I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10907

Back to Bentham, Should We? Large-Scale Comparison of Experienced versus Decision Utility

Alpaslan Akay
Olivier B. Bargain
H. Xavier Jara

JULY 2017

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10907

Back to Bentham, Should We? Large-Scale Comparison of Experienced versus Decision Utility

Alpaslan Akay
University of Gothenburg, IZA and LISER

H. Xavier Jara
University of Essex

Olivier B. Bargain
Aix-Marseille University, CNRS, EHESS and IZA

JULY 2017

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Back to Bentham, Should We? Large-Scale Comparison of Experienced versus Decision Utility*

Subjective well-being (SWB) data is increasingly used to perform welfare analyses. Interpreted as 'experienced utility', SWB has recently been compared to 'decision utility' using specific experiments, most often based on stated preferences. Results point to an overall congruence between these two types of welfare measures. We question whether these findings hold in the more general framework of non-experimental and large-scale data, i.e. the setting commonly used for policy analysis. For individuals in the British household panel, we compare the ordinal preferences either "revealed" from their labor supply decisions or elicited from their reported SWB. The results show striking similarities on average, reflecting the fact that a majority of individuals made decisions that are consistent with SWB maximization. Differences between the two welfare measures arise for particular subgroups, lending themselves to intuitive explanations that we illustrate for specific factors (health and labor market constraints, 'focusing illusion', aspirations).

JEL Classification: C90, I31, J22

Keywords: decision utility, experienced utility, labor supply, subjective well-being

Corresponding author:
Olivier Bargain
Aix-Marseille University
GREQAM
Av. Schuman
13100 Aix-en-Provence
France
E-mail: olivier.bargain@univ-amu.fr

* We are grateful to discussants and participants at the 2017 ASSA meeting, the 2016 RES Conference and various seminars (ISER, DIAL, LISER, AMSE, IZA) for useful suggestions. The usual disclaimer applies.

1 Introduction

In the standard approach to measure well-being in Economics, ordinal preferences are inferred from the observation of decisions made by supposedly rational (utility-maximizing) agents. The object derived from the 'revealed preference' approach is sometimes referred to as a *decision utility*. In contrast, recent advances in psychology, behavioral economics and 'happiness economics' claim that decision utility is unlikely to generate meaningful data on the utility associated with different experiences. These literatures recommend developing measures that focus more directly on *experienced utility* (Dolan and Kahneman, 2008), notably using subjective well-being (SWB) information. What we could call a neo-Benthamian branch of the economic literature now promotes the use of self-reported information on 'happiness' or 'life satisfaction' in order to measure human welfare. In the recent years, a rapidly growing amount of evidence has shown that SWB is not a pure statistical noise and is closely associated with objective measures of well-being and with behavior.[1] Yet, the profession does not seem ready to substitute it for revealed preferences.

In this paper, we aim to compare decision and experienced utility. This is arguably a complicated exercise since actual decisions may depart from the pursuit of happiness for a number of reasons. They may be consistent with other life goals, sometimes incompatible with short-term SWB maximization. They may reflect constraints imposed on the search for individual well-being or be altered by mistakes in decision-making. Köszegi and Rabin (2008) argue that both subjective and choice-based measures of well-being contain unique information on a person's true welfare (so that the ideal measure should make use of both types of data). Nonetheless, it is crucial to question whether there is a minimal consistency between decision and experienced utility. Even in the presence of factors that lead to discrepancies between these measures, we should find a certain degree of 'common support'. The motivation for using SWB could then be that of a better revelation of 'true' utility if decisions are affected by numerous constraints, or just an alternative approach in non-priced/non-revealed situations.

Our contribution consists of a comparison based on a nationally representative dataset (the British Household Panel Survey), focusing on the domains of leisure time and income. We suggest a way to compare, for the same individuals, the ordinal preferences revealed from their labor supply choices versus those elicited from the SWB they report. We recover indifference curves (IC hereafter) depicting the tradeoff between income and non-market time, using estimates from a random-utility model of labor supply on one side and from SWB regressions on the other. Income-leisure preferences are specified in the same way in both approaches and

[1] See Krueger and Schkade (2008) and Oswald and Wu (2010), as well as critical reviews in Senik (2008), Clark et al. (2008), Kahneman and Krueger (2006) or Fleurbaey and Blanchet (2013).

preference heterogeneity is introduced along several socio-demographic variables and personality traits. Keeping in mind that confounding factors can bias the estimation of preferences in both approaches, we suggest a setting where both subjective and revealed preferences are identified using time variation in tax-benefit rules.

This type of comparison remains rare in the literature. Studies in behavioral economics or psychology have explored the possibility of basing economic appraisal on the measurement of experienced utility (Kahneman et al., 1997) – they have contributed to explain some of the difference between experienced and decision utility, notably in the field of public good valuation (Kahneman and Sugden, 2005). More recently, the first explicit comparison has been suggested by Benjamin et al. (2012), who proxy experienced utility using SWB and decision utility using stated preferences. Several other studies have followed that also use hypothetical situations, for instance Clark et al. (2015) who elicit the relative weights placed by people on own income versus others' income. While these studies focus on a single type of decision (e.g. job opportunities in Benjamin et al., 2012), Fleurbaey and Schwandt (2015) confront respondents with a broad set of life choices and Benjamin et al. (2014b) survey a large variety of well-being aspects.

In this line of research, our work contributes to the scarce literature looking at *actual life decisions*, rather than hypothetical life scenarios underlying stated preferences.[2] Moreover, we do so for income-leisure decisions, which is a crucial domain for public economics. Indeed, this is where redistributive policies like taxes and benefits operate. Quite surprisingly, however, the bulk of the literature on welfare analysis and optimal tax design has ignored preference heterogeneity – and when taking it into account, it has relied almost exclusively on decision utility (see the discussion in Decoster and Haan, 2014, and an exception with the innovative work of Decancq et al., 2015). Finally, and most importantly, the present work is an original attempt to compare decision and experienced utility using *large-scale and non-experimental surveys*. In this way, it is very complementary to Benjamin et al. (2012). Indeed, it takes the decision-experienced utility comparison beyond the experimental domain and brings it in the setting traditionally used for policy and welfare analyses. A natural limitation is that we cannot experimentally control and manipulate the parameters that possibly explain why people do not maximize SWB. Yet, we consider heterogeneous effects that lead to suggestive interpretations regarding these factors – and we provide very extensive sensitivity checks of our comparison.

Our results first show that empirical ICs from decision utility versus experienced utility are broadly consistent with economic theory. Both approaches produce downward sloping and

[2] Other studies also look at actual choices: Benjamin et al. (2014a) consider residency choices, Fleurbaey and Schwandt (2015) study a whole set of decisions that can potentially affect SWB, and Perez-Truglia (2015) investigate consumption decisions.

convex ICs, reflecting income-leisure tradeoffs in a textbook fashion. Most importantly, ICs are strikingly similar in both approaches, conveying that people behave *on average* as if they were maximizing SWB. Our extensive robustness analysis confirms this finding (we check the sensitivity of these results to the options taken in our baseline including the SWB measure, the model functional forms, the way to 'clean' SWB data or the treatment of unobserved heterogeneity). We then consider the distribution of 'projection errors', the differences between actual and SWB-maximizing choices. Cases where revealed and subjective preferences diverge illustrate the extent to which observed decisions reflect other things than pure preferences, including constrains and possibly non-optimizing behavior. Because of these factors, subjective preferences might be closer to authentic preferences and could serve as a benchmark to tell us how much forecasting errors people make (see Odermatt and Stutzer, 2015). Consistently with the IC analysis, 'projection errors' are centered around zero and confirm that a majority of individuals actually make decisions that are in line with the maximization of income-leisure satisfaction. It is interesting to consider specific sub-groups, which carry intuitive interpretations. For instance, revealed "preferences" for men, Londoners or more educated workers put a relatively lower weight on leisure compared to subjective preferences. Finally, we conduct IC comparisons for groups possibly affected by specific constraints (such as labor market rigidities), by 'focusing illusion' or by excessive aspirations. A pattern emerges that shows how these factors are systematically responsible for the gap between both approaches. We derive implications for future research.

2 Data and Empirical Framework

2.1 Data and Sample Selection

Our analysis is based on data from the British Household Panel Survey (BHPS), a nationally representative survey collected in the United Kingdom between 1991-2008 and containing life satisfaction information since 1996. The dataset additionally provides standard information about individual and household characteristics to be used in our estimations (gender, age, education, labor market status, health). Important variables are missing for the years 2006-7, so we focus on the period 1996-2005.

We restrict our analysis to single individuals for simplicity. In the future, the literature on collective models will certainly allow retrieving completely the ordinal preferences of each spouse in a joint labor supply decision model. As yet, this is not the case, at least not without strong additional assumptions (see Chiappori and Donni, 2011, for a recent statement). As commonly

applied in the labor supply literature, we further exclude individuals in self-employment,[3] or not available for the labor market (disabled individuals, full-time students and pensioners). To comply with the labor supply nature of the model, our baseline excludes job seekers[4], but retains other inactive persons. Finally, we only retain individuals for whom all key characteristics (including socio-demographics and personality traits) are available for all years. Our selected sample includes $5,501$ person $\times$ year observations.

The key variables for our analysis are working hours and disposable income. Weekly working hours drawn from the data are denoted h_{it} for agent i at time t. Disposable income y_{it} is calculated as:

$$y_{it} = G_t(g_{it}, \mu_{it}, \zeta_{it}), \tag{1}$$

using reported gross labor income g_{it}, unearned income μ_{it} and a set of individual characteristics ζ_{it}. Function G_t represents the aggregation of all incomes and the imputation of taxes and benefits, using numerical simulations of tax-benefit rules of each period $t = 1, \ldots, T$. The set ζ_{it} represents individual characteristics that matter for tax-benefit calculations and are extracted from the data, for instance the presence of children (which conditions the calculation of child benefits, increment of income support, tax credits, etc.).

In addition, we extract information on many other individual characteristics that have been used in the literature as taste shifters in structural labor supply models and/or determinants of SWB, including gender and age, being single, widowed or divorced, health status (from very good to very poor), educational level (elementary school, high school or university), being a native or immigrant, ethnicity, household composition (including a dummy for the presence of children aged 0-2), living in London and personality traits, i.e. the so-called 'big five' (conscientiousness, neuroticism, openness, extraversion and agreeableness), measured on 1-4 scales.

'Decision' utility upon income and leisure, denoted V^D hereafter, is going to be a latent variable estimated in a structural labor supply model – as explained in detail below. 'Experienced' utility upon these two domains, denoted V^E, is retrieved from SWB data. The primary information is the answer to the life satisfaction question: "How dissatisfied or satisfied are you with your life overall?" It is measured in an ordered scale between 1 ("not satisfied at all") and 7 ("completely satisfied"). Ideally, we would like to have a question about the relative well-being drawn from the balance of income and leisure. Our data contains satisfaction on life domains that can be combined for this purpose (see also van Praag et al., 2003). We rely on satisfactions with

[3] Their labor supply decisions may considerably differ from those of salaried workers. Also, in their case, income information from surveys is considered less reliable.

[4] Individuals must answer affirmatively to the following two questions in the data: (1) "Have you actively looked for a job within the last four weeks?" (2) "Are you ready to take up a job within the next two weeks?".

income and leisure (also on 1-7 scales) as obtained from the questions "How dissatisfied or satisfied are you with the income of your household?" and "How dissatisfied or satisfied are you with the amount of leisure time you have?". We opt for a simple way to extract the variation in overall life satisfaction (denoted S_{it}) that stems from variation in income satisfaction (S_{it}^y) and leisure satisfaction (S_{it}^l). It consists in the estimation of the simple linear model:

$$S_{it} = \gamma^y S_{it}^y + \gamma^l S_{it}^l + e_{it},$$

and the use of estimated weights on each domain to compute an "income-leisure concentrated" measure $V_{it}^E = \widehat{\gamma}^y S_{it}^y + \widehat{\gamma}^l S_{it}^l$.[5] Conceptually, this is a measure of experienced utility more closely related than life satisfaction to the decision utility over income and leisure. We nonetheless use life satisfaction itself in sensitivity checks ($V_{it}^E = S_{it}$). We also experiment with alternative functional forms of the function $V_{it}^E = V^E(S_{it}^y, S_{it}^l)$ (introducing heterogeneity and more flexibility, as explained in detail later). We finally apply the same combination strategy using alternative SWB measures, namely happiness and mental health indices.[6]

2.2 Utility Estimation using Labor Supply Choices

We aim to compare ordinal preferences that are implicit in either SWB data or labor supply choices. We suggest using standard estimation methods in each case, while keeping functional forms as similar as possible (specifications are presented hereafter). We first elicit income-leisure preferences from actual labor supply choices, referring to the standard static behavioral model based on decision utility.

Modern techniques to estimate labor supply with taxation strongly rely on the discretization of the work options (e.g., Blundell et al, 2000, van Soest, 1995). We adopt this approach. An agent i at period t is assumed to face J income-labor pairs, denoted (y_{ijt}, h_{ijt}), $j = 1, ..., J$, and to choose the one maximizing utility. We opt for a thin discretization with $J = 7$ options corresponding to weekly work hours h_{ijt} from 0 ($j = 1$) to 60 ($j = 7$) with a step of 10 hours. We normalize total time available for work to 80 hours per week so that leisure is $l_{ijt} = 80 - h_{ijt}$,

[5] The two dimensions play a relatively balanced role, as we find that $\widehat{\gamma}^y/(\widehat{\gamma}^y + \widehat{\gamma}^l) = 0.468$.

[6] The latter is drawn from the General Health Questionnaire, GHQ-12, as used in several studies (e.g., Clark and Oswald, 1994; van Praag et al., 2003). It comprises 12 answers recoded on a 0-3 scale so that the GHQ ranges from 0 (lowest mental health) to 36 (highest). Happiness is measured from the question "Have you recently been feeling reasonably happy, all things considered?" with four answers recoded as: 4- more than usual, 3- same as usual, 2- less than usual and 1-much less than usual. Linear correlation between life satisfaction and GHQ-12 (resp. happiness) is 0.587 (resp. 0.462). Note that the concentrated measures obtained in these cases are "hybrid" since they extract variation in GHQ/happiness from income and leisure *satisfaction*.

ranges from 80 to 20 hours per week.[7] Thus, the utility level derived by individual i from option j at time t is written:

$$V_{ijt}^D = U_{it}^D \left(y_{ijt}, l_{ijt}; x_{it}\right) + \eta_{ijt}^D, \tag{2}$$

The deterministic utility function U_{it}^D is conditional on a vector x_{it} of individual (binary) characteristics that possibly influence work preferences, including male, age above 40, higher education, presence of children aged 0 to 2, living in London, non-white ethnic origin, migrant, above-average conscientiousness and above-average neuroticism.[8] As usual in this literature, the random component η_{ijt}^D is assumed to be i.i.d. and to follow an extreme value type I (EV-I) distribution, such that the probability to observe individual i actually choosing the alternative j at time t has an explicit conditional logit form. The latter is used to construct the likelihood for maximum likelihood (ML) estimations of the utility function U_{it}^D.

The model is built under the assumption of utility maximization, with individuals choosing among the (discrete) set of hours alternatives. This is a mere application of the revealed preference approach, which requires variation in prices. Cross-sectional variation in wage rates is typically used in this literature. Most policy studies use one year of data for convenience (sometimes because of data limitation). Yet, omitted variables, e.g. being a hard working type, may affect simultaneously work preferences and gross wages. A quasi-experimental approach retained in the literature consists in exploiting the exogenous variation in *net* wages due to changes in tax-benefit policies, following Blundell et al. (1998).[9] Since we pool 10 years of data, we obtain sufficient variation in tax-benefit schedules, compounded with spatial variation (council taxes are specific to England, Scotland, Wales and Northern Ireland). Indeed, the British tax-benefit system has experienced deep changes over the years under study, notably with the important reforms undertaken by the "New Labour" government regarding income tax, social insurance contributions, council taxes, income support and tax credits for working

[7] Disposable income $y_{ijt} = G_t(w_{it}h_{ijt}, \mu_{it}, \zeta_{it})$ is microsimulated at each option j using gross hourly wage rates w_{it} and discretized values h_{ijt}, unearned income and household characteristics, as described above. For hourly wage rates, we follow the standard approach, i.e. we calculate them as weekly earnings divided by worked hours for workers then use this information to estimate Heckman-corrected wage equation (instrument is non-labor income and the presence of children aged 0-2) in order to predict a wage rate w_{it} for non-workers.

[8] Among the 'big five', these two are shown to be those that matter for labor supply choices (see Wichert and Pohlmeier, 2010). Neuroticism is a fundamental personality trait in the study of psychology characterized by anxiety, fear, moodiness, worry, envy, frustration, jealousy, and loneliness. Conscientiousness is the personality trait of being thorough, careful, or vigilant, implying the desire to do a task well.

[9] This idea has also been used to identify the elasticity of taxable income (see Saez et al., 2012 for a survey). Note that spatial variation in tax-benefit rules has also been suggested (for instance by Hoynes, 1996, using variation in tax schedules across US states). Arguably, tax incentives also affect geographical mobility. Time variation in tax-benefit rules seems a more exogenous source of identification, possibly combined with spatial variation.

poor families (an extensive description of these reforms can be found in Blundell et al., 2000, and Adam and Browne, 2010).

2.3 Utility Estimation using SWB

Next, we estimate SWB equations in order to retrieve the ordinal preferences consistent with experienced utility. The latter, denoted V_{it}^E for individual i at period t, is proxied by life satisfaction or our concentrated income-leisure satisfaction measures, as explained before. It is modelled as a deterministic function of income-leisure utility, $U_{it}^E(y_{it}, l_{it})$, and additional controls η_{ijt}^E for individual heterogeneity in well-being responses:

$$\begin{aligned} V_{it}^E &= U_{it}^E(y_{it}, l_{it}; x_{it}) + \eta_{it}^E, \\ \text{with } \eta_{it}^E &= \lambda' z_{it} + \alpha_i + \epsilon_{it}. \end{aligned} \tag{3}$$

For a consistent comparison of both approaches, function U_{it}^E is specified exactly as U_{it}^D. Contrary to the labor supply model, SWB regressions only require information on the (y_{it}, l_{it}) pair actually chosen by the agent. With actual weekly working hours h_{it} and our normalization, leisure is $l_{it} = 80 - h_{it}$. The model can be estimated by standard linear estimation methods or, when function U_{it}^E is nonlinear (for instance in box-cox specifications), by ML.[10]

SWB measures may reflect individual heterogeneity in the way people perceived and/or report levels of well-being, which makes it hard to assume interpersonal comparability in order to extract subjective preferences on income and leisure. To "clean" SWB measures, we model well-being heterogeneity η_{it}^E. It first comprises z_{it}, a vector of the usual determinants of well-being found in the literature (cf. Clark et al., 2008).[11] Then, unobserved heterogeneity α_i can be proxied in several ways. Typically, it is estimated as fixed effects using panel data. This is not something we can easily reproduce in the 'decision utility' approach, however, so that our comparison would be biased (estimation of U_{it}^E would rely on a within-estimator while U_{it}^D would be identified on within and between variation).[12] Hence, we put some structure

[10] The (predicted) concentrated income-leisure satisfaction is a continuous variable. When using life satisfaction, measured on a 1-7 scale, we also treat this information as continuous. Alternatively using an ordered logit approach does not change the results (see also Ferrer-i-Carbonell and Frijters, 2004).

[11] More than the binary variables x_{it} used for preference heterogeneity, we include continuous variables for age (and age squared), family size, health status, home ownership, region and year.

[12] In the decision model, the estimation relies on all pairwise comparisons of work alternatives, as presented in equation (**??**), so that parameters are identified only if they vary with these alternatives. Hence, additively separable terms like α_i are not identified. Even if interacted with choices, fixed effects pose the problem of incidental parameter bias in nonlinear models. In the Appendix, we nonetheless provide comparisons when using fixed effects in SWB estimations. There, we also experiment with random and quasi-fixed effects, which make things more comparable since both within and between variations are used in this case.

on individual effects by assuming they can be proxied by detailed personality traits (not only binary conscientiousness and neuroticism, as used in x_{it}, but more generally using all the 'big five' on a 1-4 scale). These factors usually account for an important part of the individual variation in SWB (Boyce, 2010, Ravallion and Lokshin, 2001). Finally, ϵ_{it} is an i.i.d., normally distributed error term.

In SWB estimations, there is a concern regarding the role of heterogeneity in work preferences. Some of it is lifted by the fact that we introduce observed heterogeneity in the form of characteristics x_{it} in preference parameters. In robustness checks, we shall also add a random term in the coefficient for leisure. Yet, the latter is assumed to follow a normal distribution and just play the role of a random effect. In fact, it is easy to show that estimates of the utility function U_{it}^{E} will be biased if actual unobserved heterogeneity in work "preferences" (ex: having the moral obligation to work a lot to support the family) is correlated with other unobserved determinants of well-being (ex: being the morally obliged father is a source of stress in general). We assume that this correlation can be ruled out using two arguments. First, we account for some of the personality traits (conscientiousness and neuroticism) in both work preference parameters x_{it} and separately additive well-being terms z_{it}. Second, as in the case of labor supply decisions, the potential role of omitted variables can be addressed using exogeneous variation stemming from policy reforms. Precisely, two identical individuals may not make the same labor supply choice because they face different work incentives due to different tax-benefit schedules.

2.4 Specification

We have insisted on the need for a common specification of the income-leisure utility function. Discrete labor supply models do not require tangency conditions but only the comparison of utility at discrete choices, so that preferences can in principle remain very general. Well-known applications make use of a translog model (van Soest, 1995) or a quadratic specification (Blundell et al., 2000) with fixed costs of work. By contrast, relatively simple functional forms are used in the SWB literature, usually linear or log-linear in income (to capture the concave relationship with well-being). Adding leisure, Knabe and Rätzel (2010) use a log form on income and a linear or quadratic form for leisure, without interaction terms. One consideration is that SWB information is noisier than actual labor supply decisions, resulting in less precisely estimated coefficients on income or leisure in the SWB model (especially if too many interaction terms in income and leisure are used). Consequently, we adopt an intermediary position for our baseline, relying on a quadratic form in both net income and leisure:

$$U_{it}^{m}\left(y_{it}, l_{it}; x_{it}\right) = \beta_{yy}^{m} y_{it}^{2} + \beta_{ll}^{m} l_{it}^{2} + \beta_{y}^{m} y_{it} + \beta_{l}^{m}(x_{it}) l_{it} \quad \text{for } m = D, E. \tag{4}$$

We shall check the sensitivity of our results to alternative parametric forms, including the log-linear utility:

$$U_{it}^{m}\left(y_{it}, l_{it}; x_{it}\right) = \beta_{y}^{m} \ln y_{it} + \beta_{l}^{m}(x_{it}) \ln l_{it}$$

often used in SWB studies (e.g., Clark et al., 2008) and capturing some non-linearity in income and leisure, and the box-cox utility:

$$U_{it}^{m}\left(y_{it}, l_{it}; x_{it}\right) = \beta_{y}^{m} \left(\frac{y_{it}^{\lambda_y} - 1}{\lambda_y}\right) + \beta_{l}^{m}(x_{it}) \left(\frac{l_{it}^{\lambda_l} - 1}{\lambda_l}\right)$$

used in numerous empirical work (recently in labor supply studies focusing on welfare analysis, see Decoster and Haan, 2014). In all these models, preference heterogeneity is accounted for by linearly varying the leisure term with taste shifters x_{it}:

$$\beta_{l}^{m}(x_{it}) = \beta_{l0}^{m} + \beta_{l1}^{m\prime} x_{it}$$

for $m = D, E$. We shall discuss alternative specifications with, in particular, observed heterogeneity introduced in other coefficients of the model. Finally, note that we do not account for fixed costs of work. With revealed preferences, costs of work are usually not identified from preferences (or only under strong parametric assumptions, cf. Blundell et al., 2000). They may be even more difficult to identify in SWB regressions. Maybe more fundamentally for welfare analyses, work costs are difficult to interpret and can be seen as either part of the budget constraints (e.g. actual expenditure on transportation or childcare) or part of an individual's preferences (e.g. negative costs corresponding to the psychic pain of being unemployed and staying at home all day, see Clark and Oswald, 1994, and our Appendix). For this reason, we prefer to keep our framework simple. Further work could pursue the current comparison under different choices regarding fixed costs interpretations.

3 Results

3.1 Decision and Experienced Utility Functions

We first present estimated parameters for labor supply and SWB models. We highlight the fact that estimates of decision utility and of experienced utility are not directly comparable since these two welfare measures rely on different implicit scales. At this stage, our aim is simply to compare the signs and significance of key variables appearing in both models. Hence, we use our baseline quadratic specification with preference heterogeneity x_{it} on leisure only, for a direct interpretation of taste shifters on the marginal rate of substitution (MRS) between income

Table 1: Preference Estimates from Labor Supply and Subjective Well-Being

Coefficients	Utility (Labor Supply)	Subjective Well-being			
		Life Satisfaction	Income-Leisure Concentrated Satisfaction	Income-Leisure Concentrated GHQ	Income-Leisure Concentrated Happiness
	(A)	(B)	(C)	(D)	(E)
Income^2	-1.87e-05***	-4.67e-07***	-4.82e-07***	-1.63e-06***	-1.24e-07***
	(8.30e-07)	(1.54e-07)	(9.12e-08)	(2.83e-07)	(2.10e-08)
Income	0.0282***	0.000955***	0.00120***	0.00424***	0.000324***
	(0.00106)	(0.000227)	(0.000135)	(0.000418)	(3.11e-05)
Leisure^2	-0.00160***	-1.48e-05	-7.27e-05*	-0.000220*	-1.62e-05*
	(5.69e-05)	(6.56e-05)	(3.88e-05)	(0.000121)	(8.96e-06)
Leisure	0.263***	0.00251	0.00975**	0.0272*	0.00196*
	(0.00761)	(0.00781)	(0.00463)	(0.0144)	(0.00107)
x male	-0.0404***	-0.00180	0.00266	0.00794	0.000584
	(0.00237)	(0.00287)	(0.00170)	(0.00527)	(0.000392)
x over 40	-0.00127	0.00133	0.00100	0.00284	0.000206
	(0.00207)	(0.00105)	(0.000622)	(0.00193)	(0.000143)
x high educ.	-0.0216***	0.00200*	6.62e-05	0.000497	4.20e-05
	(0.00310)	(0.00114)	(0.000673)	(0.00209)	(0.000155)
x young kid	0.0975***	0.0165**	0.00274	0.00839	0.000622
	(0.00801)	(0.00680)	(0.00403)	(0.0125)	(0.000930)
x london	0.00860**	0.00729*	0.00699***	0.0212***	0.00157***
	(0.00434)	(0.00399)	(0.00236)	(0.00733)	(0.000545)
x non-white	-0.0226***	-0.00780	-0.00738*	-0.0229*	-0.00170*
	(0.00758)	(0.00712)	(0.00422)	(0.0131)	(0.000973)
x migrant	0.00136	0.000531	-0.00155	-0.00528	-0.000400
	(0.00645)	(0.00627)	(0.00372)	(0.0115)	(0.000857)
x conscientious	-0.0101***	-0.00307***	-0.00148**	-0.00467**	-0.000348***
	(0.00206)	(0.000986)	(0.000584)	(0.00181)	(0.000135)
x neurotic	0.00322	-0.00260***	0.000544	0.00173	0.000129
	(0.00206)	(0.000929)	(0.000550)	(0.00171)	(0.000127)
Log-likelihood	-12,909.25				
Pseudo R^2 / R^2	0.136	0.229	0.243	0.251	0.253
#Obs	5,501	5,501	5,501	5,501	5,501

Notes: Subjective well-being equations (columns B to E) also include additively separable controls (same variables as in leisure interaction terms plus age squared, family size, health status, home ownership, all personality traits, region and year dummies). *, **, *** indicate 1%, 5% and 10% significance levels. Standard errors in parenthesis.

and leisure. A more general characterization of the income-leisure tradeoff will be provided by comparing ICs in the next sub-section.

The results in Table 1 indicate that $U^m(y,l)$ is increasing and concave in income and leisure with both $m = D$ (the labor supply model in column A) and $m = E$ (the subjective measures of experienced utility in columns B-E). All four terms on income and leisure are statistically significant for the labor supply model and for the main SWB measure (the "concentrated" satisfaction in column C). This is also the case in sensitivity checks using concentrated GHQ and happiness (columns D and E). General life satisfaction (column B) provides a similar pattern, although the leisure terms are insignificant. As indicated by the R^2, concentrated measures fit the data slightly better than life satisfaction. The fact that the concentrated measures yield implicit preferences most similar to choice-revealed preferences is reassuring, showing that our preliminary "concentration" managed to extract well-being variation in the relevant domains of income and leisure. Overall, the first conclusion is that preferences inferred from both decisions and SWB seem consistent with economic theory.

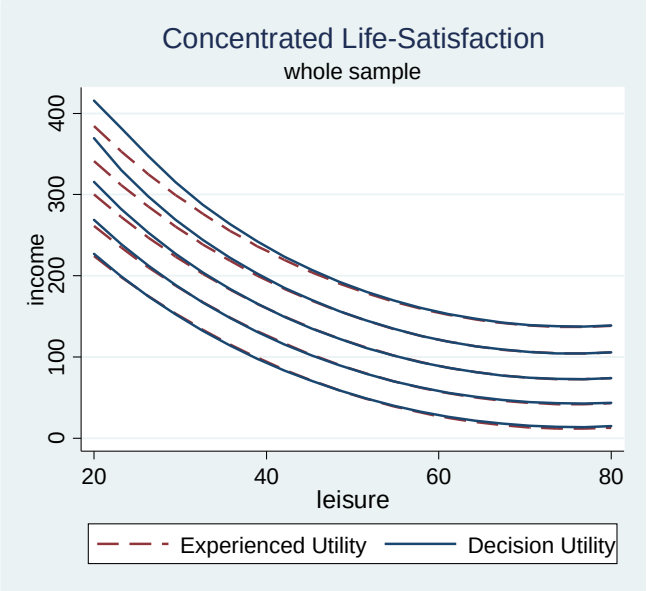
We turn to preference heterogeneity, as provided by observed characteristics in x_{it}. It corresponds to dichotomous groups, which makes interpretations easier. Other things being equal (and in particular the net wage), heterogeneous work preferences lead to different labor supply choices across groups (for instance, highly conscientious people will work more), which is indeed rationalized by the labor supply model (a negative coefficient on leisure $\times$ conscientious in column A). The SWB regression may or may not give the same pattern, depending on the type of characteristic that we consider. In our example, conscientious workers may work long hours and, at the same time, derive less (experienced) disutility than average from doing so. Indeed, this is what we observe (a negative coefficient on leisure $\times$ conscientious in columns B-E). Preference heterogeneity is similar across approaches – significant and of the same sign – for Londoner, non-white and conscientiousness. For the rest, there are no conflicting results with significant coefficients (except high education for labor supply versus life satisfaction). As expected and discussed above, taste shifters are less often significant in the case of SWB.

3.2 Patterns of Indifference Curves

We then proceed with the calculation of ICs in the income-leisure space. From this point onwards, we use only the concentrated satisfaction measure, which better fits the data and best represents the underlying utility from income and leisure experiences. Using estimates of models (2) and (3), ICs are obtained by inverting $\overline{U}^D = U^D(y,l)$ and $\overline{U}^E = U^E(y,l)$ to retrieve income as a function of leisure. We average all individual ICs drawn through a common point set at 40 hours of leisure and $\overline{y}(40)$ (the sample mean disposable income at 40 hours). We show mean ICs and those shifted by $\pm 10\%$ and $\pm 20\%$.

Baseline. In Figure 1, the solid curves represent the IC derived from the labor supply model while the dashed curves represent the IC from the SWB model. Recall that weekly leisure points range from 20 to 80 hours, corresponding to weekly work hours from 60 (overtime) to 0 (inactivity). We use our baseline quadratic specification without any restriction on the parameters. Nonetheless, the resulting average ICs comply with economic theory, i.e. displaying a monotonically decreasing and convex pattern across the leisure range. The same is true when looking at specific sub-groups of the population as we do next. The most interesting result here is the striking similarity between ICs derived from SWB measures and those from labor supply decisions. In effect, since choice-revealed preferences coincide with SWB-revealed preferences, people behave on average as if they maximized SWB.

Figure 1: Baseline Comparison of Indifference Curves (IC)



Note: Indifference Curves are obtained using estimates of concentrated life satisfaction (experienced utility) and labor supply (decision utility) on income and leisure. We use a quadratic specification with preference heterogeneity (male, age, education, presence of young kid, London, non-white, migrant, conscientious, neurotic). These variables as well as additional controls (age squared, family size, health status, home ownership, all personality traits, region and year dummies) enter the SWB equation as additively separable controls (hence, not affecting the calculation of ICs). Graphs are obtained by averaging all individual ICs, for each approach separately, drawn through a common point, defined as $(\overline{y}(40), 40)$ for the central IC and $\pm 10\%$ and $\pm 20\%$ variation in utility for the other ICs.

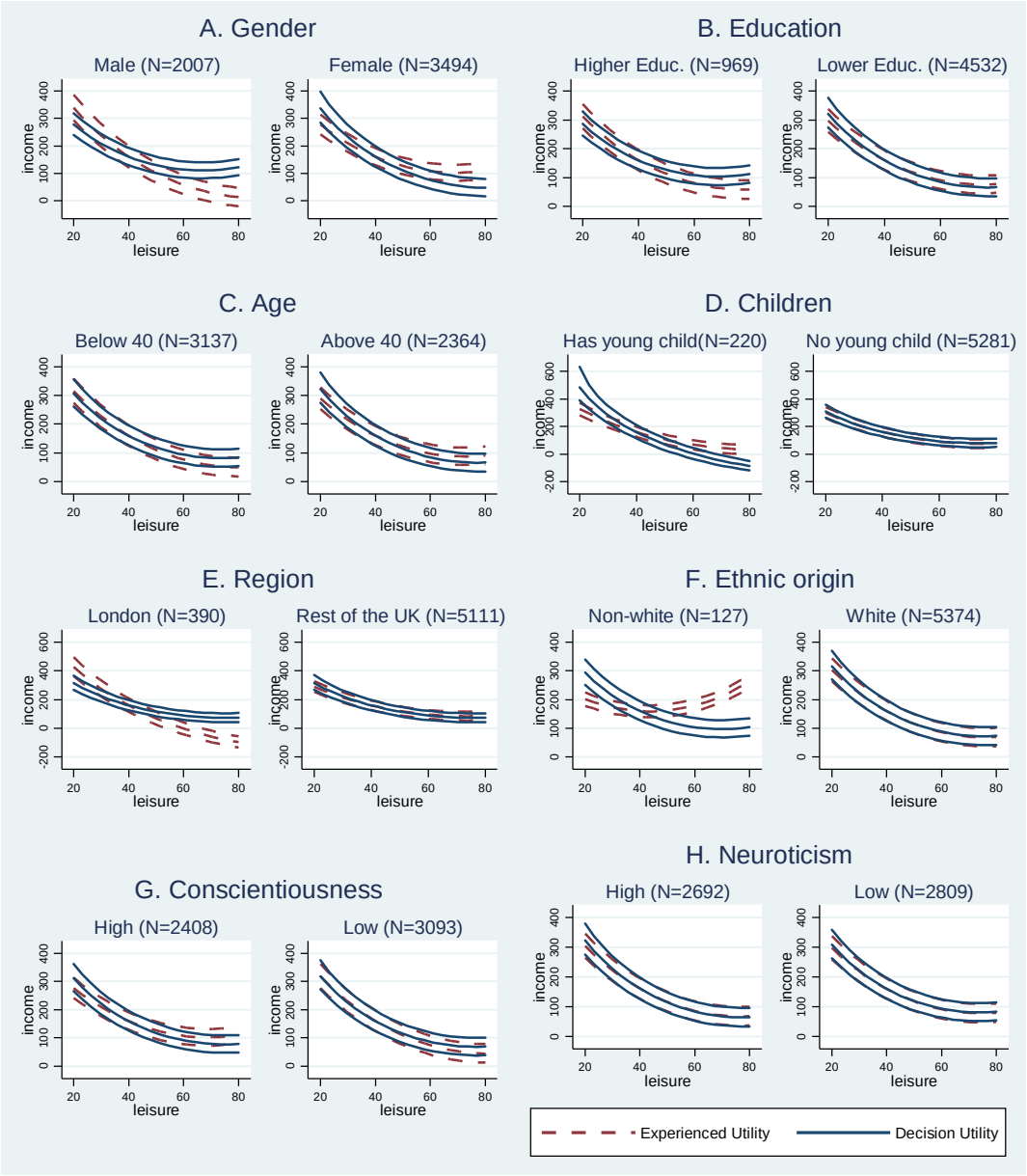
Sensitivity Analysis. The complete sensitivity check is reported in the Appendix and summarized here. First, the way we predict our 'concentrated income-leisure satisfaction' measure does

not affect the results. The checks include specifications of $V_{it}^{E} = V^{E}(S_{it}^{y}, S_{it}^{l})$ where observed heterogeneity or nonlinearity are introduced in function $V^{E}()$. Second, the way we introduce observed and unobserved preference heterogeneity in the experienced and decision utility functions $U_{it}^{m}(y_{it}, l_{it}; x_{it})$, $m = E, D$, does not affect the overall aspect of ICs. Third, the shape of ICs changes, as expected, when moving to log-linear and box-cox utility specifications. Yet the great similarity between subjective and revealed preferences remains in both cases. Fourth, we confirm that individual SWB measures need to be "cleaned" from individual-specific circumstances to recover a meaningful preference structure (see also Decancq et al., 2015). This requires a complete specification of z_{it} (notably with information on health and family status) and α_i (using personality traits or panel time-invariant effects). Finally, the explicit treatment of involuntary unemployment by double hurdle models – when using a larger sample including job seekers – provides results which are closely comparable to the baseline in which we have excluded job seekers.

Preference heterogeneity. We turn our attention to the dichotomous characteristics x_{it} used to elicit preference heterogeneity. Results are presented by means of ICs in Figure 2; we also report MRS between income and leisure (point estimates and standard errors) taken at point $(\overline{y}(40), 40)$ in Appendix Table B.1. There is a substantial heterogeneity with respect to income-leisure preferences, which can be tentatively interpreted. For instance, the fact that men work more than women is rationalized in the revealed preference approach by putting less weight on leisure in the former group (i.e. a flatter IC for men). Yet, the fact that men work more may only partly reveal difference in preferences (e.g. difference in career-orientation, preference for status, etc.). For instance, it may reflect discrimination and constraints on the labor market (e.g. forced part-time for women, especially for single mothers present in our sample of singles). This set of explanations leads to visible divergences between decision and experienced utility in Figure 2. SWB models reveal that men (women) require more (less) compensation from having little leisure than what labor supply conveys. A similar pattern emerges for other characteristics. Like men, highly educated workers or Londoners seem to work 'too much'. Like women, single parents or non-white workers seem to work 'too little'.

In conclusion, actual choices and SWB lead to similar ICs on average but show discrepancies when particular population groups are considered. These differences seem to follow an intuitive pattern, i.e. to be related to factors that hinder the maximization of SWB (status, family duties, labor market constraints, etc.). In the last subsection, we attempt to single out some of these factors using specific heterogeneity from the BHPS data.

Figure 2: Comparison of Indifference Curves: Group-Level Preference Heterogeneity



Note: Indifference Curves are obtained using estimates of concentrated life satisfaction (experienced utility) and labor supply (decision utility) on income and leisure. We use a quadratic specification with preference heterogeneity (and other additively separable controls in the SWB equation) as specified in Figure 1. Graphs are obtained by averaging all individual ICs drawn through a common point, defined as $(\overline{y}(40), 40)$.

3.3 Alternative Comparison Strategies: What do People Maximize?

Projection Errors. Another way to present our results is to test the extent to which actual choices coincide with SWB maximization. Estimates of the SWB equation are now used to predict SWB levels for each discrete labor supply alternative. Then, we calculate 'projection errors' as the difference between actual and SWB-maximizing choices. This terminology is borrowed from Loewenstein et al. (2003) and Loewenstein and Adler (1995). It implies a particular interpretation whereby SWB-maximizing errors represent failures of individuals to predict the future satisfaction levels resulting from their choices (labor supply decisions took place before the record of their SWB consequences).[13] Yet, as discussed above, 'errors' cannot be taken *prima facie* – especially if people do not aim to maximize SWB or cannot do so because of health or labor market constraints, moral obligations of being the wage earner or the main carer in the household, etc.

We calculate 'projection errors' for each individual×period in our sample. We find an average 'error' of 4.1 weekly hours, which is not significantly different from zero. This point value is well below the step of discretization (10 hours) used in the labor supply model. This result is coherent with our IC characterization, i.e. on average, labor supply choices are consistent with SWB maximization. Figure 3 also shows a single-peaked distribution of 'errors' with a mode at zero and a standard deviation (*std*) of 21.4 hours. We find that around 54% of the 'errors' are below 10 hours, i.e. in the range of a half *std* interval, which can be attributed to discretization. That is, a short majority of people are close to maximization.[14]

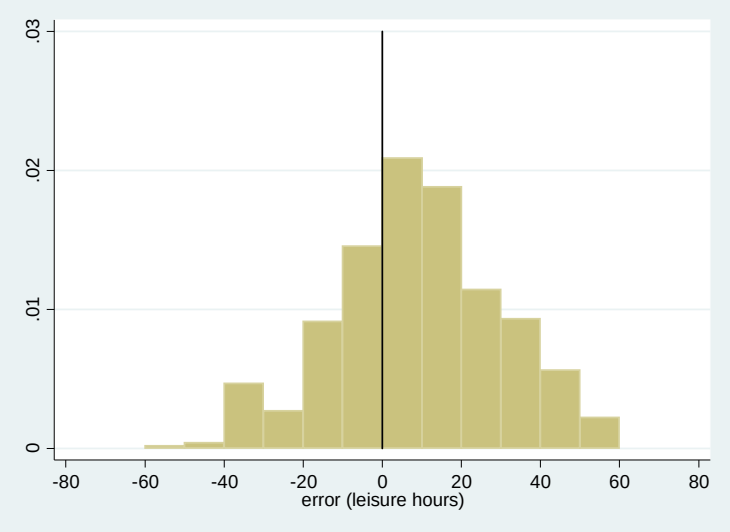
Comparison with the Literature. This is relatively remarkable in a non-experimental context. Indeed, while our results are not based on a controlled experiment, they tend to be similar to other studies showing that a majority of observed choices are consistent with the pursuit of individual satisfaction. In particular, Benjamin et al. (2012) show that most (but not all) individuals are able to predict their SWB at the moment of deciding about (hypothetical) job opportunities. Benjamin et al. (2014a), looking at actual residency choices, show that SWB scores are correlated with the ranking of actual choices (even if the tradeoffs between aspects of residency tend to be different). Fleurbaey and Schwandt (2015) ask people if they can think of changes that would increase their SWB score. About 60% cannot think of an easy improvement, i.e. feel as if they currently maximized SWB. Clark et al. (2015) also find similar relative concerns in happiness regressions and in hypothetical-choice experiments. Hence, our

[13] In line with Dolan and Kahneman (2008), Odermatt and Stutzer (2015) have recently used panel data to show that people tend to make systematic prediction errors regarding the future impact of major events on their life satisfaction, partly because of unforeseen adaptation. See also Frijters (2000) and Frijters et al. (2009).

[14] Around 70% (resp. 83%, 95%) have an error within ±1 *std* (resp. 1.5, 2 *std*).

results are in line with the optimistic view that there is an overall congruence between revealed and subjective preferences.[15]

Figure 3: Distribution of the 'Projection Error'



Note: 'projection errors' are calculated as the difference between actual and SWB-maximizing leisure choice.

3.4 Suggestive Explanations for Discrepancies

To proxy some of the factors explaining differences between "revealed" preferences and subjective preferences, we extract additional variables from the BHPS. There is arguably no perfect proxy for each factor affecting the relationship between actual choices at the time of decision and outcomes at the time of hedonic experience (each variable at use possibly lends itself to several interpretations). Hence, our reading of the results should only be taken as a first attempt to pinpoint some of the underlying mechanisms. Results are presented in the form of ICs in Figure 4, constructed as the average IC in each sub-group. We also report MRS between income and leisure at point $(\overline{y}(40), 40)$ (with standard errors) in Appendix Table B.2.

Labor Market Constraints. We first inspect differences between approaches in case of high versus low tensions on the labor market. We simply use variation in local unemployment rates across 12 regions $\times$ 10 periods. Recall that we consider a baseline sample without jobseekers. Nonetheless, we may be in the presence of frictions regarding work hours. Results in Figure

[15] Yet, it is worth stressing that other studies are a bit more negative about it. Ferrer-i-Carbonell et al. (2011) compare the estimates on job characteristics in choice equations using vignettes to those on the same characteristics in determining the respondent's own job satisfaction, finding significant differences. Perez-Truglia (2015) shows that real consumption is well predicted by life satisfaction but not by economic satisfaction.

4A convey that ICs are less comparable between approaches in highly constrained regions (the top quintile of the unemployment rate distribution) and that workers tend to work 'too much' in this case. We see three explanations for this result. First, high-unemployment regions may correspond to low-wage workers, who might have lower work "preferences" than average (due to actual preferences, low work quality, etc.) and hence are deemed working 'too much' even when they work a regular 40 hour workweek. Second, people in work may refrain from changing jobs, i.e. to adjust their working time in a SWB-maximizing way, due to high employment insecurity in these regions. In other words, their observed choice would be a second-best but a better choice than inactivity. Third, inactive workers may be less stigmatized in these regions (lower compensation at zero work hours), given than unemployment is more frequent.

Individual-Specific Constraints. We examine two other examples of optimization constraints that are discussed in studies comparing decision and experienced utility, namely health status and commuting time (see Dolan and Kahneman, 2008; Frey and Stutzer, 2008). Figure 4B compares ICs for individuals with poor and good health, respectively. ICs elicited from choices versus SWB seem to differ very strongly for constrained people only, i.e. those with poor health. For them, long work hours come with pain and require income compensation, as rationalized by the steeper IC derived from SWB. In the same way, Figure 4C conveys that working long hours requires much more compensation when commuting time is high (top quintile of the commuting time distribution).

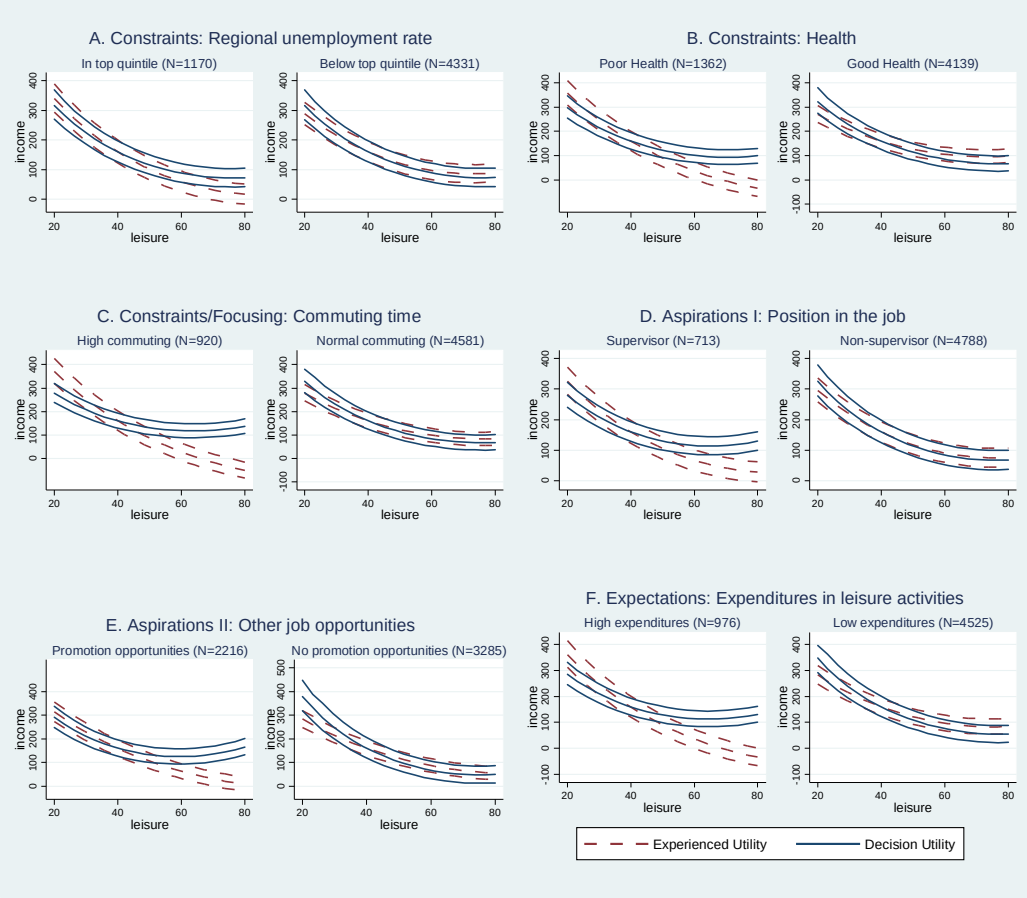
Projection Bias and Focusing Illusion. Commuting can be seen as a constraint.[16] Yet it is also associated with projection bias. People who choose far-away jobs may not be able to correctly guess well-being implications (see Kimball and Willis, 2006). This is also related to the notion of 'focusing illusion'. People focus on one aspect (income) while ignoring the effect of hedonic adaptations to a certain level of wealth (Di Tella et al, 2010, Kahneman and Thaler, 2006).[17] The consequences of a focusing illusion on work and money may include both overtime and lengthy commutes. This alternative interpretation is also consistent with Figure 4C: high commuters work more than low commuters (their choice-revealed IC is flatter) and they seem to work 'too much' given SWB-maximization.

Aspirations and Expectations. Expectations and aspirations can play a critical role on SWB (e.g. Stutzer, 2004). For instance, overly-high aspirations and wrong expectations about labor market outcomes may induce lower (experienced) utility for a time-consuming job. We start

[16] In Frey and Stutzer (2008), it is negatively correlated with SWB, even after controlling for the endogenous sorting of individuals into location choice.

[17] Kahneman et al. (2006) state that "despite the weak relation between income and global life satisfaction or experienced happiness, many people are highly motivated to increase their income. In some cases, this focusing illusion may lead to a misallocation of time".

Figure 4: Constraints, Focusing Illusion and Aspirations



Note: Indifference Curves are obtained using estimates of concentrated life satisfaction (experienced utility) and labor supply (decision utility) on income and leisure. We use a quadratic specification with preference heterogeneity (and other additively separable controls in the SWB equation) as specified in Figure 1. Graphs are obtained by averaging all individual ICs (using either SWB or Utility) drawn through a common point, defined as $(\overline{y}(40), 40)$.

with two variables, the position in the job and the opportunity for promotion, that might give an indication regarding aspirations.[18] Figures 4D and E show some large discrepancies across approaches for people with potentially high career-related aspirations (supervisors and people with high promotion opportunities). Not only do they work more than others (non-supervisor jobs and low opportunities), but they also tend to work more than what SWB would predict. Next, we proxy the level of expectations using information on expenditure on leisure activities. We assume that individuals who spend much for their leisure expect more satisfaction from free time. An alternative interpretation is that they are deprived of leisure time compared to what they would optimally choose and catch up with the quality of leisure. In Figure 4F, ICs elicited from labor supply versus SWB diverge for these people but are very similar for others.

Detailed Distributions of Projection Errors. Finally, Figure 5 shows the distribution of 'projection errors', as previously defined, for all the sub-groups (for the sake of exposition, we standardize projection errors using sample mean error and standard deviation). It is consistent with the IC characterization above. We observe that "unreasonable" sub-groups (on the left) commit larger absolute 'projection errors' while the distribution for others (on the right) is rather centered around zero (and closer to a standard normal distribution). This divergence is rather systematic. For instance, high commuters seem to work too much on the basis of IC comparison, which indeed translates here into a high concentration of negative errors (not enough leisure for SWB maximization).

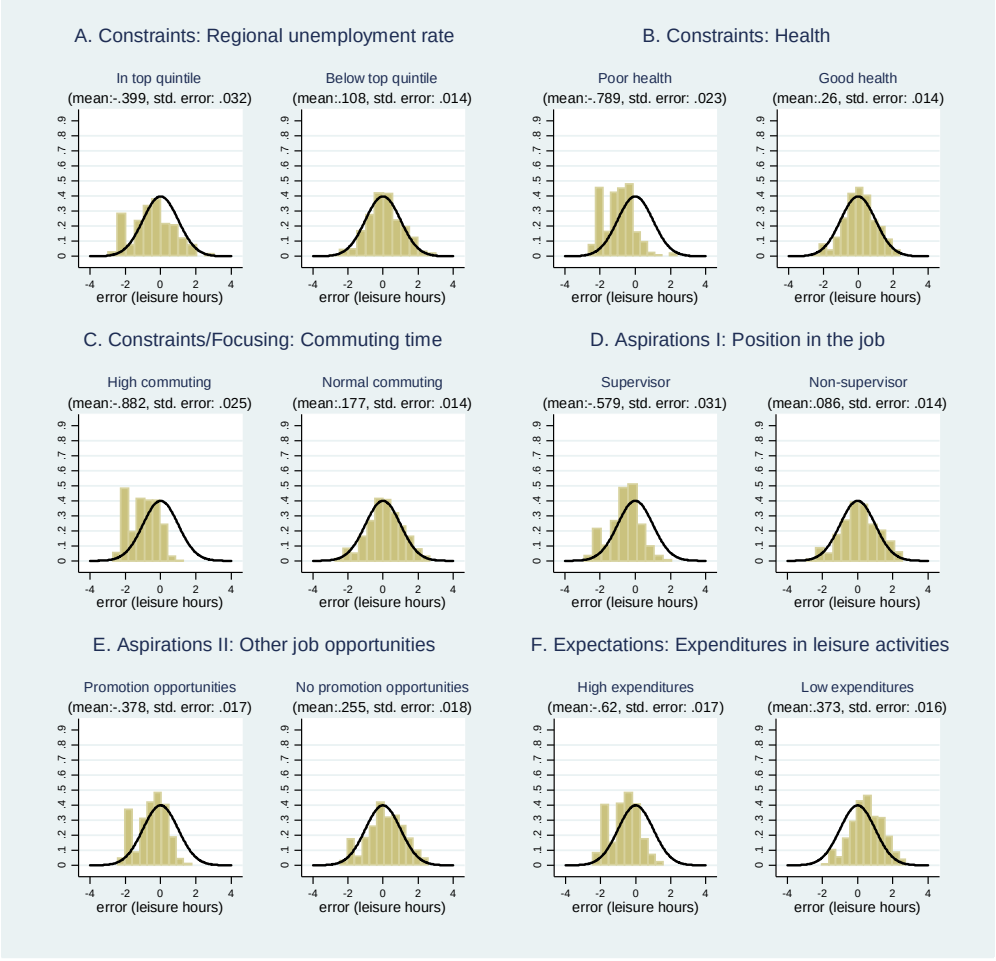
4 Concluding Discussion

This paper originally compares decision and experienced utility using a large household survey. We focus on labor supply, as a crucial dimension for the welfare analysis of redistributive policies. Estimates from both labor supply and SWB models are used to illustrate the tradeoff between leisure and net income. Indifference curves elicited from both approaches are consistent with economic theory and are remarkably comparable on average. A short majority of labor supply decisions in our sample are consistent with SWB-maximization. We have characterized cases where the two measures diverge, providing suggestive evidence for different factors affecting the relationship between the two utility concepts (constraints, projection bias/focusing illusion and aspirations).

Our results convey that after treatment, SWB data can generate consistent preference structures similar to revealed preferences, albeit with a different informational content. From there,

[18] These aspects also relate to status and relative concerns (see the influential study of Luttmer, 2015). It is not clear whether status is implicitly accounted for by the respondents in reported satisfaction about both income and leisure (and, hence, can be seen as part of subjective preferences as we measure them).

Figure 5: Distribution of Standardized 'Projection Errors' by Sub-groups



Note: 'projection errors' are calculated as the difference between actual and SWB-maximizing leisure choice. They are standardized using sample mean 'error' and standard deviation. We report mean and standard error of the standardized error as well as the number of individuals in each sub-group.

it seems restrictive to solely rely on either SWB or revealed preferences to perform welfare analysis of real-world policies. *First*, even if SWB provides a reasonable alternative, as a constraint-free proxy of true utility or in non-priced situations, it remains to be seen if it leads to similar conclusions when departing from comparisons at the mean. Indeed, we have already shown that at a disaggregated level, decision utility does not align with experienced utility for some specific groups in the population – groups whose characteristics matter for the distribution of market wages, among other things. A few studies have established welfare rankings using normative approaches that respect preference heterogeneity, either using revealed preferences (Decoster and Haan, 2014) or SWB (Decancq et al., 2015). Further research should check whether distributional analyses are very sensitive to using SWB rather than choices for welfare measurement based on (the respect of) heterogenous preferences. *Second*, the pursuit of well-being may cover only part of the dimensions guiding humans' actions (SWB can actually be seen as a sub-component of a more general objective function, see Glaeser et al., 2015). Innovative ways to combine both types of information remain to be found, as well as a suitable encompassing framework. An interesting path to follow is the normative application of behavioral economics that allows defining incomplete preferences (Fleurbaey and Schokkaert, 2013).

We see several extensions or improvements to the present work. Our comparison of decision and experienced utility in the context of non-experimental data could easily be extended to other areas in economics like transportation choices or savings (for consumption decisions, see Perez-Truglia, 2015). Also, our models are static and do not consider the dynamic nature of repeated occurrences of decision and experience. Modeling intertemporal decisions would require additional information, including actual consumption at each period (e.g., see Haan et al., 2008). Finally, our SWB estimations are not totally immune of unobserved variables that could affect work preferences /constraints (ex: moral obligation to bring money to the household, or to stay at home to care for a sick person) and, at the same time, SWB levels (the depressing effect of such a situation).[19] To solve this, further work could rely on richer data combining our approach with stated preferences.

References

[1] Adam, S. & J. Browne (2010) "Redistribution, work incentives and thirty years of UK tax and benefit reform", IFS Working paper 10/24

[2] Bargain, O., Caliendo, M., Haan, P. & Orsini, K. (2010). "'Making work pay' in a rationed

[19] As a matter of fact, Fleurbaey and Schwandt (2015) show that 'family obligations' is one of the main reason for not maximizing SWB.

labour market," Journal of Population Economics, 23(1): 323-351

[3] Benjamin, D., O. Heffetz, M. Kimball & A. Rees-Jones (2012). "What Do You Think Would Make You Happier? What Do You Think You Would Choose?" American Economic Review, 102(5), 2083–2110

[4] Benjamin, D., O. Heffetz, M. Kimball & A. Rees-Jones (2014a). "Can Marginal Rates of Substitution Be Inferred from Happiness Data? Evidence from Residency Choices," American Economic Review 104, 3498–3528.

[5] Benjamin, D., O. Heffetz, M. Kimball & N. Szembrot (2014b). "Beyond Happiness and Satisfaction: Toward Well-Being Indices Based on Stated Preference." American Economic Review, 104(9), 2698–2735

[6] Blundell, R., A. Duncan, J. McCrae, & C. Meghir (2000). "The labour market impact of the working families' tax credit", Fiscal Studies, 21 (1), 75-103

[7] Blundell, R., Duncan, A. & Meghir, C. (1998). "Estimating labor supply responses using tax reforms', Econometrica 66(4), 827–861

[8] Blundell, R. & Shephard, A. (2013). "Employment, Hours of Work and the Optimal Taxation of Low Income Families, Review of Economic Studies, 79(2), 481-510

[9] Boyce, C.J. (2010). "Understanding Fixed Effects in Human Wellbeing", Journal of Economic Psychology, 31: 1-16.

[10] Chiappori, P.A. & O. Donni (2011). "Non-Unitary Models of Household Behavior: A Survey of the Literature, Household Economic Behaviors, Molina, Alberto, Berlin: Springer.

[11] Clark, A. & A. Oswald (1994). "Unhappiness and unemployment", Economic Journal, 104, 648-659

[12] Clark, A. E., Frijters, P. & Shields, M. (2008). "Relative income, happiness and utility: an explanation for the Easterlin paradox and other puzzles", Journal of Economic Literature, 46(1), 95-144

[13] Clark, A., C. Senik, & K. Yamada (2015). "When Experienced and Decision Utility Concur: The Case of Income Comparisons", IZA DP No. 9189

[14] Decancq, K., M. Fleurbaey & E. Schokkaert (2015). "Happiness, equivalent incomes, and respect for individual preferences", forthcoming in Economica

[15] Decoster, A. & P. Haan (2015): "Empirical welfare analysis with preference heterogeneity," International Tax and Public Finance, 22 (2), 224-251

[16] Di Tella, R., Haisken-De New, J. & MacCulloch, R. (2010). "Happiness adaptation to income and to status in an individual panel," Journal of Economic Behavior and Organization, 76(3), 834-852

[17] Dolan, P. & Kahneman, D. (2008). "Interpretations of Utility and Their Implications for the Valuation of Health", Economic Journal, 118 (1), 215–234

[18] Ferrer-i-Carbonell, A. & P. Frijters (2004). "How Important is Methodology for the Estimates of the Determinants of Happiness?", Economic Journal, 114(497), 641-659

[19] Ferrer-i-Carbonell, A., B.M.S. van Praag & I. Theodossiou (2014). "Vignette Equivalence and Response Consistency: The Case of Job Satisfaction," IZA DP 6174

[20] Fleurbaey, M. (2008). Fairness, Responsibility and Welfare, Oxford University Press.

[21] Fleurbaey, M., & Blanchet, D. (2013). "Measuring Welfare and Assessing Sustainability", Oxford University Press, USA

[22] Fleurbaey, M. and E. Schokkaert (2013): "Behavioral welfare economics and redistribution", American Economic Journal: Microeconomics, 5(3):180-205, 201

[23] Fleurbaey, M. & H. Schwandt (2015): "Do People Seek to Maximize Their Subjective Well-Being?", IZA DP No. 9450

[24] Frijters, P. (2000). "Do individuals try to maximise satisfaction with life as a whole" Journal of Economic Psychology, 21, 281-304.

[25] Frijters, P., Greenwell, H., Shields, M.A., & J.P. Haisken-DeNew (2009). "How rational were expectations in East Germany after the falling of the wall?", Canadian Journal of Economics, 42(4), 1326-1346

[26] Glaeser, E.L., J.D, Gottlieb & O. Ziv (2015). "Unhappy cities", forthcoming in the Journal of Labor Economics

[27] Haan, P., V.L. Prowse & A. Uhlendorff (2008). "Employment Effects of Welfare Reforms: Evidence from a Dynamic Structural Life-Cycle Model", IZA DP 3480

[28] Hausmann, D.M. (2010). "Hedonism and Welfare Economics", Economics and Philosophy, 26(3): 321–344

[29] Hoynes, H.W. (1996). "Welfare Transfers in Two-Parent Families: Labor Supply and Welfare Participation under AFDC-UP," Econometrica, 64(2), 295-332

[30] Kahneman, D., P. Wakker, & R. Sarin (1997). "Back to Bentham? Explorations of Experienced Utility", Quarterly Journal of Economics, 112(2): 375-405

[31] Kahneman, D. & R. Sugden. (2005). "Experienced Utility as a Standard of Policy Evaluation", Environmental and Resource Economics, 32(1), 161-181

[32] Kahneman, D., & Thaler. R. (2006). "Utility Maximization and Experienced Utility", Journal of Economic Perspectives, 20(1): 221-234

[33] Kahneman, D., & Krueger, A.B., (2006). "Developments in the Measurement of Subjective Well-Being", Journal of Economic Perspectives, 22, 3–24

[34] Kahneman D., A. Krueger, D.A. Schkade, N. Schwarz, A. Stone (2006). "Would You Be Happier If You Were Richer? A Focusing Illusion", Science, 312: 1908-1910

[35] Knabe, A. & S. Rätzel (2010). "Income, happiness, and the disutility of labour", Economics Letters, 107(1), 77-79

[36] Köszegi, B. and M. Rabin (2008). "Choices, situations, and happiness", Journal of Public Economics 92 (8-9):1821-1832

[37] Krueger A.B. & Schkade, D.A. (2008). "The reliability of subjective well-being measures", Journal of Public Economics, 92: 1833–1845

[38] Loewenstein, G., & D. Adler (1995). "A Bias in the Prediction of Tastes," Economic Journal, 929–37

[39] Loewenstein, G., O'Donoghue, T., Rabin, M., (2003). "Projection bias in predicting future utility", Quarterly Journal of Economics, 118; 1209–1248

[40] Luttmer, E.F.P. (2015). "Neighbors as Negatives: Relative Earnings and Well-Being," Quarterly Journal of Economics, 120(3), 963-1002

[41] Odermatt, R. and A. Stutzer (2015). "(Mis-)Predicted Subjective Well-Being Following Life Events", IZA DP No. 9252

[42] Oswald, A. & S. Wu (2010). "Objective Confirmation of Subjective Measures of Human Well-being: Evidence from the USA", Science, 327: 576-579

[43] Perez-Truglia, R. (2015). "A Samuelsonian Validation Test for Happiness Data," Journal of Economic Psychology, 49, 74–83

[44] Ravallion, M., & Lokshin, M., (2001). "Identifying Welfare Effect from Subjective Questions", Economica, 68: 335-357

[45] Ravallion, M. (2015). "Poor, or Just Feeling Poor? On Using Subjective Data in Measuring Poverty", in Happiness and Economic Growth: Lessons from Developing Countries, Oxford University Press

[46] Senik, C. (2005). "Income distribution and well-being: what can we learn from subjective data?", Journal of Economic Surveys 19: 43–63

[47] Stutzer, A. (2004). "The Role of Income Aspirations in Individual Happiness", Journal of Economic Behavior and Organization, 54(1): 89-109

[48] van der Rijt, J.-W. (2013). "Public policy and the conditional value of happiness", Economics and Philosophy, 29 (03), 381-408

[49] van Praag, B.M.S., P. Frijters & A. Ferrer-i-Carbonell (2003). "The Anatomy of Subjective Well-Being", Journal of Economic Behavior and Organization, 51, 29-49

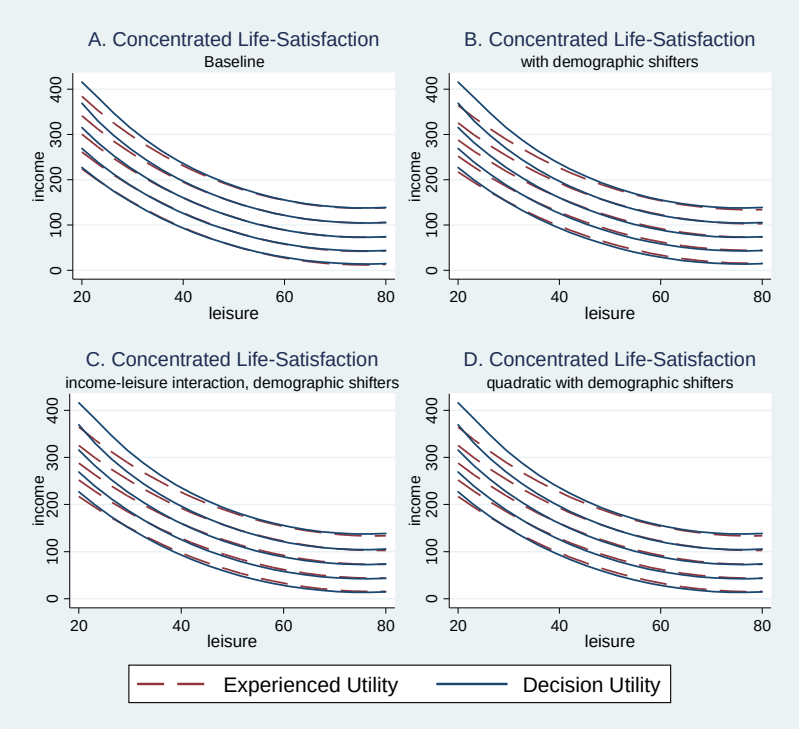
[50] van Soest, A. (1995). "Structural Models of Family Labor Supply: A Discrete Choice Approach", The Journal of Human Resources, 30 (1), 63-88

[51] Wichert, L. & W. Pohlmeier (2010). "Female Labor Force Participation and the Big Five", ZEW discussion paper 10-003

A Online Appendix A: Sensitivity Analysis

Alternative Measures of Income-Leisure Preferences. Our baseline proxy for experienced utility is a concentrated measure $V_{it}^E = \widehat{\gamma}^y S_{it}^y + \widehat{\gamma}^l S_{it}^l$ of income and leisure satisfactions. Using this measure, our primary comparison of ordinal preferences, depicted in Figure 1 in the main text, is reproduced in Figure A.1 (top left graph). We also examine results with $V_{it}^E = \widehat{\gamma}_{it}^y S_{it}^y + \widehat{\gamma}_{it}^l S_{it}^l$ (top right graph), i.e. when heterogeneity is introduced with $\gamma_{it}^y = \gamma^{y0} + x_{it}'\gamma^{y1}$ and $\gamma_i^l = \gamma^{l0} + x_{it}'\gamma^{l1}$ (the set of demographics x_i is the same as preference shifters in the model). Finally, we make the functional form more flexible, adding a term of interaction between S_{it}^y and S_{it}^l (bottom left graph) and the latter plus quadratic S_{it}^y and S_{it}^l (bottom right graph). We can see that overall comparisons are very similar to the baseline.

Figure A.1: Alternative Measures of Income-Leisure Preferences



Preference heterogeneity. We test the sensitivity of IC comparisons with respect to the specification. First, our results are not fundamentally changed with the way we introduce observed preference heterogeneity x_{it}, i.e. on leisure only, on both income and leisure, or on several terms in the quadratic specification. Our conclusions are also unchanged when adding

unobserved preference heterogeneity u_{it} as follows:

$$\beta_l^m(x_{it}) = \beta_{l0}^m + \beta_{l1}^{m\prime} x_{it} + \sigma_l^m u_{it}^m$$

for $m = D, E$. The term u_{it} is assumed to be normally distributed with mean 0 and variance 1, so the standard error is estimated as a preference parameter. In this case, estimations for both approaches can be carried out by simulated ML. In this specification, u_{it} simply plays the role of a random effect and hence cannot solve potential endogeneity issues. Fixed effects can in principle be estimated using the panel dimension of the data. Yet they are not easily accommodated with the nonlinearity of the labor supply model. Moreover, they cannot account for time-varying unobservables.

Functional form. Next, we check the sensitivity of our results to the functional form. We replace the quadratic specification by log-linear and box-cox utility functions. As explained in the main text, these two parametric forms are popular in the SWB (e.g. Clark et al., 2008) and in the labor supply literature (e.g. Decoster and Haan, 2014) respectively. They also capture some non-linearity in income and leisure. Estimations are conducted using the same preference heterogeneity x_{it}, and the same controls z_{it} for the SWB equation, as in the baseline. Estimates are reported in Table A.1, again showing interesting similarities between the two approaches regarding the signs and significance of the coefficients on income and leisure as well as for some of the heterogeneity terms (like London or conscientiousness). The ICs obtained with log-linear and box-cox models are depicted in Figure A.2. Once again, we observe strikingly similar ICs across approaches.

Sensitivity to Additive Well-being Heterogeneity in SWB Equations. A necessary asymmetry in our comparison is the presence of additive terms z_{it} and α_i in the SWB equation. They aim to purge SWB from idiosyncratic variation in well-being responses. Using the baseline quadratic model, we check whether results are sensitive to the way in which we specify this adjustment. Recall that z_{it} contains binary taste shifters x_{it}, additional individual characteristics (detailed age and age squared, family size, health status, home ownership, region, year) while α_i includes detailed 'big five' personality traits. First, we would like to drop z_{it} and α_i completely. To avoid damaging the identification of preference heterogeneity x_{it}, we keep only these variables x_{it} as additively separable controls in z_{it}. Graph A in Figure A.3 shows that the IC with this restricted SWB specification poorly represents the income-leisure tradeoff (it becomes non-monotonic).[20]

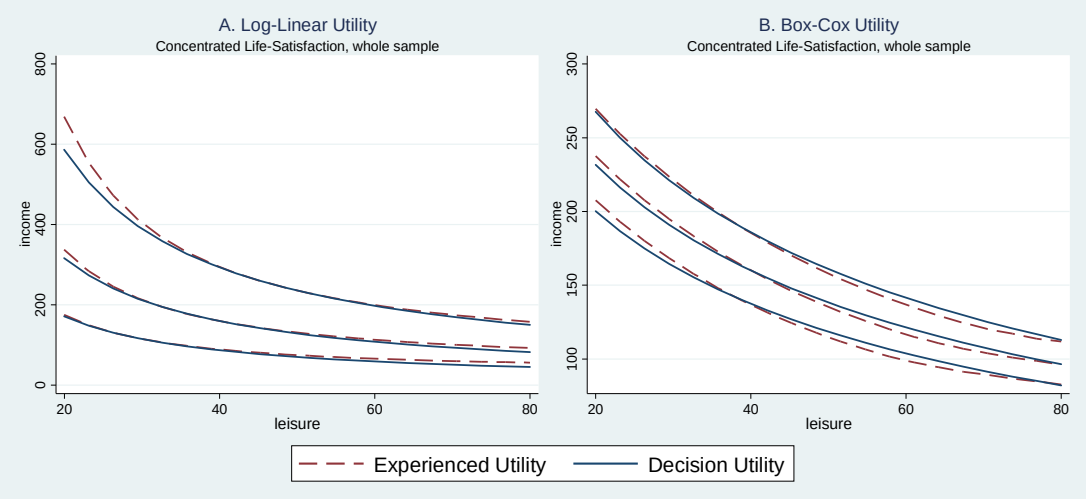
[20] In a stepwise approach, we found that two variables are especially important, namely health indicators and family size. Health is a well-known determinants of SWB. It carries a specific meaning in our comparison, as work constraints may be placed upon the choice of people with bad health (see section 3.4). Family size essentially captures the situation of single mothers (recall that we focus on singles individuals and we have

Table A.1: Preference Estimates from Labor Supply and Subjective Well-Being: Alternative Functional Forms

	Quadratic Baseline		Box-Cox		Log-linear	
Coefficients	Labor Supply	Income-Leisure Concentrated Satisfaction	Labor Supply	Income-Leisure Concentrated Satisfaction	Labor Supply	Income-Leisure Concentrated Satisfaction
Lambda income			0.118***	0.491***		
			(0.0213)	(0.0918)		
Lambda leisure			0.423***	0.618**		
			(0.0143)	(0.314)		
Income^2	-1.87e-05***	-4.82e-07***				
	(8.30e-07)	(9.12e-08)				
Income	0.0282***	0.00120***	3.396***	0.0440***	5.582***	0.105***
	(0.00106)	(0.000135)	(0.229)	(0.0140)	(0.143)	(0.0160)
Leisure^2	-0.00160***	-7.27e-05*				
	(5.69e-05)	(3.88e-05)				
Leisure	0.263***	0.00975**	1.821***	0.0186	5.797***	0.0511
	(0.00761)	(0.00463)	(0.0574)	(0.0227)	(0.165)	(0.0512)
x male	-0.0404***	0.00266	-0.489***	0.0485*	-1.079***	0.117*
	(0.00237)	(0.00170)	(0.0383)	(0.0270)	(0.0988)	(0.0696)
x over 40	-0.00127	0.00100	0.0335	0.0405**	0.115	0.00252
	(0.00207)	(0.000622)	(0.0348)	(0.0176)	(0.0969)	(0.00907)
x high educ.	-0.0216***	6.62e-05	-0.264***	0.0232	-0.627***	0.00632
	(0.00310)	(0.000673)	(0.0440)	(0.0211)	(0.119)	(0.00827)
x young kid	0.0975***	0.00274	1.704***	0.0388	4.380***	0.0843
	(0.00801)	(0.00403)	(0.180)	(0.0700)	(0.487)	(0.247)
x london	0.00860**	0.00699***	0.194***	0.138**	0.496**	0.0567
	(0.00434)	(0.00236)	(0.0721)	(0.0591)	(0.197)	(0.0563)
x non-white	-0.0226***	-0.00738*	-0.341***	-0.117	-0.833**	-0.251
	(0.00758)	(0.00422)	(0.111)	(0.0773)	(0.326)	(0.211)
x migrant	0.00136	-0.00155	0.0348	-0.0266	-0.00107	0.109
	(0.00645)	(0.00372)	(0.107)	(0.0602)	(0.297)	(0.195)
x conscientious	-0.0101***	-0.00148**	-0.166***	-0.0348*	-0.430***	-0.0183**
	(0.00206)	(0.000584)	(0.0349)	(0.0186)	(0.0973)	(0.00845)
x neurotic	0.00322	0.000544	0.0332	0.00231	0.126	0.0118
	(0.00206)	(0.000550)	(0.0351)	(0.0140)	(0.0978)	(0.00804)
Log-likelihood	-12909.25		-9799.90	-5697.71	-9447.16	
Pseudo R^2 / R^2	0.136	0.243	0.136	0.441	0.067	0.234
#Obs	5,501	5,501	5,501	5,501	5,501	5,501

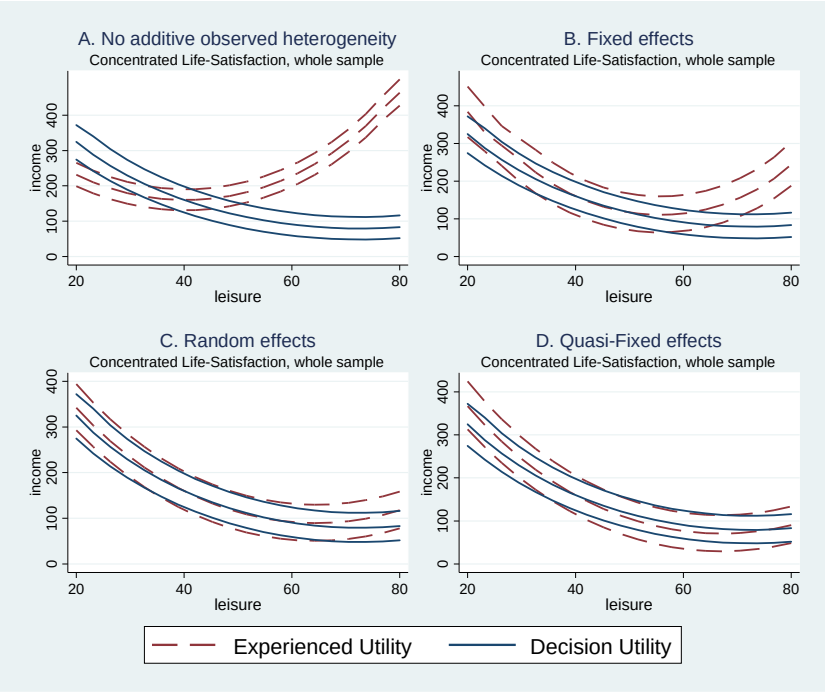
Notes: Subjective well-being equations (concentrated satisfaction) also include additively separable controls (same variables as in leisure interaction terms plus age squared, family size, health status, home ownership, all personality traits, region and year dummies). *, **, *** indicate 1%, 5% and 10% significance levels. Standard errors in parenthesis.

Figure A.2: Comparison of Indifference Curves: Alternative Functional Forms



Note: Indifference Curves are obtained using estimates of concentrated life satisfaction (experienced utility) and labor supply (decision utility) on income and leisure. We use log-linear (A) and box-cox (B) specifications in income and leisure with preference heterogeneity (male, age, education, presence of young kid, London, non-white, migrant, conscientious, neurotic). These variables as well as additional controls (age squared, family size, health status, home ownership, all personality traits, region and year dummies) enter the SWB equation as additively separable controls (hence, not affecting the calculation of ICs). Graphs are obtained by averaging all individual ICs drawn through a common point, defined as $(\overline{y}(40), 40)$.

Figure A.3: Comparison of Indifference Curves: Alternative SWB Models for the Treatment of Additive Heterogeneity



Note: Indifference Curves are obtained using estimates of concentrated life satisfaction (SWB) and labor supply (Utility) on income and leisure. We use a quadratic specification with preference heterogeneity (male, age, education, presence of young kid, London, non-white, migrant, conscientious, neurotic). These variables enter the SWB equation as additively separable controls (hence, not affecting the calculation of ICs) in model A. Only time-varying variables are added in model B, which also include fixed effects. Models C and D include all controls including additional variables as in Figure 1 (age squared, family size, health status, home ownership, all personality traits, region and year dummies) plus random or quasi-fixed effects. Graphs are obtained by averaging all individual ICs (using either SWB or Utility) drawn through a common point, defined as $(\overline{y}(40), 40)$.

Subsequently, we replace the time-invariant terms in z_{it} (like gender) and α_i (big five) by fixed effects (FE), using panel information. The IC from SWB regressions shows a reasonable pattern in Graph B of Figure A.3. Yet we observe some non-monotonicity around zero work hours, which may rationalize the fact that inactivity tends to generate subjective disutility. Importantly, let us recall that the interpretation is different here since only time variation is used to identify the SWB model (while the structural model still relies on both 'within' and 'between' variation).[21] 'Between' variation may attenuate differences (as it captures long-term trends possibly smoothed by adaptation) while "within" variation can be different (in particular, subjective appreciation of transition in or out of work may be stronger for those who experience these changes over the course of the survey).

Graphs C and D in Figure A.3 show the results with usual variables in z_{it} plus α_i modeled as random effects (RE) or quasi-fixed effects (QFE), respectively. Thus, these estimators combine within and between variation in the panel. QFE à la Mundlak are modeled as RE plus the time average of relevant time-varying controls in the estimation (health status, number of children and region). Results are relatively close to the baseline. Estimates of the FE, RE and QFE models are reported in Table A.2, showing usual results with respect to income-leisure coefficients and some of the heterogeneity terms (ex: London).

Involuntary unemployment. We have ignored potential demand-side restrictions in our baseline by excluding job seekers. When we include them in the sample, we obtain non-mononicity in the IC from SWB regressions, as seen in Figure A.4 (graph A). This shows that inactivity is forced for some observations, generating some distress (cf. Clark and Oswald, 1994). We also suggest accounting explicitly for rationing out of the labor market. A simple way is to use a so-called "double-hurdle" labor supply model (see Bargain et al., 2010). It consists of a first-stage estimation of the probability of involuntary unemployment (first hurdle) followed by the labor supply choice (second hurdle). In this case, the IC from actual choices tends to become a little flatter at zero work hours, i.e. a little closer to the IC from SWB (Figure A.4, graph B).

excluded households with multiple family units, e.g. a grown-up adult living with her parents). Single mothers in the UK have low employment rates, mainly explained by low or negative gains from work (cf. Blundell et al., 2000). Nonetheless, they may be dissatisfied not to work. Hence, both health and family size are related to the psychological costs of reduced or no activity. Not accounting explicitly for these variables in the model means that ICs associate maximum leisure (inactivity) with high financial compensation for the psychic cost of staying at home, which explains the non-monotonicity at high levels of leisure.

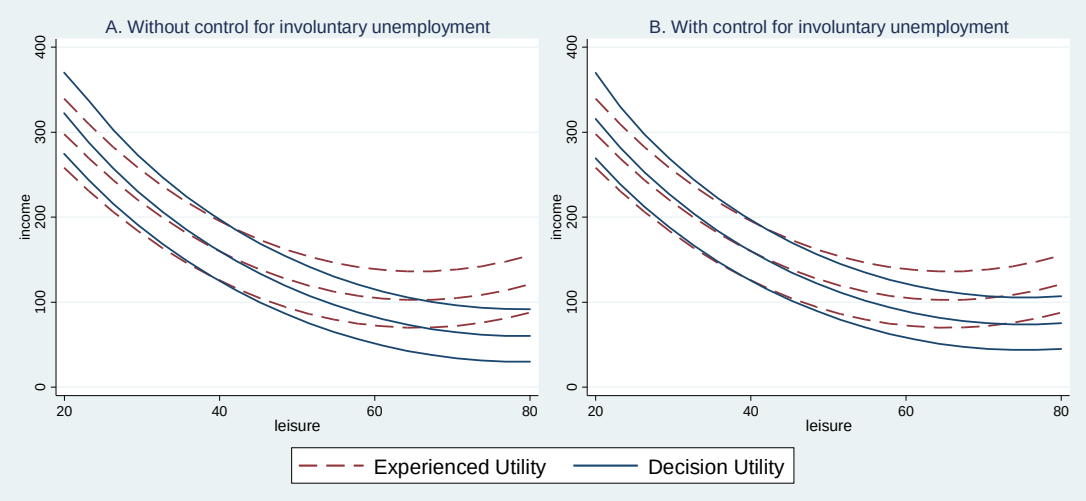
[21] See Fleurbaey and Blanchet (2013) for a discussion of SWB estimations in the context of panel data.

Table A.2: Preference Estimates: Alternative SWB Equations

Coefficients	Labor Supply	Income-Leisure Concentrated Satisfaction			
		OLS (baseline)	Fixed effects	Random effects	Quasi-Fixed Effects
Income^2	-1.87e-05***	-4.82e-07***	-4.52e-07***	-3.90e-07***	-3.83e-07***
	(8.30e-07)	(9.12e-08)	(1.48e-07)	(8.40e-08)	(8.41e-08)
Income	0.0282***	0.00120***	0.000888***	0.00105***	0.00102***
	(0.00106)	(0.000135)	(0.000231)	(0.000140)	(0.000141)
Leisure^2	-0.00160***	-7.27e-05*	-0.000140**	-0.000110***	-0.000109***
	(5.69e-05)	(3.88e-05)	(5.52e-05)	(4.03e-05)	(4.03e-05)
Leisure	0.263***	0.00975**	0.0125*	0.0139***	0.0146***
	(0.00761)	(0.00463)	(0.00719)	(0.00485)	(0.00485)
x male	-0.0404***	0.00266	-0.00113	0.000352	-0.000159
	(0.00237)	(0.00170)	(0.00288)	(0.00183)	(0.00184)
x over 40	-0.00127	0.00100	0.000231	0.000444	0.000445
	(0.00207)	(0.000622)	(0.000791)	(0.000628)	(0.000629)
x high educ.	-0.0216***	6.62e-05	0.00111	0.000491	0.000280
	(0.00310)	(0.000673)	(0.00277)	(0.000916)	(0.000917)
x young kid	0.0975***	0.00274	-0.00180	0.000167	1.17e-05
	(0.00801)	(0.00403)	(0.00394)	(0.00356)	(0.00356)
x london	0.00860**	0.00699***	0.00995**	0.00772***	0.00742***
	(0.00434)	(0.00236)	(0.00422)	(0.00267)	(0.00267)
x non-white	-0.0226***	-0.00738*	-0.00491	-0.00618	-0.00652
	(0.00758)	(0.00422)	(0.00872)	(0.00459)	(0.00459)
x migrant	0.00136	-0.00155	0.0152*	0.000784	0.00104
	(0.00645)	(0.00372)	(0.00801)	(0.00445)	(0.00445)
x conscientious	-0.0101***	-0.00148**	-0.000475	-0.00122	-0.00137*
	(0.00206)	(0.000584)	(0.00264)	(0.000801)	(0.000801)
x neurotic	0.00322	0.000544	0.00594**	0.000224	0.000287
	(0.00206)	(0.000550)	(0.00264)	(0.000756)	(0.000756)
Log-likelihood	-12,909.25				
Pseudo R^2 / R^2	0.136	0.243	0.053	0.237	0.240
#Obs	5,501	5501	5,501	5,501	5,501

Notes: Subjective well-being equations (concentrated satisfaction) also include additively separable controls (same variables as in leisure interaction terms plus age squared, family size, health status, home ownership, all personality traits (except with FE), region and year dummies). *, **, *** indicate 1%, 5% and 10% significance levels. Standard errors in parenthesis.

Figure A.4: Comparison of Indifference Curves when Including Job Seekers



Note: Indifference Curves are obtained using estimates of concentrated life satisfaction (SWB) and labor supply (Utility) on income and leisure. We use a quadratic specification with preference heterogeneity (and other additively separable controls in the SWB equation) as specified in Figure 1. Overall ICs are obtained for mean value of taste shifters and for the group mean level of SWB or Utility. Contrary to Figure 1, we use here a sample including job seekers. Fig 1B explicitly accounts for involuntary unemployment using a double hurdle labor supply model. Graphs are obtained by averaging all individual ICs (using either SWB or Utility) drawn through a common point, defined as $(\overline{y}(40), 40)$.

B Online Appendix B: Marginal Rates of Susbtitution

Table B.1: Marginal Rates of Substitution at 40 Weekly Hours: Demographic Variation

Heterogeneity		Revealed Preferences		Subjective Well-being	
		Coeff.	Std. Err.	Coeff.	Std. Err.
Gender	Female	-5.899	0.125	-4.175	1.622
	Male	-4.084	0.094	-6.715	1.659
Age	Young	-5.262	0.104	-4.691	1.451
	Old	-5.205	0.109	-5.648	1.488
Education	Low	-5.408	0.102	-5.091	1.438
	High	-4.438	0.144	-5.154	1.547
Children	No	-5.062	0.090	-4.997	1.424
	Yes	-9.447	0.432	-7.607	4.153
London	No	-5.210	0.096	-4.629	1.435
	Yes	-5.596	0.204	-11.296	2.659
Ethnicity	White	-5.261	0.096	-5.264	1.446
	Non-white	-4.247	0.342	1.775	4.159
Migrants	Native	-5.235	0.096	-5.147	1.446
	Migrant	-5.297	0.298	-3.665	3.667
Conscientiousness	Low	-5.436	0.108	-5.719	1.476
	High	-4.982	0.103	-4.309	1.451
Neuroticism	Low	-5.166	0.104	-4.848	1.447
	High	-5.311	0.108	-5.366	1.476
Mean		-5.237	0.095	-5.102	1.437

Table B.2: Marginal Rates of Substitution at 40 Weekly Hours: Explanatory Variables

Heterogeneity		Revealed Preferences		Subjective Well-being	
		Coeff.	Std. Err.	Coeff.	Std. Err.
Regional unemployment	High	-5.127	0.135	-5.566	1.829
	Low	-5.265	0.099	-4.961	1.468
Health	Poor	-4.836	0.119	-7.765	1.796
	Good	-5.380	0.103	-3.994	1.463
Commuting	High	-4.065	0.140	-7.347	2.851
	Low	-5.546	0.110	-4.570	1.500
Supervise	Yes	-4.179	0.141	-6.086	3.226
	No	-5.483	0.106	-4.967	1.487
Opportunities	Yes	-4.141	0.109	-6.019	1.919
	No	-6.734	0.168	-4.820	1.726
Leisure expenditure	High	-4.446	0.109	-7.465	1.896
	Low	-5.835	0.126	-4.717	1.708