, P., LICHTENSTEIN, S. & MCCLELLAND, G. 1993. Preference reversals and the measurement of environmental values. *Journal of Risk & Uncertainty,* 6**,** 5-18.
IVERSEN, L., SABROE, S. & DAMSGAARD, M. 1989. Hospital admissions before and after shipyard closure. *British Medical Journal,* 299.
JAHODA, M. 1982. *Employment and unemployment: A social-psychological analysis*. Cambridge, Cambridge University Press.
JOHNS, H. & ORMEROD, P. 2007. *Happiness, economics and public policy.* London, Hobbs.
JONES-LEE, M., LOOMES, G. & PHILIPS, P. 1995. Valuing the prevention of non-fatal road injuries: Contingent valuation vs. standard gambles. *Oxford Economic Papers,* 47**,** 676- 695.
JOYCE, K., PABAYO, R., CRITCHLEY, J. & BAMBRA, C. 2010. Flexible working conditions and their effects on employee health and wellbeing. *Cochrane Database of Systematic Reviews* 2.
KAHNEMAN, D. 2000. *Choices, values, frames.* New York, Sage.
KAHNEMAN, D. 2012. *Thinking, fast and slow*. Penguin.
KAHNEMAN, D., KRUEGER, A., SCHKADE, D., SCHWARZ, N. & STONE, A. 2004. A survey method for characterizing daily life experience: The day reconstruction method. *Science,* 306**,** 1776-1780.
KAHNEMAN, D., FREDRICKSON, B., SCHREIBER, C. & REDELMEIER, D. 1993. When more pain is preferred to less: Adding a better end. *Psychological Science,* 4**,** 401-405.
KAHNEMAN, D. & KNETSCH, J. 1992. Valuing public goods: The purchase of moral satisfaction. *Journal of Environmental Economics and Management,* 22**,** 57-70.
KAHNEMAN, D., KRUEGER, A., SCHKADE, D., SCHWARZ, N. & STONE, A. 2006. Would you be happier if you were richer? A focusing illusion. *CEPS Working Paper,* 125.
KAHNEMAN, D. & KRUEGER, A. 2006. Developments in the measurement of subjective well-being. *The Journal of Economic Perspectives,* 20**,** 3-24.

KAHNEMAN, D., RITOV, I. & SCHKADE, D. 1999. Economic preferences or attitude expressions?: An analysis of dollar responses to public issues. *Journal of Risk & Uncertainty,* 19**,** 203-235.
KAHNEMAN, D., SCHKADE, D. & SUNSTEIN, C. 1998. Shared outrage and erratic awards: The psychology of punitive damages. *Journal of Risk & Uncertainty,* 16**,** 49-86.
KAHNEMAN, D. & SNELL, J. 1992. Predicting a changing taste: Do people know what they will like? *Journal of Behavioral Decision Making,* 5**,** 187-200.
KAHNEMAN, D. & SUGDEN, R. 2005. Experienced utility as a standard of policy evaluation. *Environmental and Resource Economics,* 32**,** 161-181.
KAHNEMAN, D. & THALER, R. 2006. Anomalies: Utility maximization and experienced utility. *The Journal of Economic Perspectives,* 20**,** 221-234.
KAHNEMAN, D. & TVERSKY, A. 1979. Prospect Theory: An analysis of decision under risk. *Econometrica,* 47**,** 263-291.
KAHNEMAN, D., WAKKER, P. & SARIN, R. 1997. Back to Bentham? Explorations of experienced utility. *The Quarterly Journal of Economics,* 112**,** 375-405.
KAMMANN, R. & FLETT, R. 1983. Affectometer: A scale to measure current level of general happiness. *Australian Journal of Psychology,* 35**,** 259-265.
KASSENBOEHMER, S. & HAISKEN-DENEW, J. 2009. You're fired! The causal negative effect of entry unemployment on life satisfaction. *Economic Journal,* 119**,** 448-462.
KAVETSOS, G., DIMITRIADOU, M. & DOLAN, P. 2014. Measuring happiness: context matters. *Applied Economics Letters,* 21**,** 308-311.
KEEFE, V., REID, P., ORMSBY, C., ROBSON, B., PURDIE, G. & BAXTER, J. 2002. Serious health events following involuntary job loss in New Zealand meat processing workers. *International Journal of Epidemiology,* 31**,** 1155-1161.
KELMAN, M. 2005. Hedonic psychology and the ambiguities of "welfare". *Philosophy & Public Affairs,* 33**,** 391-412.
KIECOLT-GLASER, J., MCGUIRE, L., ROBLES, T. & GLASER, R. 2002. Psychoneuroimmunology: Psychological influences on immune function and health. *Journal of Consulting and Clinical Psychology*. 70**,** 537–547.
KIMBALL, M. & WILLIS, R. 2006. Utility and happiness. *University of Michigan Paper*.
KLINE, R. 2005. *Principles and practice of structural equation modeling,* New York, Guildford Press.
KNIGHT, J., SONG, L. & GUNATILAKA, R. 2009. Subjective well-being and its determinants in rural China. *China Economic Review,* 20**,** 635-649.
KORPI, T. 1997. Is utility related to employment status? Employment, unemployment, labor market policies and subjective well-being among Swedish youth. *Labour Economics,* 4**,** 125-147.
KOSZEGI, B. & RABIN, M. 2008. Choices, situations, and happiness. *Journal of Public Economics,* 92**,** 1821-1832.
KOUNTOURIS, Y. & REMOUNDOU, K. 2011. Valuing the welfare cost of forest fires: a life satisfaction approach. *Kyklos,* 64**,** 556-578.
KREKEL, C. & ZERRAHN, A. 2017. Does the presence of wind turbines have negative externalities for people in their surroundings? Evidence from well-being data. *Journal of Environmental Economics and Management,* 82, 221-238.

KRUEGER, A. & SCHKADE, D. 2008. The reliability of subjective well-being measures. *Journal of Public Economics,* 92**,** 1833-1845.
KUEHNLE, D. & WUNDER, C. 2015. Using the life satisfaction approach to value daylight savings time transitions. Evidence from Britain and Germany. *SOEP paper,* 744.
LACHOWSKA, M. 2017. The effect of income on subjective well-being. Evidence from the 2008 economic stimulus tax rebates. *Journal of Human Resources,* 52**,** 374-417.
LAYARD, R. 2006. *Happiness. Lessons from a new science,* London, Penguin.
LAYARD, R., NICHOLL, S. & MAYARZ, G. 2008. The marginal utility of income. *Journal of Public Economics,* 92**,** 1846–1857.
LEVINSON, A. 2012. Valuing public goods using happiness data: The case of air quality. *Journal of Public Economics,* 96**,** 869-880.
LINDQVIST, E., ÖSTLING, R. & CESARINI, D. 2018. Long-run effects of lottery wealth on psychological well-being. *IFN Working Paper,* 1220.
LIST, J. 2003. Does market experience eliminate market anomalies? *The Quarterly Journal of Economics,* 118**,** 41-71.
LOEWENSTEIN, G. & ADLER, D. 1995. A bias in the prediction of tastes. *The Economic Journal,* 105**,** 929-937.
LOEWENSTEIN, G., O'DONOGHUE, T. & RABIN, M. 2003. Projection bias in predicting future utility. *The Quarterly Journal of Economics,* 118**,** 1209-1248.
LOEWENSTEIN, G. & SCHKADE, D. 1999. Wouldn't it be nice? Predicting future feelings. In Kahneman, D., Diener, E., & Schwarz, N. (Eds.), *Well-being: The foundations of hedonic psychology*, Russell Sage Foundation.
LOEWENSTEIN, G. & UBEL, P. 2008. Hedonic adaptation and the role of decision and experience utility in public policy. *Journal of Public Economics,* 92**,** 1795-1810.
LOOMES, G. 2006. (How) Can we value health, safety and the environment? *Journal of Economic Psychology,* 27**,** 713-736.
LUCAS, R., DIENER, E. & SUH, E. 1996. Discriminant validity of well-being measures. *Journal of Personality and Social Psychology,* 71**,** 616-628.
LUECHINGER, S. 2009. Valuing air quality using the life satisfaction approach. *Economic Journal,* 119**,** 482-515.
LUECHINGER, S. & RASCHKY, P. 2009. Valuing flood disasters using the life satisfaction approach. *Journal of Public Economics,* 93**,** 620-633.
LUTTMER, E. 2005. Neighbors as negatives: Relative earnings and well-being. *Quarterly Journal of Economics* 120**,** 963-1002.
LYKKEN, D. & TELLEGEN, A. 1996. Happiness is a stochastic phenomenon. *Psychological Science,* 7**,** 186-189.
LYUBOMIRSKY, S., KING, L. & DIENER, E. 2005. The benefits of frequent positive affect: Does happiness lead to success? *Psychological Bulletin,* 131**,** 803-855.
MACKERRON, G. 2011. Happiness economics from 35,000 feet. *Journal of Economic Surveys,* 26, 705-735.
MACKERRON, G. & MOURATO, S. 2009. Life satisfaction and air quality in London. *Ecological Economics,* 68**,** 1441-1453.
MAHASUWEERACHAI, P. & PANGJAI, S. 2018. Does piped water improve happiness? A case from Asian rural communities. *Journal of Happiness Studies,* 19, 1329-1346.

MALOUFF, J. & SCHUTTE, N. 1989. Shaping juror attitudes: Effects of requesting different damage amounts in personal injury trials. *The Journal of Social Psychology,* 129**,** 491 - 497.
MANNING, M., FLEMING, C. & AMBREY, C. 2016. Life satisfaction and individual willingness to pay for crime reduction. *Regional Studies,* 50, 2024-2039.
MARSH, K. & BERTRANOU, E. 2012. Can subjective well-being measures be used to value policy outcomes? The example of engagement in culture. *Cultural Trends,* 21**,** 299-310.
MATRIX. 2009. Lifelong learning and well-being: An analysis of the relationship between adult learning and subjective well-being. *IFLL Public Value Paper,* 3.
MCDONALD, R. & POWDTHAVEE, N. 2018. The shadow prices of voluntary caregiving: Using panel data of well-being to estimate the cost of informal care. *IZA Discussion Paper,* 11545.
MCFADDEN, D. & GREGORY, L. 1993. Issues in the contingent valuation of environmental goods: Methodologies for data collection and analysis. In Hausman, J. (ed.). *Contingent valuation: A critical assessment,* Amsterdam, North-Holland.
MCGILL, A. 2000. Counterfactual reasoning in causal judgments: Implications for marketing. *Psychology & Marketing,* 17, 323-343.
MCKENZIE, L. 1957. Demand theory without a utility index. *The Review of Economic Studies,* 24, 185–189.
MEDVEC, V., MADEY, S. & GILOVICH, T. 1995. When less is more: Counterfactual thinking and satisfaction among Olympic medalists. *Journal of Personality & Social Psychology,* 69, 603-610.
MENDOZA, Y., LOYOLA, R., AGUILAR, A. & ESCALANTE, R. 2019. Valuation of air quality in Chile: The life satisfaction approach. *Social Indicators Research,* 145**,** 367-387.
MENZ, T. & WELSCH, H. 2012. Life-cycle and cohort effects in the valuation of air quality: Evidence from subjective well-being data. *Land Economics,* 88**,** 300-325.
MENZEL, P., DOLAN, P., RICHARDSON, J. & OLSEN, J. 2003. The role of adaptation to disability and disease in health state valuation: A preliminary normative analysis. *Social Science and Medicine,* 55**,** 2149–2158.
METCALFE, R. 2009. Hedonic pricing: preferences and experiences. *London Imperial College*.
METCALFE, R., POWDTHAVEE, N. & DOLAN, P. 2011. Destruction and distress: Using a quasi-experiment to show the effects of the September 11 attacks on mental well-being in the United Kingdom. *The Economic Journal,* 121, 81-103.
MOORE, S. 2006. The value of reducing fear: an analysis using the European Social Survey. *Applied Economics,* 38**,** 115-117.
MORRIS, J. & COOK, D. 1991. A critical review of the effect of factory closures on health. *British Journal of Industrial Medicine,* 48, 1-8.
MOSTELLER, F. & NOGEE, P. 1951. An experimental measurement of utility. *Journal of Political Economy,* 59, 371-404.
MOURATO, S., ATKINSON, G., OZDEMIROGLU, E., NEWCOMBE, J. & DE GARIS, Y. 2005. Does a cleaner Thames pass an economic appraisal?: The value of reducing sewage overflows in the River Thames. *Water International,* 30**,** 174-183.

MURPHY, G. & ATHANASOU, J. 1999. The effect of unemployment on mental health. *Journal of Occupational and Organizational Psycholoy,* 72**,** 83-99.
MURPHY, J., ALLEN, P., STEVENS, T. & WEATHERHEAD, D. 2005. A meta-analysis of hypothetical bias in stated preference valuation. *Environmental and Resource Economics,* 30**,** 313-325.
MURTIN, F., BOARINI, R., CORDOBA, J. & RIPOLL, M. 2017. Beyond GDP: Is there a law of one shadow price? *European Economic Review,* 100**,** 390-411.
NG, Y-K. 1997. A case for happiness, cardinalism, and interpersonal comparability. *The Economic Journal,* 107**,** 1848-1858.
NG, Y-K. 2003. From preference to happiness: Towards a more complete welfare economics. *Social Choice and Welfare,* 20**,** 307-350.
NISBETT, R. & KANOUSE, D. 1969. Obesity, food deprivation, and supermarket shopping behavior. *Journal of Personality and Social Psychology,* 12**,** 289-294.
NORTHCRAFT, G. & NEALE, M. 1987. Experts, amateurs, and real estate: An anchoring and adjustment perspective on property pricing decisions. *Organizational Behavior and Human Decision Processes,* 39**,** 84-97.
NOZICK, R. 1974. *Excerpt from anarchy, state, and utopia,* US, Basic Books.
NUSSBAUM, M. 2000. The costs of tragedy: Some moral limits of cost-benefit analysis. *The Journal of Legal Studies,* 29**,** 1005-1036.
OECD. 2006. *Cost-benefit analysis and the environment: Recent developments. OECD.*
OECD. 2018. *Cost-benefit analysis and the environment*. OECD.
OSWALD, A. & POWDTHAVEE, N. 2008. Death, happiness, and the calculation of compensatory damages. *Journal of Legal Studies,* 37**,** 217-251.
PARFIT, D. 1984. *Reasons and persons,* Oxford Scholarship Online.
PAVOT, W. & DIENER, E. 1993. The affective and cognitive context of self-reported measures of subjective well-being. *Social Indicators Research,* 28**,** 1-20.
PAVOT, W., DIENER, E., COLVIN, C. & SANDVIK, E. 1991. Further validation of the satisfaction with life scale: Evidence for the cross-method convergence of well-being measures. *Journal of Personality Assessment,* 57**,** 149-161.
PISCHKE, J. 2010. Money and happiness: Evidence from the industry wage structure. *CEP Discussion Papers,* 1051.
PLASSMANN, H., O' DOHERTY, J. & RANGEL, A. 2007. Orbitofrontal cortex encodes willingness to pay in everyday transactions. *Journal of Neuroscience,* 27, 9984-9988;
PLOTT, C. 1996. *Rational individual behavior in markets and social choice processes: The Discovered Preference Hypothesis,* London, Macmillan.
POUNDSTONE, W. 2010. *Priceless: The hidden psychology of value*. UK, Oneworld Book.
POWDTHAVEE, N. 2008. Putting a price tag on friends, relatives, and neighbours: Using surveys of life satisfaction to value social relationships. *Journal of Socio-Economics,* 37**,** 1459-1480.
POWDTHAVEE, N. 2010. How much does money really matter? Estimating the causal effects of income on happiness. *Empirical Economics,* 39**,** 77-92.
PROVENCHER, B., LEWIS, D. & ANDERSON, K. 2012. Disentangling preferences and expectations in stated preference analysis with respondent uncertainty: The case of invasive species prevention. *Journal of Environmental Economics and Management,* 64, 169-182.

RANDALL, A. 1982. Economic surplus concepts and their use in benefit-cost analysis. *Review of Marketing and Agricultural Economics,* 50, 135-164.
READ, D. 2007. Utility theory from Jeremy Bentham to Daniel Kahneman. *Thinking and Reasoning,* 13**,** 45-61.
READ, D. & VAN LEEUWEN, B. 1998. Predicting hunger: The effects of appetite and delay on choice. *Organizational Behavior and Human Decision Processes,* 76**,** 189-205.
REHDANZ, K. & MADDISON, D. 2005. Climate and happiness. *Ecological Economics,* 52**,** 111-125.
REHDANZ, K., WELSCH, H., NARITA, D. & OKUBO, T. 2015. Well-being effects of a major natural disaster: The case of Fukushima. *Journal of Economic Behavior & Organization,* 116**,** 500-517.
ROBINSON, L. & HAMMITT, J. 2011. Behavioral economics and the conduct of benefit-cost analysis: Towards principles and standards. *Journal of Benefit-Cost Analysis,* 2.
SAGOFF, M. 2000. Environmental economics and the conflation of value and benefit. *Environmental Science & Technology,* 34, 1426-1432.
SAGOFF, M. 2003. On the relation between preference and choice. *Journal of Socio-Economics,* 31**,** 587-598.
SALES, S. & HOUSE, J. 1971. Job dissatisfaction as a possible risk factor in coronary heart disease. *Journal of Chronic Diseases,* 23, 861-873.
SAMPLES, K. & HOLLYER, J. 1990. Contingent valuation of wildlife resources in the presence of substitutes and complements, In Johnson, R. & Johnson, G. (Eds.), *Economic valuation of natural resources: issues, theory, and applications*. Boulder, Westview Press.
SANDEL, M. 2013. *What money can't buy: The moral limits of markets*, Penguin.
SANDVIK, E., DIENER, E. & SEIDLITZ, L. 1993. Subjective well-being: The convergence and stability of self-report and non-self-report measures. *Journal of Personality,* 61**,** 317-342.
SARRIAS, M. & JARA, B. 2019. How much should we pay for mental health deterioration? The subjective monetary value of mental health after the 27F Chilean earthquake. *Journal of Happiness Studies*.
SCHIMMACK, U., BÖCKENHOLT, U. & REISENZEIN, R. 2002a. Response styles in affect ratings: Making a mountain out of a molehill. *Journal of Personality Assessment,* 78**,** 461-483.
SCHIMMACK, U., DIENER, E. & OISHI, S. 2002b. Life satisfaction is a momentary judgment and a stable personality characteristic: The use of chronically accessible and stable sources. *Journal of Personality,* 70**,** 345-384.
SCHKADE, D. & KAHNEMAN, D. 1998. Does living in California make people happy? A focusing illusion in judgments of life satisfaction. *Psychological Science,* 9**,** 340-346.
SCHNEIDER, U. & KLEINDIENST, J. 2016. Monetising the provision of informal long-term care by elderly people: estimates for European out-of-home caregivers based on the well-being valuation method. *Health & Social Care in the Community,* 24, 81-91.
SCHULZE, W., BOYCE, R. & MCCLELLAND, G. 2018. Contingent valuation of natural resource damages due to injuries to the upper Clark Fork River Basin. Sagwan Press.

SCHWARZ, N. 2012. Why researchers should think "real-time": A cognitive rationale. In Mehl, M. & Conner, T. (Eds.), *Handbook of research methods for studying daily life*, The Guilford Press.
SCHWARZ, N. & CLORE, G. 2003. Mood as information: 20 years later. *Psychological Inquiry,* 14**,** 296-303.
SCHWARZ, N. & STRACK, F. 1999. Reports of subjective well-being: Judgmental processes and their methodological implications. In Kahneman, D., Diener, E., & Schwarz, N. (eds.). *Well-being: The foundations of hedonic psychology*, New York, Russell Sage.
SCHWARZ, N., STRACK, F. & MAI, H-P. 1991. Assimilation and contrast effects in part-whole question sequences: A conversational logic analysis. *The Public Opinion Quarterly,* 55**,** 3-23.
SCOLLON, C., KIM-PRIETO, C. & DIENER, E. 2003. Experience sampling: Promises and pitfalls, strengths and weaknesses. *Journal of Happiness Studies,* 4**,** 5-34.
SEN, A. 1977. Rational Fools. *Philosophy and Public Affairs,* 6**,** 317-344.
SEN, A. 1999. *Development as Freedom,* Oxford, Oxford University Press.
SHIZGAL, P. 1999. *On the neural computation of utility: Implications from studies of brain stimulation reward.,* York, Russell Sage.
SHORTT, S. 1996. Is unemployment pathogenic? A review of current concepts with lessons for policy planners. *International Journal of Health Services,* 26**,** 569-589.
SIMONSON, I. 1990. The effect of purchase quantity and timing on variety-seeking behavior. *Journal of Marketing Research,* 27**,** 150-162.
SLOVIC, P., FINUCANE, M., PETERS, E. & MACGREGOR, D. 2002. Rational actors or rational fools: implications of the affect heuristic for behavioral economics. *Journal of Socio-Economics,* 31**,** 329-342.
SLOVIC, P. & LICHTENSTEIN, S. 2006. *The construction of preference,* New York, Cambridge University Press.
SMITH, D., SHERRIFF, R., DAMSCHRODER, L., LOEWENSTEIN, G. & UBEL, P. 2006. Misremembering colostomies? Former patients give lower utility ratings than do current patients. *Health Psychology,* 25**,** 688-695.
SMITH, V. & MOORE, E. 2010. Behavioral economics and benefit-cost analysis. *Environmental and Resource Economics,* 46**,** 217-234.
SMITH, V. 1993. Nonmarket valuation of environmental resources: An interpretive appraisal. *Land Economics,* 69**,** 1-26.
SMITH, T. 1979. Happiness. *Social Psychology Quarterly*, 42, 18-30.
STEVENSON, B. & WOLFERS, J. 2008. Economic growth and subjective well-being: Reassessing the Easterlin Paradox. *NBER Working Paper,* 14282.
STRACK, F., MARTIN, L. & SCHWARZ, N. 1988. Priming and communication: Social determinants of information use in judgments of life satisfaction. *European Journal of Social Psychology,* 18**,** 429-442.
STRACK, F., SCHWARZ, N. & GSCHNEIDINGER, E. 1985. Happiness and reminiscing: The role of time perspective, affect, and mode of thinking. *Journal of Personality and Social Psychology*, 49, 1460–1469.
STUDNICKA, M., STUDNICKA-BENKE, A., WOGERBAUER, G., RASTETTER, D., WENDA, R., GATHMANN, P. & RINGEL, E. 1991. Psychological health, self-reported physical health and health service use. *Social Psychiatry and Psychiatric Epidemiology*, 26, 86–91.

STUTZER, A. & FREY, B. 2004a. Reported subjective well-being: A challenge for economic theory and economic policy. *Schmollers Jahrbuch,* 124**,** 191-231.
STUTZER, A. & FREY, B. 2004b. Stress that doesn't pay: The commuting paradox. *Institute for Empirical Research in Economics, University of Zurich Working Paper Series.*
STUTZER, A. & FREY, B. 2010. Recent advances in the economics of individual subjective well-being. *IZA Discussion Paper,* 4850.
SUGDEN, R. 2005. Anomalies and stated preference techniques: A framework for a discussion of coping strategies. *Environmental and Resource Economics,* 32**,** 1-12.
SUGDEN, R. 2015. Looking for a psychology for the inner rational agent. *Centre for Behavioural and Experimental Social Science and School of Economics. University of East Anglia.*
SUGDEN, R. 2018. *The community of advantage. A behavioural economist's defence of the market,* Oxford, Oxford University Press.
SUMNER, L. 1996. *Welfare, happiness and ethics,* Oxford, Clarendon Press.
SUNSTEIN, C. 2015. The ethics of nudging. *Yale Journal on Regulation,* 32.
SUNSTEIN, C. 2018. *The cost-benefit revolution,* MIT Press.
TELEGRAPH. 2014. *20 amazing facts from 20 years of the National Lottery*. www.telegraph.co.uk/finance/personalfinance/11228989/National-Lottery-20-amazing-facts-from-20-years-of-creating-millionaires.html
TELLA, R., MACCULLOCH, R. & OSWALD, A. 2003. The macroeconomics of happiness. *The Review of Economics and Statistics,* 85**,** 809-827.
TER DOEST, L. & DE JONGE, J. 2006. Testing causal models of job characteristics and employee well-being: A replication study using cross-lagged structural equation modelling. *Journal of Occupational and Organizational Psychology,* 79**,** 499-507.
TREASURY. 2018. Green Book.
TSURUMI, T. & MANAGI, S. 2015. Environmental value of green spaces in Japan: An application of the life satisfaction approach. *Ecological Economics,* 120**,** 1-12.
TVERSKY, A. & GRIFFIN, D. 1991. Endowment and contrast in judgments of well-being. In Strack, F., Argyle, M. & Schwarz, N. (Eds.), *International series in experimental social psychology, Vol. 21. Subjective well-being: An interdisciplinary perspective,* New York, Pergamon Press.
URRY, H., NITSCHKE, J., DOLSKI, I., JACKSON, D., DALTON, K., MUELLER, C., ROSENKRANZ, M., RYFF, C., BURTON, H. & DAVIDSON, R. 2004. Making a life worth living: Neural correlates of well-being. *Psychological Science,* 15**,** 367-372.
VAN DEN BERG, B. & FERRER-I-CARBONELL, A. 2007. Monetary valuation of informal care: the well-being valuation method. *Health Economics,* 16**,** 1227-1244.
VAN DEN BERG, J., VERBERG, C., BERKHOUT, J., LOMBARTS, M., SCHERPBIER, A. & JAARSMA, A. 2015. A qualitative interview study on the positive well-being of medical school faculty in their teaching role: job demands, job resources and role interaction. *BMC research notes,* 8.
VAN PRAAG, B. & BAARSMA, B. 2005. Using happiness surveys to value intangibles: The case of airport noise. *Economic Journal,* 115**,** 224-246.

VAN PRAAG, B., FRIJTERS, P. & FERRER-I-CARBONELL, A. 2003. The anatomy of subjective well-being. *Journal of Economic Behavior & Organization,* 51**,** 29-49.
VEENHOVEN, R. 2004. *Happiness as an aim in public policy: the greatest happiness principle,* Hoboken (NJ), John Wiley & Sons.
VEENHOVEN, R. 2010. Capability and happiness: Conceptual difference and reality links. *Journal of Socio-Economics,* 39**,** 344-350.
WADDELL, G. & BURTON, A. 2006. *Is work good for your health and well-being?* London, TSO (The Stationery Office).
WARR, P. 1987. *Work, unemployment and mental health,* Oxford, Oxford University Press.
WARR, P. 1994. A conceptual framework for the study of work and mental health. *Work and Stress* 8**,** 84-97.
WELSCH, H. 2002. Preferences over prosperity and pollution: Environmental valuation based on happiness surveys. *Kyklos,* 55**,** 473-494.
WELSCH, H. 2006. Environment and happiness: Valuation of air pollution using life satisfaction data. *Ecological Economics,* 58**,** 801-813.
WELSCH, H. 2007. Environmental welfare analysis: A life satisfaction approach. *Ecological Economics,* 62**,** 544-551.
WELSCH, H. 2008a. The social costs of civil conflict: Evidence from surveys of happiness. *Kyklos,* 61**,** 320-340.
WELSCH, H. 2008b. The welfare costs of corruption. *Applied Economics,* 40**,** 1839-1849.
WELSCH, H. & FERREIRA, S. 2014. Environment, well-Being, and experienced preference. *International Review of Environmental and Resource Economics,* 7, 205-239.
WELSCH, H. & KÜHLING, J. 2009. Using happiness data for environmental valuation: issues and applications. *Journal of Economic Surveys,* 23**,** 385-406.
WHITTINGTON, D., SMITH, V., OKORAFOR, A., OKORE, A., LIU, J. & MCPHAIL, A. 1992. Giving respondents time to think in contingent valuation studies: A developing country application. *Journal of Environmental Economics and Management,* 22**,** 205-225.
WILSON, T. & GILBERT, D. 2003. Affective forecasting. *Advances in Experimental Social Psychology,* 35**,** 345-411.
WINKELMANN, L. & WINKELMANN, R. 1998. Why are the unemployed so unhappy? Evidence from panel data. *Economica,* 65**,** 1-15.
WIRTZ, D., KRUGER, J., SCOLLON, C. & DIENER, E. 2003. What to do on spring break? The role of predicted, on-line, and remembered experience in future choice. *Psychological Science,* 14**,** 520-524.
WONG, S. 2006. *The foundations of Paul Samuelson's revealed preference theory: A study by the method of rational reconstruction.* Oxford, Routledge.
WRIGHT, L., PEASGOOD, T. & MACLENNAN, S. 2017. *A guide to wellbeing economic evaluation*. What Works Centre for Wellbeing.

Valuing Non-Market Goods using Subjective Wellbeing Data

Addendum to PhD Thesis

February 2020

Daniel Fujiwara

This Addendum contains the corrections to the PhD Thesis requested by the examiners and is to be read alongside the full Thesis.

Corrections and additions to Chapter 1

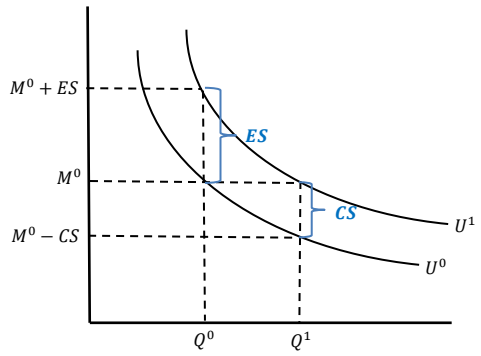
None.

Corrections and additions to Chapter 2

1. The thesis uses a range of terminology related to the concept of wellbeing and so I provide further clarification here on how I define and use the terminology throughout the thesis:
    - **Welfare** is an overall term for quality of life or how one's life is going. I use the term Welfare and its definition interchangeably with the terms **Wellbeing** and **Quality of life** in this thesis; all terms are used to refer to how one's life is going overall.
    - Within the concept of **Welfare/Wellbeing/Quality of life** there are various specific metrics and I discuss three such metrics in the thesis. These represent three different ways in which we can measure **Welfare/Wellbeing/Quality of life**:
        i. **Subjective wellbeing** (also termed as **Mental states** in this thesis), which refers to people's self-reports about how their lives are going;
        ii. **Preferences** which assume that wellbeing consists of the extent to which one's preferences are satisfied - "what would be best for someone is what would best fulfil his desires" (Parfit, 1984. P.4).
        iii. **Objective list** accounts of wellbeing, which are based on normative assumptions about basic human needs and rights.
    - **Preference** is related to the economist's term **utility**: utility is what is generated through the satisfaction of preference – the more preferences you satisfy, the higher your level of utility. Therefore, under this definition, **utility is equivalent to welfare/wellbeing/quality of life**.
    - **Utility maximisation** is the act of satisfying as many of one's preferences as possible given a budget and resource constraint. In the economist's terminology it is therefore the maximisation of **welfare/wellbeing/quality of life**.

- The term **Preference realisation** refers to the satisfaction of preferences (e.g. consuming things that you want), which in turn leads to higher utility and wellbeing.

2. The figure below illustrates Compensating surplus (CS) and Equivalent surplus (ES) for a non-market good (Q). The change Q^0 to Q^1 represents a positive impact on utility from the non-market good which moves the individual from indifference curve U^0 to indifference curve U^1. M^0 is the initial level of income. Here CS is equivalent to the WTP for the good and ES equals the WTA compensation to forgo the non-market good.



Corrections and additions to Chapter 3

1. I add to the list of first papers published on wellbeing valuation (WV) in 2002 to also include Clark and Oswald (2002) and Welsch (2002). Alongside Ferrer-i-Carbonell and Van Praag (2002) these were the first papers to be published on the WV method.

   It should also be noted that various working papers used the WV method in the 1990s and that the method was first presented in 1993 at the Economics of Happiness conference at the London School of Economics and Political Science.

2. I provide a discussion on how WV compares and relates to the hedonic pricing method. Broadly speaking as hedonic pricing is limited in terms of the types of non-market goods it can value (only non-market goods that impact on house

prices) WV has a wider scope. However, there are some interesting areas of overlap. A key issue for hedonic pricing is the extent to which the house market is functioning - if there are market imperfections, house prices will not reveal the true value of non-market goods like air quality and safety (Welsch and Ferreira, 2013). Problems include government regulation in housing markets, lack of information (e.g. about air quality) when people purchase their homes and transaction costs. In these cases, WV can be used to estimate the surplus value of the non-market good, i.e. the value over and above any impact on house prices (compensation received through housing prices) of the non-market good (Ferreira and Moro, 2009; Luechinger, 2009; Levinson, 2012; Van Praag and Baarsma, 2005; Cohen, 2008) because WV can capture all impacts of things like air quality even if people are not consciously aware of them (Welsch and Ferreira, 2013). This is achieved by *not* controlling for house prices in the wellbeing regression, allowing them to vary and for the effect of the house price change to be internalised such that the coefficient on the non-market good represents the surplus effect on subjective wellbeing (SWB). In this sense WV can complement the hedonic pricing method when there are housing market imperfections and the degree of complementarity is a positive function of the level of market imperfection.

3. Chapter 3 discusses at length the issues related to measuring SWB and provided evidence to support the use of the life satisfaction measure. An interesting recent study to add here is Bond and Lang's (2019) critique of happiness measures. They state that if happiness is not reported/scaled in the exact same way across individuals it is impossible to rank groups based on their mean levels of happiness, unless the distribution of happiness in one group stochastically dominates the distribution of the other. This relates to the broader issue of interpersonal comparability of wellbeing measures, which was discussed at length in the thesis and I presented arguments in favour of interpersonal comparability (e.g. Kahneman, 2000; Diener et al., 1999; Duncan, 2010). In particular van Praag et al. (2003) and van Praag (1991) find that there is a common reporting function for wellbeing measures across society.

This discussion is a relative issue since interpersonal comparability is something that is important for all valuation methods. Whilst there are issues related to interpersonal comparability for wellbeing measures, the same can also be said for preferences and indeed as argued in section 4.3.2, interpersonal incomparability issues may even be more serious in preference-based valuation methods such as stated preference. Preference-based valuation methods like contingent valuation do not use or recommend methods for addressing interpersonal comparability such as Harsanyi's extended preference method (see Hausman and McPherson, 2006) and hence are just as susceptible to interpersonal comparability issues as WV. These types of critique are therefore important but they are not something that should dismiss the WV method per se, rather it is an important area for future research in WV. A particularly useful question to address for the purposes of non-market valuation would be which form of welfare measure - SWB or preferences – is *more* interpersonally comparable?

4. Section 3.3.4.2.3 (SWB measurement issues) discusses the problems of interview mode effects and priming and question order effects for the measurement of life satisfaction. I noted there that recent studies have been unable to replicate the results from Schwarz and Strack's seminal work on these types of contextual biases, hinting that these issues may not be such a concern as first thought (Haybron, 2010; Diener and Suh, 1997). Therefore, it is my assessment that WV is unlikely to be prone to these issues, however, we should be aware of the consequences when/if these types of biases arise in life satisfaction surveys.

   If the interviewer mode and question ordering are constant for all respondents in the survey then this should not be a problem for WV since we would expect the effect of question ordering to even out over large surveys as previous questions will invoke positive memories for some people and negative ones for others. And the constant interviewer mode will have the same effect on all people either increasing or decreasing their life satisfaction scores but this is not a problem since it is the ratio of impacts (coefficients) that we are interested in rather than the absolute values and the ratio of coefficients will

remain constant under a monotonic change in life satisfaction ratings. However, where interviewer mode and priming biases exist this will potentially cause problems if interviewer mode and question ordering were not constant across the sample. This would potentially invalidate WV studies that use, for example, samples that come from different interviewer modes unless the mode was randomly assigned or unless we control for the variable with which interviewer mode is correlated (e.g. elderly populations may be more likely to do surveys over the phone or face-to-face rather than online). We can also control for question ordering in the statistical analysis if priming is an issue in the survey. In summary, there seems to be a low risk of WV studies being affected by interviewer mode and priming effects, but where they are there are ways to nullify the impact and so these biases should not pose a serious threat to WV.

Additional References (for papers not already included in thesis bibliography)

BOND, T. & LANG, K. 2019. The sad truth about happiness scales. *Journal of Political Economy*, 127, 1629-1640.

WELSCH, H. & FERREIRA, S. 2013. Environment, well-being and experienced preference. *International Review of Environmental and Resource Economics*, 7, 205-238.

Corrections and additions to Chapter 4

1. Equation (13) (p.110) requires further clarification. It does not make sense to multiply a discrete change by a derivative (the numerator term in equation (13)) *unless* the derivative is from a linear functional form (where the change in the dependent variable is constant for all levels of change in the independent variable), or is from a model where the variable Q is binary such that two discrete changes are in effect being multiplied together. As explained on p.115 the focus of the thesis is on binary Q variables (i.e. the presence or not of the non-market good), and in equation (10), which is a precursor to equation (13), I explain that the equation is intended as a basic format and that

it becomes more complex when the Q and M variables take on a non-linear format.

To be clear, therefore, equation (13) is not applicable when the non-market good variable Q is a non-binary (for example continuous) variable that has a non-linear relationship with SWB. For example, this could be levels of air pollution which have higher marginal impacts on SWB the higher the initial levels of pollution. The non-market good that is used in the case study in the thesis (unemployment) is binary and hence equation (13) can be applied to it. I do not provide an exposition of equation (13) here for the case of non-binary non-linear Q variables because (a) there are numerous ways in which the non-linear relationship between Q and SWB can be estimated and so it is not possible to set out a generic functional form for this here, and (b) the focus of this thesis is on binary Q variables.

2. The income coefficient in wellbeing regressions that use an endogenous income variable tends to be small due to a number of reasons. First, findings from the literature that instruments for income (or uses exogenous changes in income) would suggest that people who are less satisfied with life to begin with tend to earn more, potentially with the aim of improving their lives (reverse causality) or that there are factors that we cannot fully control for, such as personality traits, that make some people less satisfied with their lives and more likely to earn higher incomes (endogeneity). This would result in a downward bias in the income coefficient. Second, income is hard to measure in surveys because of all of the different sources of income that people have, often leading to measurement error in the income variable. Measurement error in independent variables in regression analysis leads to downward bias in the coefficient. Third, people adapt to higher incomes causing a diminishing marginal utility of income and since the wellbeing regressions I review and estimate in the thesis have generally come from high income countries such as the UK, absolute changes in income will generally have smaller effects on life satisfaction and it may be that relative changes in income take greater importance. A final problem in wellbeing regressions is that income impacts

on wellbeing through various channels, for example through health, education and housing quality, but often these factors are controlled for in the same regression model which nullifies these indirect effects, leading to a reduction in the size of the coefficient for income.

3. To provide context for the income analysis and results in the thesis, the following table provides a review of income coefficient results for a range of studies that focus on life satisfaction. I focus only on studies that are directly comparable to the methods used here – models that use multivariate regression analysis and where income is defined as log of household income and life satisfaction is used as the measure of wellbeing. This narrows down the range of possible comparison studies as a wide range of methods and variable definitions are used. Studies come from the UK and other developed countries and I focus only on studies that use large secondary national data sets and exclude small primary data studies. The response scale to the life satisfaction question varies across surveys. To make the results comparable to the BHPS I have rescaled the coefficients to a seven-point scale by dividing by the number of points on the original scale and multiplying by 7. These figures can be compared to the results presented in Chapter 6.

Author(s)	Country	Data set	Income coefficient	Life satisfaction scale	7 point scale conversion
Frey et al. (2009)	UK	Euro-Barometer	0.168	4	0.294
Powdthavee (2010)	UK	BHPS	0.105	7	0.105
Garcia-Mainar et al. (2015)	Spain	Quality of Working Life Survey	0.140	11	0.089
Schneider and Kleindienst (2015)	Europe	SHARE	0.171	11	0.109
Murtin et al. (2017)	World	Gallup	0.480	11	0.305
Knoll & Pitlik (2016)	Europe	ESS	0.147	11	0.094
Brenig & Proeger (2016)	Europe	ESS	0.486	11	0.309
Maccagnan et al. (2019)	England & Wales	Crime Survey for England and Wales	0.217	11	0.138

Barrington-Leigh & Behzadnejad (2017)	Canada	Canadian Community Health Survey	0.173	6	0.202
Ólafsdóttir et al. (2017)	USA	Health and Retirement Study	0.006	5	0.008
Ambrey et al. (2014)	Australia	HILDA	0.184	11	0.117
Kuehnle & Wunder (2014)	UK	BHPS	0.183	7	0.183
Danzer & Danzer (2011)	Ukraine	ULMS	0.169	5	0.237
				AVERAGE	**0.168**

The mean coefficient size for log of household income in OLS models is 0.168 across these studies. This is comparable to the income coefficient estimated using the BHPS in this thesis of 0.104 (see Chapter 6). Incidentally, this is almost identical to Powdthavee's (2010) estimate using the BHPS of 0.105.

The size of the income coefficient in the literature varies depending on whether OLS, fixed effects (FE) or random effects (RE) models are used and they tend to be lower in FE and RE models. A small number of the above studies also ran FE and RE models with the mean coefficient size for log of household income estimated at 0.059 (based on 4 studies).

4. I clarify here that hedonic pricing methods for valuation do not solely pick up use values. For example, house prices will reflect so-called option value for things like nearby parks, beaches, museums and other amenities. This is the value that people place on the option of using/visiting a site in the future although they don't use it now.

Additional References (for papers not already included in thesis bibliography)

AMBREY, C., FLEMING, C. & CHAN, A. 2014. Estimating the cost of air pollution in South East Queensland: An application of the life satisfaction non-market valuation approach. *Ecological Economics*, 97, 172-181.

BRENIG, M. & PROEGER, T. 2016. Putting a price tag on security: Subjective well-being and willingness-to-pay for crime reduction in Europe. *CEGE Discussion Papers*, 278.

GARCIA-MAINAR, I., MONTUENGA, V. & NAVARRO-PANIAGUA, M. 2015. Workplace environmental conditions and life satisfaction in Spain. *Ecological Economics*, 119, 136-146.

KNOLL, B. & PITLIK, H. 2014. Who benefits from big government? A life satisfaction approach. *WWW For Europe Policy Paper*, 14.

ÓLAFSDOTTIR, T., ÁSGEIRSDOTTIER, T. & NORTON, E. 2017. Valuing pain using the subjective wellbeing method. *NBER Working Paper Series*, 23649.

MACCAGNAN, A., TAYLOR, T. & WHITE, M. 2019. Valuing the relationship between drug and alcohol use and life satisfaction: Findings from the Crime Survey for England and Wales. *Journal of Happiness Studies*.

MURTIN, F., BOARINI, R., CORDOBA, J. & RIPOLL, M. 2017. Beyond GDP: Is there a law of one shadow price? *European Economic Review*, 100, 390-411.

Corrections and additions to Chapter 5

1. The equations (19) and (20) that are set out in section 5.2 in reference to the 3S-WV approach suppressed the other determinants of SWB, which was not clearly stated in the text. Therefore, I rectify the equations to read as follows (so that they are in line with previous equation (8) in the text):

   (19) $SWB_i = f(\ln(M_i), X_i^M)$

   Where M_i is the income of individual i and X_i^M is a vector of other determinants of SWB that are correlated with M_i. Note that X_i^M may contain the non-market good variable Q_i.

   (20) $SWB_i = f(Q_i, X_i^Q)$

   Where Q_i is the non-market good as experienced by individual i and X_i^Q is a vector of other determinants of SWB that are correlated with Q_i. Note that X_i^Q may contain the income variable M_i.

2. I find that using the lottery wins IV in the control function approach leads to a 10-fold increase in the size of the coefficient on income (p.160), which is in

line with previous IV studies in the literature which find a 10-12 fold increase. I provide further clarification here as to why, therefore, my estimation procedure and results contribute to the literature. There are two reasons why the method used in the thesis is novel and adds to the literature:

i. As explained in the thesis the rationale for some of the previous instruments used for income is not clear and there are reasons to argue that the core identifying assumptions may not hold in many cases. I have argued that lottery wins, although not perfect, represents the most robust instrument for income in current data sets. Therefore, I would argue that although it is useful to compare my lottery wins results with the wider IV literature we should note that some of the results in the other IV literature may have come about through chance where the instrument used is not robust. Using a robust instrument in lottery wins has demonstrated that we can be confident that the increase in the income coefficient is around 10-fold when using an IV, so at the lower end of findings in the wider literature.
ii. Secondly in this thesis I use the control function method rather than the more traditional two-stage least squares (2SLS) approach, which means that the estimated impact of income on SWB is for the sample average rather than for the narrower complier subset as per 2SLS. It is therefore not possible to directly compare my results with the wider IV literature and so any comparison should be made with these caveats. The results of the income instrument using the control function here are more generalisable than 2SLS models which is a contribution to the literature.

Overall for these reasons the findings from the wider literature are used to provide context rather than a test of the results in the thesis.

3. As I set out on p.151 unless we know how often people play the lottery, lottery wins do not represent exogenous changes in income with respect to SWB. The BHPS data set does not contain data on how often people play lotteries and so my solution is to adjust for past lottery win amounts, which is claimed will

more accurately reflect current lottery playing frequency than observable socioeconomic factors/variables which have been used as control variables in the lottery IV literature to date. I reiterate that this is an assumption and that previous lottery win amounts is not a perfect indicator of current lottery playing frequency and hence it is a partial fix for this issue.

4. I provide here more details in relation to the euphoria effect of lottery wins and why it is not an issue for the lottery wins instrument I use in the thesis. As I set out on p.156 the euphoria experienced from winning in a lottery has the potential to invalidate the exclusion restriction in the IV set up since lottery wins would impact directly on SWB through a channel other than through income. I hypothesise that this is not a problem in my analysis as following Gardner and Oswald (2002) and Imbens et al. (2001) I compare lottery winners of different amounts and exclude entirely people who do not win or play lotteries. The IV (Z) = 0 for people with £1-£99 of lottery wins and Z = 1 for people with lottery wins of £100 and over (restricted to a maximum of £50,000).

   Here both groups are winners and will feel some happiness due to having won. Larger winners (the $Z = 1$ group) may feel *more* euphoria than smaller winners (the $Z = 0$ group), but this would be because the level of euphoria experienced at winning the lottery is correlated with win size, which suggests that it is the money prize that causes happiness - precisely the effect that we are interested in for the instrument. A second important point to note is that the euphoria felt from the act of winning itself may only be temporary anyway and not picked up in the life satisfaction responses at the time of the survey. Hence, I argue that the exclusion restriction assumption is satisfied for lottery wins under this set up.

   In the table below I present sensitivity analysis of the control function model with the lottery winnings instrument (from p.160) with different cut-off thresholds for definitions of small versus medium/large-sized lottery wins. I present the original results with the £100 threshold against models where the threshold is doubled to £200; then £400 and then a final large threshold of

£1,000. Of the sample of lottery winners in the BHPS only 3.1% of people have higher annual lottery wins than £1,000, which will make the regression results less stable and reliable. The original threshold of £100 was chosen as it allowed me to retain a large sample of winners in the Z=1 instrument category (over 20% of the lottery winning sample had annual winnings higher than £100).

Table 1 sets out the key results from the second stage of the control function, which shows the impact of income on the dependent variable life satisfaction. The key result used in WV in the thesis is the coefficient on log (household income) of 1.103 from the £100 threshold model. Moving the lottery winnings threshold to £200 or £400 does not have a material impact on the results with the income coefficient staying at around 1.0 to 1.1, which is well within the 90% confidence interval of the result from the original £100 threshold model (C.I. = 0.609 – 1.598). There is, however, an increase in the income coefficient in the £1,000 threshold model to 1.563. This model, however, should be ignored as the sample size with Z=1 for the instrument category is very low which makes the results less robust as can be seen by the substantially lower level of statistical significance; the income coefficient is now only significant at the 10% level in this model. The results here suggest that the coefficient on income from the control function using lottery wins as an instrument are robust to changes in the lottery winnings threshold for the lottery instrument.

Table 1. Sensitivity analysis of the control function results

	£100 threshold	**£200 threshold**	**£400 threshold**	**£1,000 threshold**
Independent variables	**Coefficient (S.E)**	**Coefficient (S.E)**	**Coefficient (S.E)**	**Coefficient (S.E)**
log (household income)	1.103***	1.028***	1.141**	1.563*
	(0.252)	(0.334)	(0.483)	(0.803)
previous lottery wins	-0.00001***	-0.00001**	-0.00001**	-0.00001*
	(0.000)	(0.000)	(0.000)	(0.000)
$\hat{\vartheta}_2$	-1.108***	-1.032***	-1.145**	-1.566*
	(0.26)	(0.34)	(0.488)	(0.807)
$\hat{\vartheta}_2 \cdot \ln(M)$	0.011*	0.011*	0.012*	0.012*
	(0.006)	(0.006)	(0.006)	(0.006)

constant	-5.777** (2.53)	-5.017 (3.349)	-6.152 (4.847)	-10.380 (8.05)
Observations	10,328	10,328	10,328	10,328

*Notes: Significance: *** 0.01 level; ** 0.05 level; * 0.1 level. Results of the second stage of the control function model from equation (26). Huber-White heteroscedasticity-robust standard errors.*

A second issue related to use of lottery wins is that a pound (£1) of lottery wins may not have the same effect as a pound (£1) generated through other mechanisms such as earned income. This, however, is also not a problem for the WV approach because valuation in economic theory is interested with how exogenous changes in income/wealth can compensate people in place of the non-market good and so we would want to estimate the effect of unearned income shocks on wellbeing rather than the wellbeing effect of changes in earned income.

Corrections and additions to Chapter 6

1. It is difficult to measure the causal effect of unemployment on wellbeing because there are many factors that will drive both likelihood of unemployment and wellbeing, which may be hard to measure (e.g. motivation, ability, work ethic) which will lead to endogeneity bias. There is also the potential problem of reverse causality, whereby people with lower initial levels of wellbeing become unemployed rather than the other way around.

   I use a self-reported redundancy variable to measure the causal effect of unemployment on wellbeing. As discussed in section 6.3.2.1. there are a number of reasons why redundancy may not be exogenous in some circumstances. In relation to the variable and methods used in this thesis the exogeneity assumption would be violated if (i) the likelihood of being made redundant was driven by factors that also impact on an individual's life satisfaction (and these factors are not controlled for in the analysis), which would lead to endogeneity bias; and/or (ii) the redundancy variable was

reported with error leading to a downward bias in the coefficient on redundancy.

Arguably, since becoming unemployed is a significant event the reasons for unemployment should be clear to people and so measurement error should be minimal. The main challenge for exogeneity in the redundancy variable here will therefore be around unobservable or hard to measure confounding variables. This will include factors such as productivity, motivation, work ethic, risk aversion, how likeable the individual is at work and their soft skills. However, the balance tests conducted in Table 8 of the thesis (which also tested for reverse causality by including a lagged life satisfaction variable) provide support for the exogeneity assumption for the redundancy variable as measured in the BHPS.

2. In Table 2 below I also run the redundant unemployed model in equation (31) with individual fixed effects. Even with fixed effects included in the model there is essentially no change in the size of the coefficient on the redundant unemployed variable which is -0.352 in the original OLS model and -0.349 in the fixed effects model, both significant at the 1% level. However, there is considerable change in the coefficients on all other variables including the income coefficient which reduces significantly as expected.

Table 2. Redundant unemployed model (original OLS and fixed effects)

Independent variable	**(OLS)**		**(Fixed Effects)**	
	Coefficient	**Standard error**	**Coefficient**	**Standard error**
Redundant unemployed	-0.352***	0.059	-0.349***	0.061
Log (household income)	0.153***	0.011	0.030*	0.016
Self-employed	0.118***	0.031	0.070*	0.038
Retired	0.385***	0.038	0.098	0.065
Student	0.011	0.04	-0.044	0.047
Maternity leave	0.397***	0.087	0.270***	0.101
Sick leave	-1.078***	0.083	-0.774***	0.107
Government training	-0.092	0.142	-0.191	0.163
Other job status	-0.165*	0.09	-0.229**	0.092
Constant	3.634***	0.111	4.869***	0.163
Observations	26,849		26,849	

R-squared	0.030	0.024

*Notes: Significance: *** 0.01 level; ** 0.05 level; * 0.1 level. Equation (31) estimated using OLS and with fixed effects. Huber-White heteroscedasticity-robust standard errors. Employed is the reference case for the employment status variables.*

Although not requested in the examiners' comments I also ran equation (32) (with the full set of control variables) with individual fixed effects and found the same result: there was no significant change in the size of the coefficient on redundant unemployed (-0.337 in the original OLS model and -0.328 in the fixed effects model, both significant at the 1% level), but there were significant changes in the coefficients on all other variables.

The results of the fixed effects models provide further support for the exogeneity of the redundant unemployed variable.

3. In equation (33) there is an interactive term to assess whether some variables moderate the impact of unemployment on life satisfaction. This is the term ($MOD_{it} \cdot RU_{it}$). MOD_{it} is a vector of four moderating variables - gender, age, educational attainment and marital status – which are interacted with the redundant unemployed variable RU_{it}. As (33) is an interactive model ($MOD_{it} \cdot RU_{it}$) is added as a separate variable and the four moderating variables (gender, age, educational attainment and marital status) are also included in the vector X_{it} so that we can estimate the additional effect of the moderating term ($MOD_{it} \cdot RU_{it}$).

   Huber-White robust standard errors are used in the model due to potential heteroscedasticity in the data.

Corrections and additions to Chapter 7

None.