Back to Bentham: Should We? Large-Scale Comparison of Decision versus Experienced Utility for Income-Leisure Preferences

Alpaslan Akay
University of Gothenburg, IZA

Olivier Bargain
Aix-Marseille (Aix-Marseille School of Economics),
CNRS & EHESS, IZA

H. Xavier Jara
Institute for Social and Economic Research
University of Essex

INSTITUTE FOR SOCIAL
& ECONOMIC RESEARCH

E·S·R·C
ECONOMIC
& SOCIAL
RESEARCH
COUNCIL

Non-technical summary

The standard approach to measure well-being in Economics is based on inferring preferences information from the observed decisions made by supposedly rational agents. At the same time, research on self-reported "happiness" or "life satisfaction" has increasingly been used as a way to measure individual well-being and to derive information about preferences for different life domains.

The aim of this paper is to compare preferences information elicited from subjective wellbeing data to that inferred from actual labour supply choices. More precisely, we focus on income-leisure preferences, which are closely related to redistributive policies. We make use of data from the British Household Panel Survey to estimate both models on the same panel of individuals.

Our results suggest that, on average, there are striking similarities between income-leisure preferences derived from both approaches, reflecting that overall, people's decisions are not inconsistent with subjective wellbeing maximisation. Nevertheless, preferences information differs across approaches when looking at specific population groups. We investigate these differences and explore potential explanatory channels, such as constraints or 'errors' related to aspirations, expectations, past experiences, among others.

Back To Bentham: Should We?

Large-Scale Comparison of Decision versus Experienced Utility for Income-Leisure Preferences*

Alpaslan Akay, Olivier Bargain, H. Xavier Jara

January 2015

Abstract

Subjective well-being (SWB) is increasingly used as a way to measure individual well-being. Interpreted as "experienced utility", it has been compared to "decision utility" using specific experiments (Kahneman et al., 1997) or stated preferences (Benjamin et al. 2012). We suggest here an original large-scale comparison between ordinal preferences elicited from SWB data and those inferred from actual choices (revealed preferences). Precisely, we focus on income-leisure preferences, closely associated to redistributive policies. We compare indifference curves consistent with income-leisure subjective satisfaction with those derived from actual labor supply choices, on the same panel of British households. Results show striking similarities between these measures on average, reflecting that overall, people's decision are not inconsistent with SWB maximization. Yet, the shape of individual preferences differ across approaches when looking at specific subpopulations. We investigate these differences and test for potential explanatory channels, particularly the roles of constraints and of individual "errors" related to aspirations, expectations or focusing illusion. We draw implications of our results for welfare analysis and policy evaluation.

Key Words : decision utility, experienced utility, labor supply, subjective well-being.

JEL Classification : C90, D63

1. Introduction

The standard approach to measure well-being in Economics consists in inferring ordinal preferences from the observation of decisions made by supposedly rational agents (rationality primarily means utility-maximizing behavior). The object derived from this "revealed preference" approach is sometimes referred to as a *decision utility*. For instance, in the context of labor supply, indifference curves in the income-leisure space can be retrieved from variation in labor supply decisions following variation in the wage rate.[1] At the same time, research based on *experienced utility*, as proxied by self-reported information on subjective experience of "happiness" or "life satisfaction", has recently emerged in Economics. It draws from decades of research in psychology and other social sciences.[2] Further back in time, the search for welfare measures associated with pleasure and pain following an experience was originally suggested by utilitarians such as Jeremy Bentham (see for instance Kahneman and Sugden, 2005).

In this paper, we aim to compare these two types of welfare measures. The motivation is manifold. *First*, it is worrying to observed a growing divide between the two approaches, which we rather see as complementary. The relatively a-theoretical approach based on "subjective well-being" (SWB therein) to measure human welfare has brought a lot more realism into the utility function and covers dimensions which could not be addressed with revealed preferences.[3] On the other hand, the traditional approach relies on behavioral models that support policy analyses and policy design.[4] Hence, there is a pragmatic aim at reconciling these approaches. *Second*, there is a more profound need to compare these welfare measures, in particular to understand to which extent they are related. SWB measures do not only cover a multitude of domains that determine welfare but also reflect constraints imposed on human pursuit of well-being or, broadly speaking, optimization errors that may take us away from happiness maximization (see the stimulating discussion in Kimball and Willis,

---

[1] Often, in absence of wage variation, it is possible to assume homogenous preferences within a certain gender-age-education group, say, and to retrieve ordinal preferences by using cross-sectional variation in both wages and labor supply choices in this group. After choosing a particular specification of the utility function, behavioral parameters can be recovered by standard estimation methods (see for instance Blundell and MaCurdy, 1999). In the present paper, we use panel data and, hence, exploit both time and group variation in wage/labor supply choices.

2 See surveys in Clark et al. (2008) or Layard (2005), among others, and critical reviews in Kahneman and Krueger (2006), Fleurbaey (2009), Fleurbaey and Blanchet (2013).

[3] For instance, it allows studying the role of adaptations (e.g., Loewenstein and Ubel, 2008), positional concerns (e.g., Senik, 2009) or the welfare impact of circumstances like unemployment (e.g., Clark and Oswald, 1994).

[4] For instance, structural models of labor supply are extensively used to simulate redistributive reforms or to define optimal tax schedules (see Blundell and Shepard, 2013, as an example among many).

2006). On the other hand, actual decisions (and underlying objective functions captured in revealed preferences) may account for life goals that are not part of SWB as we measure it (freedom, pride, etc.). Even so, SWB and actual choices may lead to similar ordinal "preferences" regarding the relationship between key variables like income and leisure. Checking this degree of consistency between the two measures is the driving question behind this study. *Third*, recent research has attempted to operationalize welfare comparisons based on some fairness principles while respecting individual preference heterogeneity (see the general approach in Fleurbaey, 2008). These studies have indifferently used revealed preferences (for instance Decoster and Haan, 2014, or Bargain et al., 2013) or SWB (Decancq et al., 2013) in their empirical applications. We consider the present study as a first step in understanding possible normative implications of using one or the other approach for "fair" welfare analysis.

Admittedly, studies using experimental data in psychology and behavioral economics have helped to explain the differences between experienced and decision utility, and in particular to understand why and how people make decisions which do not maximize happiness. Specific comparisons of these two welfare measures have been produced in the field of public good valuation (SWB versus contingent valuation, e.g., Kahneman and Sugden, 2005), in small-scale experiments (see the survey of Kahneman et al., 1997) or using stated preferences (asking people if they would choose what they think would maximize their SWB, see Benjamin et al. 2012, 2014).[5] Yet, to our knowledge, there is hardly any research trying to (i) compare experienced utility (elicited from SWB) with decision utility inferred *from actual choices* on the same households, and (ii) to do so in *large-scale* household surveys, especially when valuing income against leisure time. We believe the domain of "income-leisure" preferences is particularly important due to its intimate connection with social justice theories and normative questions.[6]

We proceed with a comparison of ordinal preferences elicited from people's self-reported SWB or revealed from the same people's labor supply choices, using a large nationally representative dataset of British households. That is, we recover indifference curves (IC therein) from SWB regressions on one side, and from estimation of a random-utility model of labor supply on the other. Preference heterogeneity is introduced with usual taste shifters (e.g., male and female, high and

---

[5] A broad field of behavioral economics also focuses on labor market outcomes, for instance the cab driver studies starting with Camerer et al. (1997).

[6] The way we measure welfare and assess preference heterogeneity across individuals may have strong implications when designing an optimal redistribution system or using normative characterization that hold people responsible for their preferences (for instance, when asking: "should we hold the poor responsible for being so if poverty is not only on account of low productivity but also different valuation of leisure?").

low educated, old and young, migrant and native), psychological traits (we extract two relevant variables among the "big five" that likely affect work choices: being neurotic and conscientious),[7] and variables that may proxy "optimization errors" (ex: information related to personal aspirations) or labor market constraints.

Results first show that empirical ICs from decision utility versus experienced utility are broadly consistent with economic theory: Both approaches produce ICs which are downward sloping, reflecting income-leisure trade-offs in a textbook fashion. The most interesting result is that a sufficiently rich SWB model specification generates strikingly similar ICs to those obtained from the labor supply model, at least on average. This strong empirical regularity across approaches conveys that there is – on average – no systematic "error" in this domain, i.e. people choose working time as if they were maximizing SWB. Our conclusions are broadly unaffected when changing model specifications (functional forms, treatment of unobserved individual effects) and ICs generated by alternative SWB measures. Admittedly, ICs diverge for specific groups, which may reflect specific behaviors (for instance, "projection bias"), constraints in the optimization process (labor market constraints) or other possible channels through which sub-optimal behavior may appear (for instance expectations, aspirations and the "focusing illusion"). We suggest an extensive analysis of these factors.

The paper is structured as follows. In Section 2, we present the data, selection and the empirical approach to elicit ordinal preferences. Section 3 presents our main empirical results, estimates and patterns of indifference curves with both approaches, while Section 4 explores the cases when differences between approaches occur. Section 5 concludes by drawing implications for welfare measurement and policy analysis.

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 classic information about individual and household characteristics to be used in our estimations (gender, age, education, labor market status).

---

7 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.

We select a sample for the years 1996-2005.[8] We restrict our analysis to single individuals, which enables us to neglect interactions within the household in the context of joint work decisions. As it is usually done in the labor supply literature, we further exclude individuals in self-employment because their labor supply decisions may differ considerably from those of salaried workers and because income information from surveys is considered less reliable in their case. Disabled individuals, full-time students and pensioners are also excluded in order to keep only those individuals available for the labor market. We take out all non-workers that are classified as 'involuntary unemployed' (actively looking for a job but, possibly temporarily, rationed out of the labor market), to comply with the labor supply nature of the model and discard possible interpretations in terms of demand-side constraints.[9] Finally we keep only individuals for whom all key characteristics (including socio-demographics and personality traits) are available all years. The final sample includes 1,881 individuals, and 5,501 person x year observations, for our empirical analysis.

The key variables for our analysis are working time and disposable income. We make use of variables on weekly working hours combined with employment status to allocate each worker in a discrete category, as explained below. Household disposable income (income after tax payment and benefit receipt) is microsimulated using tax-benefit rules of each year and, for each household, information on earnings and demographics that affect the payment of taxes or the receipt of transfers. In addition, we extract information on many other individual characteristics which have been used in the literature as taste shifters in structural labor supply models and/or determinants of SWB, including gender and age; single, widowed or divorced; health status (very good health to very poor health); educational level (elementary school, high school or university); native or immigrant; ethnicity (simplified to white or non-white origin); number of household members (mainly children or elderly dependents) and a dummy for the presence of young children (aged 0 to 2); living in London; and personality traits.

An important question in our study is which measure of well-being should be used to compare decision and experienced utility in the income-leisure domain. For our baseline, we shall make use of the answer to the *life satisfaction* question: "How dissatisfied or satisfied are you with your life overall?" The well-being is measured in an ordered scale between 1 and 7, where 1 means "not satisfied at all" and 7 means "completely satisfied". We shall also suggest a way to extract satisfaction in the

---

[8] Important variables are missing for years 2006-7 and we avoid the onset of the crisis (2008) for consistency with a pure labor supply interpretation.

[9] 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?".

domain of income-leisure preferences from the general SWB index. To do so, we will rely on satisfaction with income and free time 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?". As for the life satisfaction question, answers are ordered from 1 to 7 ("not satisfied at all" to "completely satisfied"). We shall also check the sensitivity of our results when using two other SWB measures, including the *General Health Questionnaire* (GHQ-12)[10] and answers to the *happiness* question.[11] We can already indicate that all three measures are highly correlated.[12]

2.2. Labor Supply Model and Estimations of "Decision Preferences"

We first elicit "income-leisure" revealed preferences from actual labor supply choices. The behavioral model used here is a classic example, in contemporary economics, of a model conceptually based on "decision utility" (e.g., Blundell et al, 2000). Agents are assumed to choose among J discrete work choices. Let y_{ijt} be the weekly disposable income of individual *i* and h_{ijt} her weekly working hours for choice $j = 1,...,J$ and period *t*. The utility level derived from this choice is written:[13]

$$V_{ijt}^{D} = U_i^{D}(y_{ijt}, l_{ijt}) + \varepsilon_{ijt} \tag{1}$$

with $U_i^{D}()$ the utility function determined by income (equivalent to consumption in this static model) and $l_{ijt} = T - h_{ijt}$ her "leisure" time (i.e., total time *T* minus formal hours of work). Non-market time is associated to leisure while in principle we could remain agnostic about its very nature, which may well incorporate other uses like time dedicated to domestic production. Disposable income y_{ijt} for choice *j* and period *t* is written:

$$y_{ijt} = G_t(w_{it}, h_{ijt}, \mu_{it}, \zeta_{it}) \tag{2}$$

where w_{it} is her gross hourly wage rate, μ_{it} her non-labor income and the function $G_t()$ represents the function transforming gross income and individual characteristics into disposable income. It is numerically simulated on the basis of tax-benefit rules at period *t*, and depends on non-labor income and weekly labor income $w_i h_{ij}$ (but also on wage w_i and work duration h_{ij} independently because of tax rules - for instance an

---

10 GHQs are originally designed to measure mental health (Clark and Oswald, 1994; van Praag et al., 2003; Clark et al., 2008). To build the GHQ-12 index, each answer is recoded so that answers are scaled from 0 to 3 and summed to give an overall scale from 0 (lowest mental health level) to 36 (highest level).

11 "Have you recently been feeling reasonably happy, all things considered?". There are four answers to the question which are recoded as: 4- more than usual, 3- same as usual, 2- less than usual and 1-much less than usual.

12 Linear correlation between life satisfaction and GHQ-12 (resp. happiness) is 0.587 (resp. 0.462).

13 In our application, J=7 and corresponds to weekly work hours from 0 to 60, with a step of 10 hours.

eligibility condition of the British Working Tax Credit on working a minimum of 16 hours per week). This function also depends on a set ζ_{it} of individual characteristics (for instance, the presence of children, which conditions the calculation of taxes and transfers - like child benefits or increment of income support due to the presence of dependent children). A standard Heckman-correction model is used to estimate wage equation. Wage rates are then predicted for all workers.[14]

The model is derived under the assumption of (decision) utility maximization, with individuals revealing their preferences by choosing among the set of discrete hours alternatives. Revealed preferences usually require variation in price, typically time variation in commodity price. In the present context, we benefit from time and cross-sectional variation in the *net* wage rate, which allows identifying preferences under parametric assumptions (for a non-parametric approach of revealed preference in the context of labor supply, see Cherchye and Vermeulen, 2008). Let us be more specific on the important question of model identification. A standard critique about identification of behavioral parameters based on cross-sectional variation (i.e. between individuals) is that *gross* wages are potentially endogenous to preferences (for instance workaholic agents were also hard-working during their studies and are likely to have higher wage rates today). Identification based on instrumentation of wage rates hinges on the quality of instruments, which is always questionable; identification on nonlinearity and discontinuity in tax-benefit rules requires extrapolation across demographic groups and may not circumvent the problem. Therefore, an additional source of exogenous variation in *net* wages can be obtained by spatial variation in tax-benefit rules (as in Hoynes, 1996, who use variation in tax schedules across US states) and time variation in these rules (as Blundell et al., 1998, who rely on tax-benefit reforms over the years). We follow this identification strategy by relying on some regional variation (e.g., council tax varies between the four main regions) and especially on 10 years of data during which the British tax-benefit system has experienced important changes.[15]

Note also that this approach requires only the comparison of utility at discrete choices (rather than usual tangency conditions), so that preferences can in principle remain very general. For comparability with the SWB estimations, however, we shall use simple and interpretable specifications. Our baseline shall rely on a quadratic form in disposable income and "leisure" for the deterministic part of the utility function:

---

[14] Wages being calculated as earnings divided by working time, they may be contaminated by the same measurement error as the dependent variable (working time), creating a so-called "division bias". Predicting wages for all (and not only for non-workers) tend to reduce this bias (see Bargain et al., 2013).

[15] Notably important changes in the income tax schedule, social insurance contributions, council taxes, income support generosity and tax credits for working poor families under "New Labour". See Blundell et al. (2000) and Adam and Browne (2010).

$$U_{ijt}^D = \alpha_{yy} y_{ijt}^2 + \alpha_{ll} l_{ijt}^2 + \alpha_y y_{ijt} + \alpha_l (x_{it}) l_{ijt} \tag{3}$$

The labor supply model is kept as simple as possible in order to ensure a clear interpretation of the preference parameters. In particular, the model does not include fixed costs of work because they are usually not identified (see Bargain et al., 2013), or only at the price of parametric assumptions which could not be carried on for the SWB estimation and, hence, would make our comparison fail.[16] An important point for our analysis is the source of heterogeneity across individuals. Preference heterogeneity is accounted for through the leisure term:

$$\alpha_l (x_{it}) = \alpha_{l0} + \alpha_l' x_{it} \tag{4}$$

which varies linearly with a vector x_{it} of binary taste shifters, including dummies for males, 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. These factors will be used to check if both revealed and SWB-consistent preferences indicate the same trends within each groups (for instance, that conscientious workers work more hours and get more SWB than non-conscientious if they actually do so).

In total utility $V_{ijt}^D = U_{ijt}^D + \varepsilon_{ijt}$, the random component ε_{ijt} is assumed to be i.i.d. and following an extreme value type I (EV-I) distribution, such that the probability, P_{ikt}, for each individual i to choose alternative k at time t, has an explicit conditional logit form:

$$P_{ikt} = \Pr(V_{ikt}^D \geq V_{ijt}^D, \forall j = 0,...,J) = \frac{\exp U_{ikt}^D}{\sum_{j=0}^{J} \exp U_{ijt}^D}. \tag{5}$$

The preference parameters are easily obtained by means of maximum likelihood techniques. Once the model is estimated, indifference curves are calculated using the model parameters by inverting expression (3) to retrieve income as a function of leisure for a fixed level of utility.

2.3. Subjective Well-Being and Estimation of "Experienced Preferences"

In the growing literature using "happiness" or "life satisfaction" data, SWB estimations are most often based on simple linear models controlling for observed and

---

[16] Also, fixed costs are difficult to interpret and can be seen as part of the individual's preferences (psychic cost of working) or part of constraints imposed by the demand side (negative work costs would correspond for instance to psychic cost of remaining unemployed, see Clark and Oswald, 1994). We ruled out potential demand-side restrictions on the labor market by taking out rationed job seekers. We have also experimented estimation including these group and dealing explicitly with rationing thanks to a double-hurdle labor supply model (see Bargain et al., 2010). Results, available from the authors, were not very different, possibly due to the small number of rationed workers on the British labor market of the years under study.

unobserved determinants of well-being. Hence, the experienced utility derived from income and leisure, for individual i making choice j at time t, is written as:

$$V_{ijt}^E = U_{ijt}^E + \lambda' z_{it} + \tau_t + \theta_i + \eta_{ijt}$$

$$\text{with} \quad U_{ijt}^E = \beta_{yy} y_{ijt}^2 + \beta_{ll} l_{ijt}^2 + \beta_y y_{ijt} + \beta_l(x_{it})l$$

(6)

We specify the income-leisure function as in the deterministic part of the labor supply model and, still for comparability purposes, introduce observed heterogeneity in preferences as follows:

$$\beta_l(x_{it}) = \beta_{l0} + \beta_l' x_{it}.$$

(7)

The vector x_i of taste shifters contains the same variables as the labor supply model above (male, age above 40, etc.). One important point to highlight is that SWB measures are the result of various occurrences in one's life (enjoying a certain level of income and leisure, but also having a house, being married or not, etc.). These subjective measures involve also high noise and measurement issues. Therefore in order isolate the effect of income, leisure and the taste shifters appearing in the leisure interactions, we proceed with two adjustments. *First*, as explained below, we construct a SWB measure which concentrates the experience of income-leisure choices by combining information from life satisfaction in general with satisfaction in the income and leisure domains. *Second*, we control for a rich set of characteristics that may shift individual well-being levels, some of which are well-known correlates or determinants of SWB in the literature (see Clark et al., 2008). In equation (6), this includes observable characteristics (age, age squared, education, marital status, male, presence of children aged 0-2, ethnicity, migrant, home ownership, health status, year and region), represented by vector z_{it}, time effects τ_t and a series of unobserved effects θ_i (η_{it} is a usual i.i.d. error term). Standard practice in the SWB literature consists in modeling individual-specific influences θ_i.[17] In our baseline, we do so using the big-5 personality traits, which usually account for an important part of the variation in SWB (Boyce, 2010, Ravallion and Lokshin, 2001), then by alternatively modeling θ_i as individual effects using panel information.[18]

---

[17] These individual effects lift some of the concerns related to interpersonal comparability (Kahneman and Sugden, 2005). Our problem is slightly different, however, as we construct ordinal preference measures. It is rather related to the fact that the relationship between income and leisure revealed by SWB regressions may not be interpreted as preferences if SWB measures are not purged from individual heterogeneity like aspirations (see discussion in Decancq et al. 2013). Individual effects play that role in the present application.

18 Note that life satisfaction takes a discrete value on a 1-7. Yet, results are very similar whether we use linear estimations of (6) or ordered probit/logit models (see also Ferrer-i-Carbonell and Frijters, 2004). We favor the linear approach since it makes the introduction of individual effects much easier. We have nonetheless tried estimations of fixed effect ordered logit (the "Blow and Cluster" model of Baetschmann et al., 2011), which again gives no major difference in signs and significance of the model parameters compared to other specifications.

An ideal SWB measure for our study would be the direct question about satisfaction levels obtained from actual income-leisure decisions. Our dataset does not include such a measure of well-being, yet there are measures available for life domains that can be used for that purpose (these are usually found to be strongly interrelated with overall life satisfaction because of common explanatory variables, see van Praag et al., 2003). Hence, we approximate experienced utility from income-leisure by *concentrating* the overall SWB index ("life satisfaction" in our baseline) using "satisfaction from income" and "satisfaction from leisure" (as previously defined in the data section). Let S_{it} be life satisfaction as reported by individual i at time t, S_{it}^{y} her satisfaction with income and S_{it}^{l} her satisfaction with leisure time. We estimate

$$S_i(S_i^y, S_i^l) = \gamma_y S_{it}^y + \gamma_l S_{it}^l + e_{it}, \tag{8}$$

then use estimated weights on each relevant domains of satisfaction to predict our "income-leisure concentrated" SWB measure:[19]

$$U_{it}^E = \hat{\gamma}_y S_{it}^y + \hat{\gamma}_l S_{it}^l. \tag{9}$$

We consider this proxy for income-leisure (experienced) utility as the most closely related to the (decision) utility conceptualized in the labor supply model. For sensitivity checks, we shall also apply this approach to the "happiness" and "GHQ-12" measures of well-being. Yet, "income-leisure concentrated" happiness or GHQ-12 will be taken with caution since they rely on income/leisure satisfactions (we do not have happiness or mental health measures experienced in the domains of income or leisure).

3. Results

We present estimated parameters for our baseline labor supply model and SWB model (3.1), then compare these methods by means of ICs overall (3.2) and for various population groups (3.3), asking ourselves whether people choose working time in a way that maximize their SWB (3.4).

3.1. Decision and Experienced Utility Functions

We first present estimated parameters for our baseline labor supply model and SWB model. It is important to highlight the fact that parameter estimates are not directly comparable since these two types of welfare measures rely on different implicit scales. Only Marginal Rates of Substitution (MRS) between variables of interest could be compared, which is what we do more generally by comparing ICs in 3.2. At this stage, our aim is simply to compare signs and significance of key variables appearing

---

19 The estimated relative weight on income is $\tilde{\gamma}_y / (\tilde{\gamma}_y + \tilde{\gamma}_l)$ 0.468.

in both models. Table 1 reports coefficients on income and leisure terms only, including interaction terms with leisure for preference heterogeneity.[20] Utility is increasing and concave in income and leisure with both decision and experienced preferences. This is statistically significant for the labor supply model and for the "concentrated" measures of SWB.[21] General life satisfaction also gives a similar pattern, even if leisure terms are insignificant.

Turning to preference heterogeneity, it is assumed to take place across broad groups of characteristics like gender, age, etc.[22] Other things being equal (the net wage in particular), heterogeneous work preferences will lead to different labor supply choices across groups (say, highly conscientious people will work more), which is indeed rationalized by the labor supply model as different (decision) preferences (a lower leisure weight α_l on conscientiousness). The SWB regression may or may not give the same thing (in our example, a lower leisure weight β_l on conscientiousness). Interpretations may depend on the type of characteristic we consider. For instance, while conscientious workers may spend more hours at work, it is intuitive to think they will also derive less (experienced) disutility from doing so. Other variables like age or education are more difficult to interpret as they affect many other dimensions.

---

20 Estimates of controls in equation (6) give the standard results found in the SWB literature: SWB is U-shaped in age, positively correlated with health and education. Complete results are available from the authors.

[21] That the latter give the most similar informational content to revealed preferences is (unsurprisingly) consistent with the fact that they manage to extract well-being variation in the relevant domains of income and leisure.

[22] That is, we must assume that within a cell defined by specific characteristics (for instance, being a young, low-educated male living in London), preferences are homogenous and constant over the period under investigation. For richer specifications with unobserved heterogeneity in preferences, see section 3.2.

Table 1: Comparing parameter estimates from structural models and subjective well-being regressions

Coefficients	Labor Supply	Subjective Well-being					
		Life Satisfaction	Income-Leisure Concentrated Satisfaction	GHQ	Income-Leisure Concentrated GHQ	Happiness	Income-Leisure Concentrated Happiness
$Income^2$	-1.87e-05***	-4.67e-07***	-4.82e-07***	-4.61e-07	-1.63e-06***	2.12e-08	-1.24e-07***
	(8.30e-07)	(1.54e-07)	(9.12e-08)	(7.32e-07)	(2.83e-07)	(8.64e-08)	(2.10e-08)
Income	0.0282***	0.000955***	0.00120***	-0.000662	0.00424***	-9.32e-07	0.000324***
	(0.00106)	(0.000227)	(0.000135)	(0.00108)	(0.000418)	(0.000128)	(3.11e-05)
$Leisure^2$	-0.00160***	-1.48e-05	-7.27e-05*	-6.70e-05	-0.000220*	4.00e-06	-1.62e-05*
	(5.69e-05)	(6.56e-05)	(3.88e-05)	(0.000312)	(0.000121)	(3.68e-05)	(8.96e-06)
Leisure	0.263***	0.00251	0.00975**	-0.00531	0.0272*	-0.00165	0.00196*
	(0.00761)	(0.00781)	(0.00463)	(0.0371)	(0.0144)	(0.00438)	(0.00107)
x male	-0.0404***	-0.00180	0.00266	0.0178	0.00794	0.00402**	0.000584
	(0.00237)	(0.00287)	(0.00170)	(0.0136)	(0.00527)	(0.00161)	(0.000392)
x over 40	-0.00127	0.00133	0.00100	-0.000123	0.00284	-9.21e-05	0.000206
	(0.00207)	(0.00105)	(0.000622)	(0.00499)	(0.00193)	(0.000589)	(0.000143)
x high educ.	-0.0216***	0.00200*	6.62e-05	0.0130**	0.000497	0.00164***	4.20e-05
	(0.00310)	(0.00114)	(0.000673)	(0.00540)	(0.00209)	(0.000637)	(0.000155)
x young kid	0.0975***	0.0165**	0.00274	0.0845***	0.00839	0.00854**	0.000622
	(0.00801)	(0.00680)	(0.00403)	(0.0324)	(0.0125)	(0.00382)	(0.000930)
x london	0.00860**	0.00729*	0.00699***	0.00855	0.0212***	0.00382*	0.00157***
	(0.00434)	(0.00399)	(0.00236)	(0.0190)	(0.00733)	(0.00224)	(0.000545)
x non-white	-0.0226***	-0.00780	-0.00738*	-0.0553	-0.0229*	-0.00250	-0.00170*
	(0.00758)	(0.00712)	(0.00422)	(0.0339)	(0.0131)	(0.00399)	(0.000973)
x migrant	0.00136	0.000531	-0.00155	0.00343	-0.00528	-0.000822	-0.000400
	(0.00645)	(0.00627)	(0.00372)	(0.0298)	(0.0115)	(0.00352)	(0.000857)
x conscientious	-0.0101***	-0.00307***	-0.00148**	-0.0101**	-0.00467**	-0.000306	-0.000348***
	(0.00206)	(0.000986)	(0.000584)	(0.00469)	(0.00181)	(0.000554)	(0.000135)
x neurotic	0.00322	-0.00260***	0.000544	-0.00397	0.00173	-5.42e-05	0.000129
	(0.00206)	(0.000929)	(0.000550)	(0.00442)	(0.00171)	(0.000521)	(0.000127)
Log-likelihood	-12,909.25						
R-Squared	0.136	0.229	0.243	0.228	0.251	0.093	0.253
#Obs	5,501	5,501	5,501	5,501	5,501	5,501	5,501

Notes: Labor supply model estimated by MNL. Subjective well-being models estimated using OLS with the following additive controls: age and age squared, male, education, marital status, presence of children aged 0-2, household size, ethnicity, migrant, health status, home ownership, personality traits, region and year dummies.

**, **, *** indicate 1%, 5% and 10% significance levels. Standard errors in parenthesis*

As seen in the lower part of Table 1, taste shifters are more often significant in the labor supply model than in SWB regressions. This is certainly due to the fact that SWB measures are relatively noisy, and despite the treatment of individual heterogeneity described in the previous section. Nonetheless, when β_l parameters are significant, we generally observe the same sign as α_l in the labor supply model. This is the case for the presence of young children, London, non-white and conscientiousness. Sign and significance of these factors also show some regularity across the different proxies for experienced utility (the six last columns of Table 1). Notice that concentrated measures better fit the data, as indicated by the R2 in Table 1. Some of the variables can be given reasonable interpretations. For instance, the presence of

young kids is naturally associated with lower work duration on average, while high-wage workers who do work a lot despite having a child may see their SWB depressed much. Inversely, as mentioned above, conscientious workers may do long hours and experience higher SWB than less conscientious with the same work. In contrast, education, reduced to two broad groups for binary variation, shows conflicting results between labor supply estimates and SWB estimates. More educated workers may work more because of higher work preferences, yet experienced utility tells us they draw more overall satisfaction from higher leisure levels.[23]

3.2. Patterns of Indifference Curves

We now go one step further and use these estimates to depict the shape of indifference curves (ICs) in the income-leisure space. From this point onward, we use only concentrated SWB measures, which better fit the data and best represent the underlying utility from income and leisure experiences.

Figure 1: Baseline Comparison of Indifference Curves (IC)
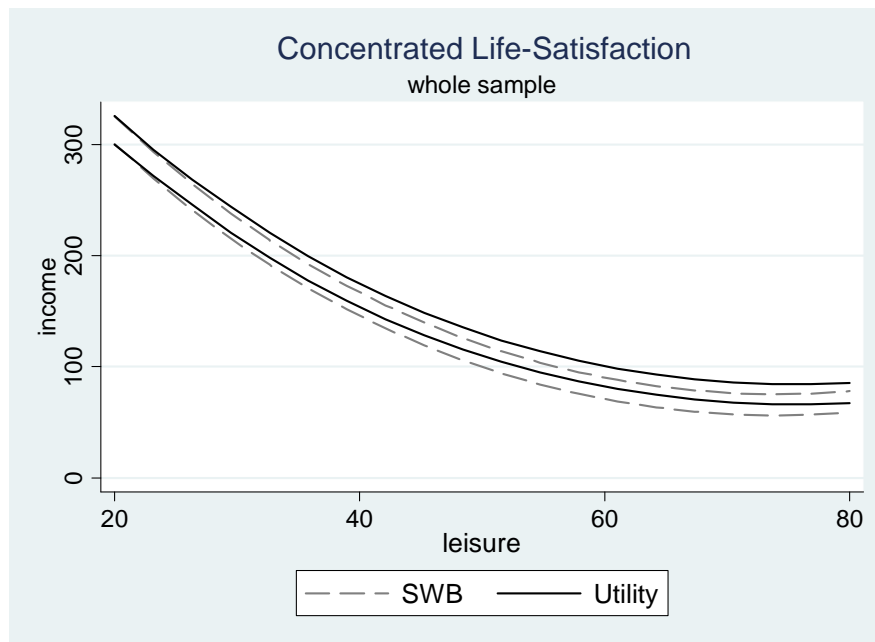
Note: SWB and Utility denote the Indifference Curves obtained using concentrated life satisfaction estimation and labor supply model estimation respectively. The SWB equation includes additive controls (male, age, age squared, education, marital status, male, presence of children aged 0-2, ethnicity, migrant, health status, home ownership, household size, region and time dummies) and individual effects (personality traits).

---

[23] Admittedly, these interpretations are valid if the models are well identified. We have argued that our approach relies on strong sources of exogenous variation in net wages. Yet, there may still be a lack of "common support" between sub-groups for some characteristics (for instance, if the high-educated are systematically better paid than the low-educated, and work more for this reason). "Optimization errors" may also be larger within particular sub-groups (for instance if the young do not evaluate correctly the income-leisure combination that suits their mental health best, compared to the old). Yet it is difficult to interpret the role of general preference shifters like age and education. In section 4, we shall turn to other, more interpretable sources of variation across individuals.

Baseline. ICs for the whole sample are derived from previous estimates and mean characteristics, setting welfare level to the empirical average of decision utility or experienced utility.[24] In Figure 1, the black solid curves represent confidence intervals of the IC derived from the labor supply model ("Utility") while the gray dashed curves represent confidence intervals of the IC from the SWB estimation ("SWB"). Total time available is T=80 hours per week so that leisure points ranging from 20 to 80 hours correspond to weekly work hours from 60 (overtime) to 0 (inactivity). Recall that we use quadratic utility but did not impose any restriction on the parameters (other specifications will be shown later). Resulting ICs nonetheless comply with economic theory, i.e. displaying a monotonically decreasing and convex pattern, almost all along the leisure range. The most interesting result is the striking similarity between ICs derived from SWB measures and from labor supply decisions.

Alternative Well-Being Measures. In Figure 2, we present ICs obtained with concentrated GHQ (Figure 2A) and concentrated happiness (Figure 2B), using the same specification as in the baseline. ICs are again similar to the IC derived from labor supply choices, confirming that alternative SWB measures capture well the overall variation in well-being from income and leisure satisfaction. The concentrated SWB using life satisfaction (Figure 1) is nonetheless the most consistent measure and the most comparable to the IC based on labor supply decisions. While indices used in Figure 2 do not tell completely different stories on the implicit tradeoff between income and leisure, they show an interesting non-monotonicity for maximum leisure (inactivity). SWB-based ICs go up a little bit, which may capture compensation needed when being inactive (the psychic cost of staying at home, cf. Clark and Oswald, 1994).[25]

---

[24] Deterministic utility in the random utility model (decision utility) is calculated for each individual at his/her observed choice using observed individual characteristics and the estimated parameters. We do the same for SWB (experienced utility). In each case, the mean utility is then used together with mean utility parameters to retrieve income as a function of leisure.

[25] We find that this compensation is even larger, and appears also in Figure 1, when keeping job seekers in the sample – for whom the psychic cost of (forced) unemployed is possibly even higher.

Figure 2: Alternative SWB Measures (Concentrated GHQ and Happiness)
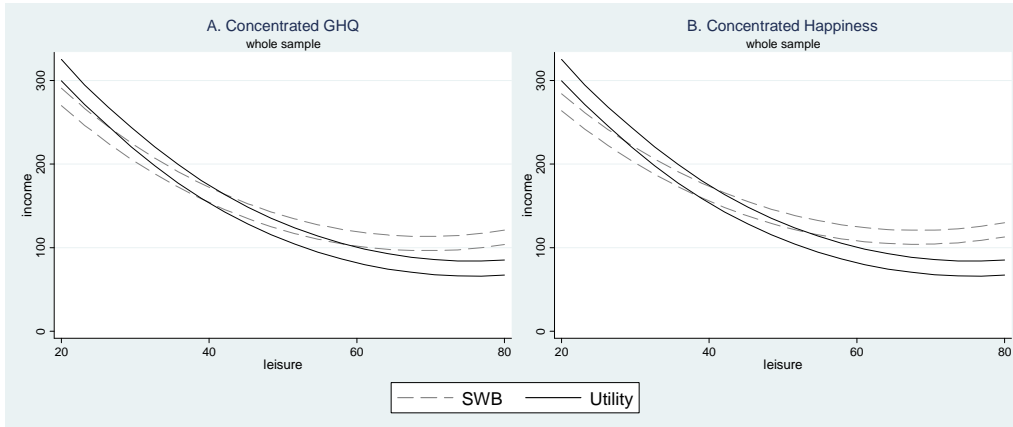
Note: SWB and Utility denote the Indifference Curves obtained using concentrated GHQ (figure A) or Happiness (figure B) and labor supply model estimation respectively. SWB equations include additive controls and individual effects as listed in Figure 1.

Sensitivity to Additive Heterogeneity in SWB Equations. Results with the labor supply model are very stable and particularly robust to alternative specifications of the deterministic utility (results available from the authors). Therefore, we focus here on sensitivity checks with respect to SWB estimations. In Figure 3A, we redefine the deterministic part of the utility in equation (6) to be most similar to specification of the labor supply model. That is, we drop additive heterogeneity z_{it} (except the terms interacted with leisure which also enter here), τ_t and θ_i. The IC with the restricted SWB specification poorly represents the income-leisure tradeoffs. In a stepwise approach, we estimate the SWB model by adding determinants of well-being one by one. We observe that the key controls that make the two ICs similar are household size and health indicators. The importance of the latter in SWB regression is well-known in this literature and carries a specific meaning in our comparison at it may represent some additional constraints put on the choice of low health people, as discussed in the next section. Household size captures the effect of single mothers whose low employment rate is due to high fixed costs of work (leading to high responses to financial incentives, cf. Blundell et al., 2000). That is, they may be dissatisfied to stay at home and have no social role/life, despite the "rational" choice of doing so if gains to work are small or even negative.[26]

These results highlight the importance of a rich SWB model specification that "clean" individual well-being measures from individual-specific situations and idiosyncratic noise. This aspect becomes particular important when trying to interpret SWB-based

---

[26] Further exploration shows that dummies for presence of one or two children gives the same result, conveying that the divergence between model is not due to outliers (single individuals living with many dependents) but simply single parents with just one or two children (recall that we controlled only for presence of *young* children in the leisure interaction terms).

relationship between income and leisure as ordinal preference (cf. Decancq et al., 2013).

Figure 3: Alternative SWB Models for the Treatment of Additive Heterogeneity
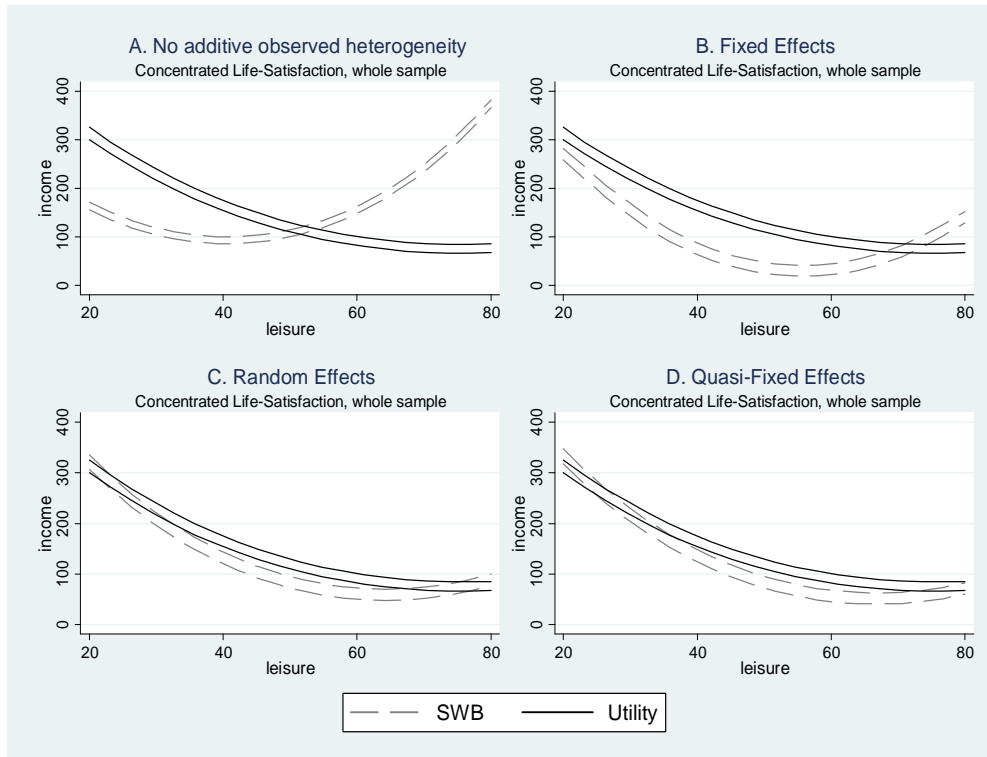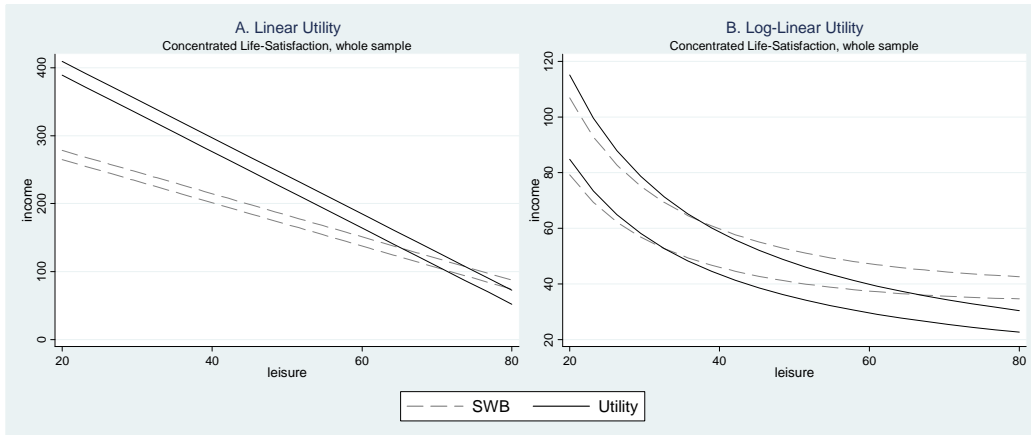

Note: SWB and Utility denote the Indifference Curves obtained using concentrated life satisfaction estimation and labor supply model estimation respectively. For Figure A, we do not include additive controls and individual effects in the SWB regression so the latter is most similar to the specification used in the labor supply model. Figure B, C and D include usual controls and individual effects for SWB as listed in Figure 1.

We pursue this specification check by addressing the nature of unobserved heterogeneity θ_i. In the baseline, we have used personality characteristics (as done, for instance, in Boyce, 2010). Alternatively, we now suggest SWB estimations using panel information to model θ_i as fixed effects (FE), random effects (RE) and quasi fixed effects (QFE), as presented in Figures 3B, 3C and 3D respectively. The FE model allows unobserved individual characteristics to be taken into account in the most flexible way (allowing possible correlation with income and leisure). The IC from SWB regressions shows a reasonable pattern, even if inactivity tends to generate a lot more subjective disutility. Importantly, let's recall that the interpretation is different here since only time variation (within estimator) is used to identify the model.[27] "Between variation" may attenuate differences (as it captures long-term trends which are

---

[27] See Fleurbaey and Blanchet (2013, p. 197) for a discussion of SWB estimations in the context of panel data.

smoothed by adaption) 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). Other estimators combine within and between variation. While RE do not deal with possible correlation income and leisure (and lead to very similar results as in the baseline), QFE à la Mundlak offer an intermediary approach whereby correlation is ruled out when controlling for 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 (see also estimates of these alternative models in Appendix Table A1).

<div align="center">Figure 4: Alternative Functional Forms</div>

Note: SWB and Utility denote the Indifference Curves obtained using concentrated life satisfaction estimation and labor supply model estimation respectively. SWB equations include additive controls and individual effects as listed in Figure 1.

Functional Forms and preference heterogeneity. Our comparison is necessarily carried out for specific forms of the deterministic utility function. The baseline models, in equations (3) and (6), rely on a quadratic function of income and leisure, widely used and appreciated as a relatively flexible specification in the labor supply literature (see Blundell and MaCurdy, 1999). We test the sensitivity of IC comparisons with respect to two other specifications which are popular in the SWB literature. The simplest yet most restrictive form is a *linear* utility function of income and leisure, $U_i(y_{ij}, l_{ij}) = \alpha_y y_{ij} + (\alpha_{l0} + \alpha_l' x) l_{ij}$. The second is a *log-linear* utility, $U_i(y_{ij}, l_{ij}) = \alpha_y \ln(y_{ij}) + (\alpha_{l0} + \alpha_l' x) \ln(l_{ij})$, often used in SWB studies (e.g., Clark et al., 2008) and capturing some non-linearity in income and leisure. Estimations are conducted using the same observed controls as in the baseline (estimates are reported in Appendix Table A2, showing again some similarities in signs and significance of coefficients with the two approaches). The ICs obtained with linear and log-linear models are depicted in Figures 4A and 4B respectively. The IC derived from SWB is flatter than the one elicited from labor supply choices in the restrictive linear specification; the shape of

implicit preferences is much more similar in both approaches when using the more flexible log-linear form. Let us finally remark that specification checks with more (rather than less) flexible forms would have to deal with the fact that interaction terms on leisure (and higher order terms for income and leisure in a polynomial form) are not significant in absence of larger datasets.

Previous results were obtained by assuming that only observed heterogeneity could explain different labor supply choices. We have also experimented with unobserved heterogeneity in relative income-leisure preferences. We have estimated models following the logic of unobserved heterogeneity in labor supply models (e.g., van Soest, 1995), that is, adding to x_{it} a zero-mean, normally distributed random effect (uniform across alternatives in the labor supply model, yet varying for each period). Then we have replaced it by a Mundlak's quasi-fixed effects, taking mean value of health status, number of children and region over time (and, hence, accounting for a proper worker's individual effect, uniform across alternatives in the labor supply model and time-invariant).[28] Treatment of this heterogeneity when constructing ICs logically implied to set new random terms to zero. Results (unreported) are very similar to Figure 1.[29]

3.3. Indifference Curves with Preference Heterogeneity

We now turn our attention to the *observed* (hence interpretable) characteristics used to elicit preference heterogeneity with respect to income and leisure. We calculate specific ICs for different groups in order to investigate how preference heterogeneity compares in both approaches.[30] In order to ease the comparison, we represent ICs where the bundle $(y, l) = (y(40), 40)$ is fixed as a common rotating point for all groups (ICs are hence derived for the model-specific and group-specific value of well-being obtained for this combination). We use the quadratic form with baseline specification. Figure 5 reports ICs for subgroups, revealing interesting observations. First there is substantial individual heterogeneity in terms of income-leisure preferences, which is possible to interpret intuitively (see section 3.1). Second, we see again similar ICs with both approaches. Yet they tend to deviate more for some groups (ex: the highly educated compared to the low-educated workers). Specific reasons explaining these differences or similarities are discussed more in detail in section 4.

---

[28] Fixed effects were not used since they would create incidental parameter problems in the nonlinear labor supply model.

[29] Note that the baseline already accounts for individual effects to the extent that we include two important personality traits (neuroticism and conscientiousness).

[30] Information on how MRS vary with individual heterogeneity was contained in Table 1 estimates, yet only in comparison to the implicit reference group for the leisure coefficient (young white woman with lower education, no young children, who is not a migrant and with low degrees of conscientiousness and neuroticism). We calculate here ICs for sub-groups mean characteristics (for instance, mean characteristics of the high-educated and of the low-educated).

Figure 5. Indifference Curve with Group-level Preference Heterogeneity
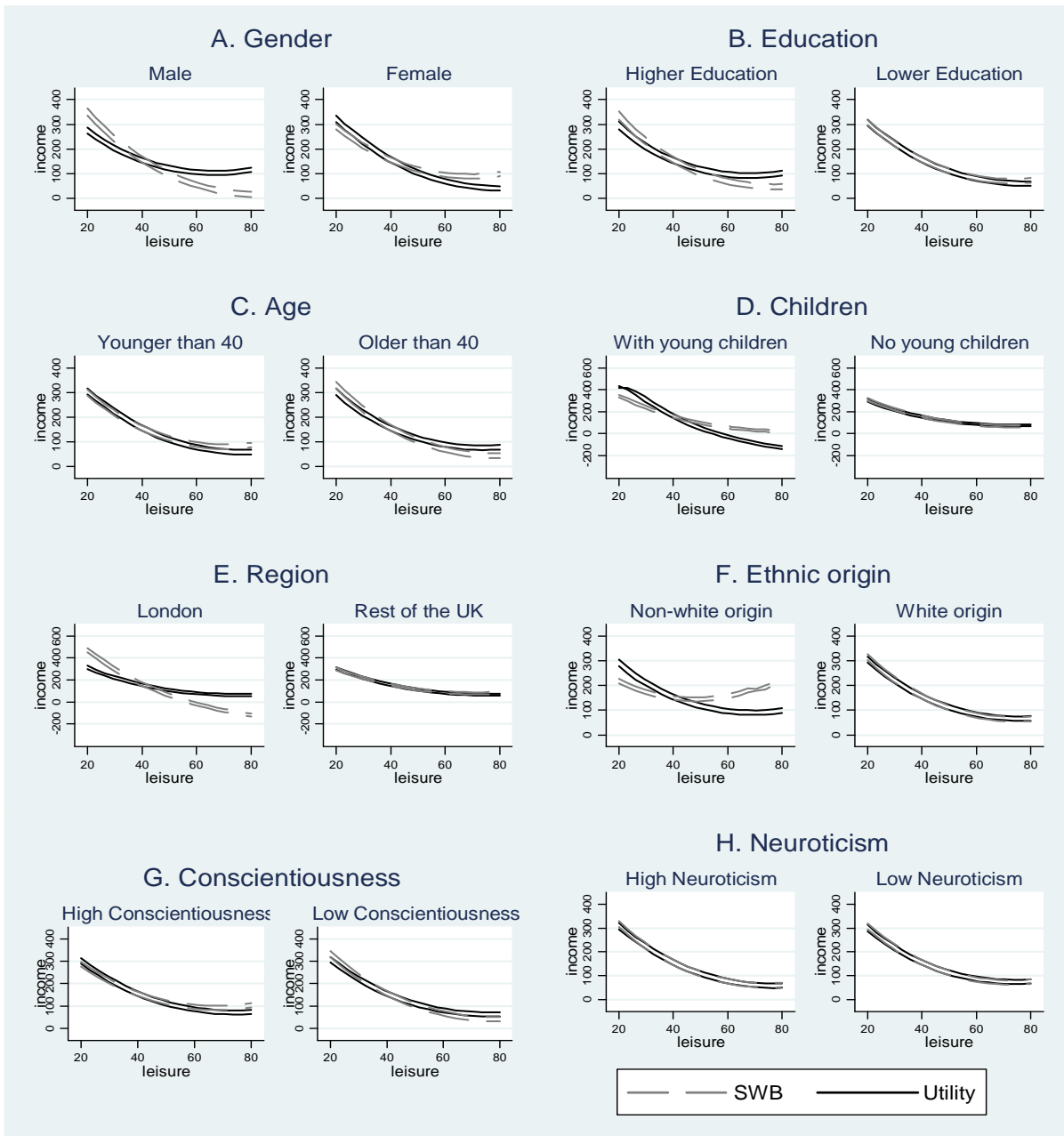
Note: SWB and Utility denote the Indifference Curves obtained using concentrated life satisfaction estimation and labor supply model estimation respectively. SWB equations include additive controls and individual effects as listed in Figure 1.

3.4. Alternative Comparison Strategies: What do People Maximize?

From previous results, it seems that labor supply choices are consistent with SWB maximization overall, yet not for some specific groups. Some people may make mistakes, change their mind, aspire or develop wrong expectations after the actual choice was made, etc. (see Kahneman and Thaler, 2006). In a recent study, Benjamin

et al. (2012) use an experimental design to investigate whether individuals are able to predict their SWB (measured as life satisfaction or happiness) at the moment of decision.[31] Another way to present our results is to test, in a related fashion, for the extent to which SWB determinants do predict choices. That is, estimates of the SWB equation are 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.[32] This terminology, borrowed from Loewenstein et al. (2003) and Loewenstein and Adler (1995), implies a first interpretation whereby SWB-maximizing errors represent failures of individuals to predict their own future choices (labor supply decisions took place before the record of their SWB consequences). Arguably, "errors" may in fact represent other types of behavioral features or simply the fact that people are constrained to some extent in their optimization, as extensively discussed in the next section.

Figure 6: Distribution of the 'Projection Error' for different preferences over hours worked
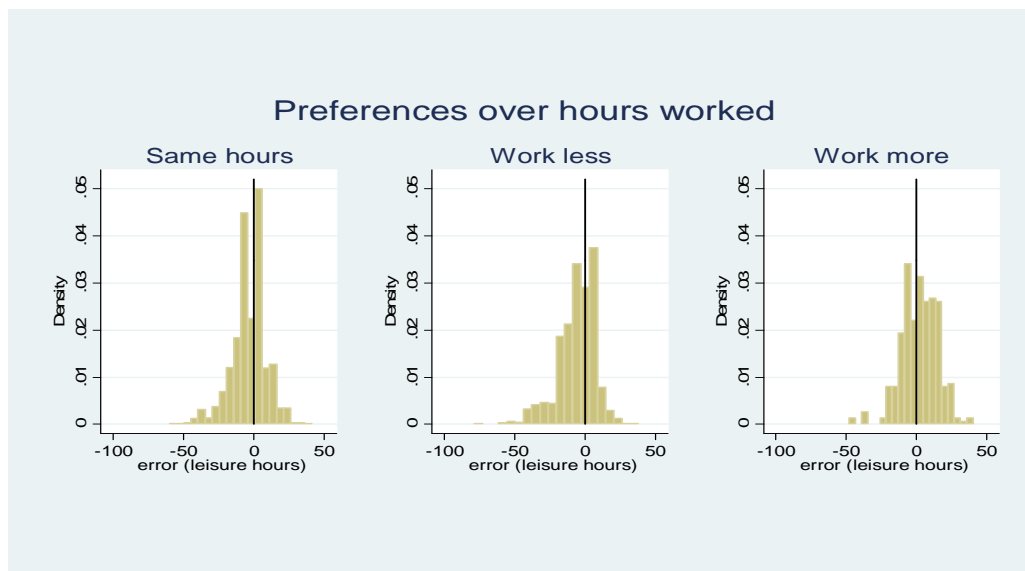
Note: the 'projection error' is calculated as the difference between actual and SWB-maximising leisure choice.

---

31 They find 83% of the subjects are able to perfectly predict their SWB when making a choice. Scores range from well below 50% to above 95% depending on the choice situations, subject pools, survey methods and SWB measure.

32 In this way, we obtain a basic evaluation of whether implicit preferences elicited from SWB actually do optimize the experienced outcome. Of course, this 'error' cannot be taken prima facie given that agents cannot always optimize (labor market constraints) or do not want to optimize (for instance, they cannot work over time, even if they want, due to family duties). There is also a possible time shift: static models assume that decision and experience occur simultaneously, while individuals may get used to past decisions (habituation) or find it difficult to change their choices (for instance due to labor market frictions).

Our calculations suggest that the mean 'error' is almost null (2.93 hours of leisure, which is much lower than the 10-hour step used for discrete labor supply modeling), with a standard deviation of 17.2 hours. Almost 50% of the people are in the range of a half standard deviation interval.[33] The 'error' distribution shown in Figure 6 (left graph) illustrates the high density at zero and the dispersion, with most people in a range of 'error' between 0 and 20 hours (in absolute value). Admittedly, there is still a significant number of people far off the mark, with labor choices that give too much or too little importance to their leisure needs for life satisfaction. In the two other graphs, we use the answer to a question on whether people would like to work more or less than what they do. We take this subjective information with caution as many interpretations are possible. Nonetheless, the distribution for people who declare that they would like to work less (middle graph) is left-skewed, indicating that they work too much compared to SWB-maximization. Those who report a wish to work more show a higher concentration in the right tail, i.e. too much leisure time compared to SWB-maximization. These findings give confidence in the method: self-reported perception about the gap between actual and desired hours tends to match with our revealed `errors'. The latter may be consistent with "projection bias" but also be the reflection of relative concerns or the effect of some constraints (for instance, the unavailability of part-time jobs for those working full time who like to work less). We now investigate these different explanatory channels.

4. Explaining Differences between Experienced and Decision Utility

We discuss factors that possibly explain differences between choice- and SWB-based ICs. These factors can affect the relationship between actual choices at the time of decision and outcomes at the time of hedonic experience. They are related to whether individuals are able to choose (market constraints may restrict choices) and whether they choose what maximizes experienced utility. Individuals may not pursue the highest SWB, or make errors when predicting the right balance between income and leisure, due to a variety of factors, including projection bias, focusing illusion and high expectations/aspirations. We suggest variable drawn from the BHPS to proxy some of these factors. This last section is more exploratory since each proxy possibly lends itself to several other interpretations - hence, our reading of the results should be taken only as a first attempt to pinpoint the underlying mechanisms.

Labor Market Constraints. We first inspect differences between approaches in an extreme case, i.e. when demand-side constraints are particularly strong. We use 12

---

33 Since the deviation is a continuous measure, we refer to meaningful brackets as proportions of a standard deviation (std). The number of individuals with errors within +/- 0.25 std is 25.6%; within +/- 0.33 std is: 35.9%; within +/- 0.5 std is: %47.7; within +/- 1 std is: 70%; within +/-1.5 std is: 83%; within +/-2 std is: 94.7%. A small share of the errors, those less than 10 hours, can be due to the approximation with discrete hour options.

regions x 10 periods, using local unemployment rates for each cell. Even if we took out job seekers, some of the (seemingly voluntary) unemployed in our data may actually be discouraged workers who face high job search on the local labor market, or in-work people who refrain from changing job (and, hence, adjusting their working time) because of high employment insecurity in their region. We expect that the difference in ICs would be larger when there are large constraints. The results are presented in Figure 7A, consistently showing less comparable ICs across approaches in highly constrained regions (top quintile of the unemployment rate distribution). Those working overtime (more than 40 hours, i.e. leisure below 40) seem to suffer more. It can also be that among those out of work (leisure at 80), the (seemingly voluntary) inactive workers suffer less, which can be interpreted as the situation of workers who internalize that they do not have to go through the high work costs of a depressed labor market or that inactivity carries less stigma when it is widespread.[34] It is not possible to discriminate between interpretation given the difference between choice- and SWB-based ICs is merely a question of slope, not of level.

Individual-Specific Constraints. Other, individual-specific constraints might explain `errors' in labor supply choices. We examine two examples of optimization constraints which are also discussed in studies comparing decision and experienced utilities: health status and commuting time (see Dolan and Kahneman, 2008; Frey and Stutzer, 2008). Figure 7B compares ICs for individuals with poor and good health respectively. ICs elicited from choices and from SWB differ very significantly for constrained people only, i.e. those with poor health. For them, long work hours come with pain and required income compensation, as rationalized by the steeper IC derived from SWB. In other words, an hour of leisure is more valuable for those with bad health compared to what their labor supply choices suggest.

Commuting, Projection bias and Focusing Illusion. Commuting may also be seen as a constraint for those whose job location has changed. In this case, results in Figure 7 (left-hand side) mean that working long hours requires more compensation that actually received when it takes time to go home after a long day.[35] In fact, longer commuting time is often associated to projection bias (people who choose far away jobs may not be able to correctly guess well-being implications, see Kimball and Willis, 2006) or focusing illusion (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

---

[34] Inversely, inactive workers in low-unemployment markets, on the right-hand side graph, seem to require income compensation for staying out of the labor market.
[35] Commuting is negatively correlated with SWB, even after controlling for endogenous sorting of individuals into location choice (Frey and Stutzer, 2008).

illusion may lead to a misallocation of time").[36] Figure 7C (left-hand side) indeed shows that high commuters work more and are more prone to make "mistakes" when choosing their working time. In contrast, those will little commuting time seem to make "reasonable" work choices as well (the two ICs broadly overlap in their case).

Figure 7: Constraints and Projection bias/Focusing Illusion
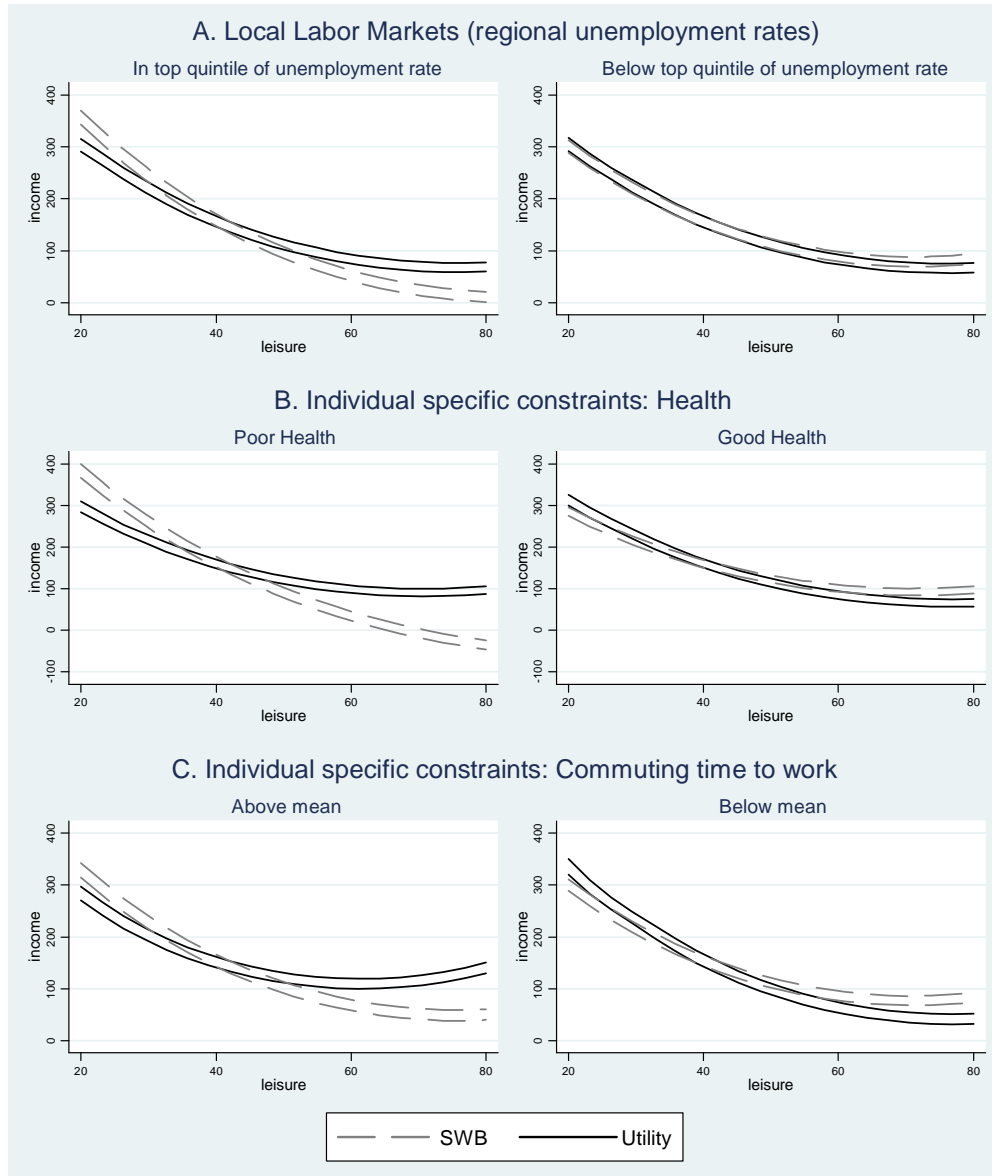
Note: SWB and Utility denote the Indifference Curves obtained using concentrated life satisfaction estimation and labor supply model estimation respectively. SWB equations include additive controls and individual effects as listed in Figure 1.

---

[36] This phenomenon, known as "focusing illusion", is identified in various situations in the behavioral economics/psychology literature (Kahneman and Thaler, 2006), for instance the illusion that living in a sunny place (Schkade and Kahneman, 1998) or having more money (Kahneman et al., 2006) would make us happier. The latter case occurs especially when people ignore the effect of hedonic adaptations to a certain level of wealth (Di Tella and MacCulloch, 2010). Consequences may be overtime work and lengthy commutes.
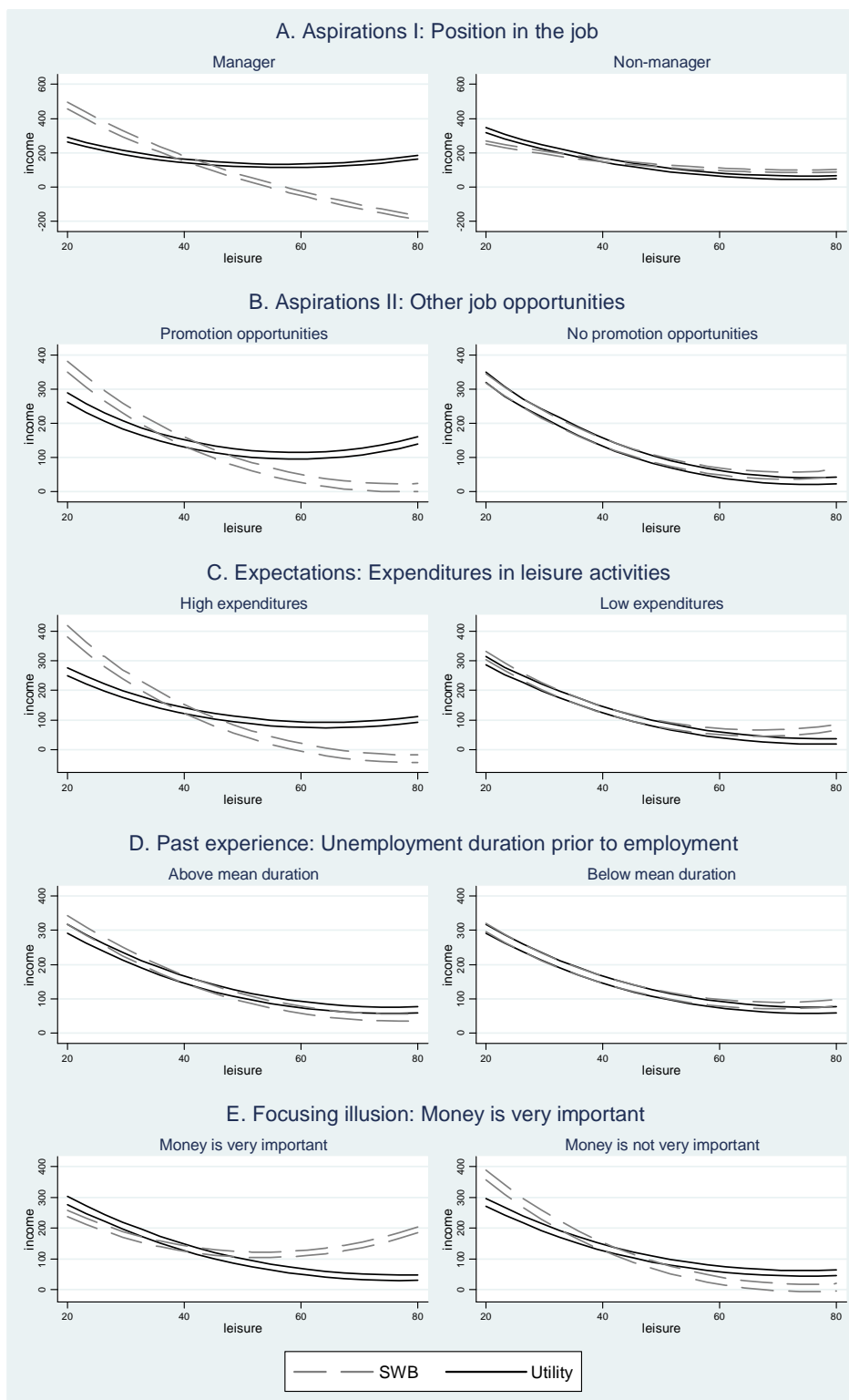
Expectations and Aspirations. Expectations and aspirations can play a critical role on SWB (e.g., Stutzer, 2004). Too high aspirations and wrong expectations about labor market outcomes may induce lower (experienced) utility for a time-consuming job, for instance. Our SWB regressions include a set of taste shifters that might be correlated with expectations and aspirations (for instance, as suggested above, males may aspire to career opportunities in a way that they work "too much" compared to their ideal income-leisure balance).[37] The problem is that these characteristics (like male, or being highly educated) capture many other dimensions. Therefore, we use additional variables that may proxy individual heterogeneity in the degree of expectations/aspirations. ICs are depicted in Figure 8. The first two variables, the position in the job and the opportunity for promotion, are related to aspirations. A higher job position and better promotion possibilities may be related to higher aspirations. Indeed, we observe some discrepancy between ICs: managers and people with high promotion opportunities work too much compared to what preferences revealed by SWB (Figures 8A and B, left-hand side). In contrast, ICs broadly overlap for non-managers and those with low promotion chances. A last finding is that when controlling for these aspiration proxies, the gap between ICs depicted on Figure 5 for some groups tends to disappears, notably for men and high-educated. This is highly consistent with the point made above on the fact that some of these general groups (like male) coincide with specific channels explaining sub-optimal decisions.

The third variable (Figure 8C) attempts to proxy the level of expectations using information on activities during leisure time. We assume that individuals who spend much on leisure activities expect more satisfaction from leisure 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). ICs elicited from labor supply and from SWB do diverge for these people, while they are very similar for others.

---

[37] A similar mechanism has been suggested for migrants in Knight and Gunatilaka (2010). Evidence in Figure 5 is consistent with this interpretation (a steeper IC from SWB compared to labor supply choices for migrants).

Figure 8: Expectations, Aspirations, Focusing Illusion

A. Aspirations I: Position in the job

Manager | Non-manager

B. Aspirations II: Other job opportunities

Promotion opportunities | No promotion opportunities

C. Expectations: Expenditures in leisure activities

High expenditures | Low expenditures

D. Past experience: Unemployment duration prior to employment

Above mean duration | Below mean duration

E. Focusing illusion: Money is very important

Money is very important | Money is not very important

SWB --- Utility

Note: SWB and Utility denote the Indifference Curves obtained using concentrated life satisfaction estimation and labor supply model estimation respectively. ICs from labor suppy are kept fixed to the baseline while ICs from SWB varies with the heterogeneity sources described in graphs A-E. SWB equations include additive controls and individual effects as listed in Figure 1.

Past Experience and the Focus on Money. Labor supply decisions might be influenced by past employment and unemployment experiences. That is, people who experienced long unemployment spells may make biased evaluations about the hedonic outcome of their decision. ICs in Figure 8D are highly consistent with the idea that the past experiences affect the relationship between decision and experience: those who have experienced above average unemployment periods seem to work more than necessary for well-being, while ICs are similar for those who experienced short unemployment spell.

Individuals can also give an exaggerated focus on a particular aspect of their life, leading to systematic optimization mistakes at the moment of decision because they ignore the overall significance of the decision. We have seen some indirect evidence of it based on commuting time. While people may overstate the importance of money, the opposite bias might exist, i.e. the fact that some people can give extreme attention to non-financial aspects of their work (its social impact, for instance). While it is uneasy to test focusing illusion in the context of non-experimental data, we suggest using the answers to the question: "how important is having a lot of money to you". The answers of the question are obtained with a scale from 1 ("Not important at all") to 10 ("Very important"). We use answers 8-10 for the type valuing money the most. This question does of course relate to different income-leisure preferences as well. Yet we do not observe marked differences in choice-based preferences between the two groups on Figure 8E. The main difference between them, consistent with preference heterogeneity, comes from SWB responses: those who care for money have steeper SWB-based ICs (left-hand side graph), meaning that they would not need compensation for long hours but suffer more when working part-time or being inactive. Those who think money is unimportant would need to be compensated when working a lot.

5. Concluding Discussion

This paper is one of the first comparing decision utility and experienced utility using a large household survey and for a type of decision, labor supply, which is at the core of analyses on social justice and optimal redistributive policies. The two welfare measures are not directly comparable due to scale differences (and general interpersonal non-comparability) so we retrieve ordinal preference representations (indifference curves, IC) which are consistent with labor supply choices on one side (revealed preferences) and SWB levels on the other, for the same individuals.

Our analysis suggests that ICs elicited from labor supply and SWB equations are consistent with economic theory. They are remarkably comparable on average, which is consistent with (but not a proof of) the fact that there is no systematic sub-optimal

labor supply choices in terms of the pursuit of well-being. We indeed find that the gap between actual and SWB-maximizing decisions is basically null on average. We nonetheless characterize marked differences for specific sub-groups, either due to constraints faced by agents or to particular optimization errors.

We believe that these results have important implication in normative economics and welfare analysis. Should we take the Benthamite concept of welfare more seriously (as suggested in the seminal paper of Kahneman et al., 1997)? Or does SWB constitute an alternative to traditional welfare measures which is only conditionally valuable (van der Rijt, 2013)? Our paper suggests interesting empirical similarities between approaches. Yet there is no philosophical reference point to establish that these empirical regularities imply conceptual relationships. Thus, we have suggested tentative identification of various forms of potential errors and influences affecting the relationship between these two utilities, as suggested in the literature. These include choice constraints, focusing illusion, expectations, aspirations. In most cases, we found intuitive correspondences between both approaches suggesting the relationship between them may not only be empirical but also conceptual.

These results would mean that SWB data is an informative measure generating consistent preference structures similar to the structural model – yet preferences with a different informational content. Yet, to be able to recover preferences, we have highlighted the fact that rich SWB specifications are required (specifications which omit important observed and unobserved determinants of SWB produce a poor comparison). With this requirement, SWB measures might be considered as a fitting alternative to construct overall welfare measures. If interpersonal comparability remains an issue, relationship between key variables like income and leisure (or other choice variables) may be derived to construct ordinal preferences or one-dimensional equivalent income based on explicit normative foundation (like the degree of responsibility attached to individual preferences, see Fleurbaey, 2008).

Finally, our approach suggests comparing choice-consistent preferences (a more appropriate terminology than "revealed preference") with SWB-consistent preference in the context of non-experimental data, which can easily be applied to other areas in economics. Given the difficulties related to (labor market) constraints and partial equilibrium effects in our context, it may actually be more appropriate to conduct our comparison exercise in less constrained areas like consumption or transportation choices. We also note that there are important limitations to our approach that we could not deal with due to data restrictions. Most importantly, our models are static and do not consider the dynamic nature of repeated occurrences of decision and experience. Modeling intertemporal decisions would require information on

consumption, labor supply states and the policy context at every period (e.g., see Haan et al., 2008). This is left for future research.

References

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

Bargain, O., Caliendo, M., Haan, P., and Orsini, K. (2010). "'Making work pay' in a rationed labour market," *Journal of Population Economics*, 23(1): 323-351.

Bargain, O., A. Decoster, M. Dolls, A. Peichl, D. Neumann, S. Siegloch (2013). "Welfare, heterogeneity and responsibility: Evidence from Europe", *Social Choice and Welfare*, 41(4), 789-817.

Bargain, O., K. Orsini and A. Peichl (2013): "Labour supply elasticities: A complete characterization for Europe and the US", Journal of Human Resources, 49(3),723-838

Benjamin, D., O. Heffetz, M. Kimball and 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

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

Blundell, R. and MaCurdy, T. (1999). Labor Supply: A Review of Alternative Approaches, in O. Ashenfelter and D. Card (eds), Handbook of Labor Economics, Vol.3A, Elsevier, Amsterdam, pp. 1559-1695

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

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

Baetschmann, G., K. Staub, and R. Winkelmann (2011). "Consistent Estimation of the Fixed Effects Ordered Logit Model", IZA Discussion Paper No. 5443.

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

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

Camerer, C., Babcock, C., Lowenstein, G. and Thaler, R. (1997). 'Labor supply of New York cabdrivers: one day at a time', *Quarterly Journal of Economics*, 112: 407–441.

Cherchye, L. and F. Vermeulen (2008), "Nonparametric analysis of household labor supply: goodness-of-fit and power of the unitary and the collective model", Review of Economics and Statistics, 90, 267-274

Clark, A. and A. Oswald (1994). "Unhappiness and unemployment", *Economic Journal*, 104, 648-659.

Clark, A. E., Frijters, P. and 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.

Decancq, K.,, M. Fleurbaey and E. Schokkaert (2013). "Happiness, equivalent incomes, and respect for individual preferences", mime

Decoster A. and P. Haan (2014). "Empirical welfare analysis in random utility models of labour supply", *International Tax and Public Finance*, Vol. 21

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

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

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

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

Fleurbaey, M. (2009). "Beyond GDP: the quest for a measure of social welfare", *Journal of Economic Literature*, 47(4), 1029-1075.

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

Frey, B., and A. Stutzer (2008). "Stress that Doesn't Pay: The Commuting Paradox", *Scandinavian Journal of Economics,* 110(2), 339-366.

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

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

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

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

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

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

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

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.

Kimball, M., and R. Willis (2006). "Utility and happiness", Mimeo.

Knight, J., Gunatilaka, R., (2010). "Great expectations? The subjective well-being of rural–urban migrants in China", *World Development*, 38 (1), 113–124.

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

Layard, R. (2005), Happiness: Lessons for a new Science, Penguin

Loewenstein, G., and D. Adler (1995). "A Bias in the Prediction of Tastes," *Economic Journal*, 929–37

Loewenstein, G., and Ubel, P.A. (2008). "Hedonic adaptation and the role of decision and experience utility in public policy," *Journal of Public Economics*, 92: 1795–1810

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

Ravallion, M., and Lokshin, M., (2001). "Identifying Welfare Effect from Subjective Questions", *Economica*, 68: 335-357.

Schkade, D.A. and Kahneman, D. (1998). "Does living in California make people happy? A focusing illusion in judgments of life satisfaction", *Psychological Science*, 9, 340 - 346

Senik, C. (2009). "Direct Evidence on Income Comparison and their Welfare Effects", *Journal of Economic Behavior and Organization*, 72, 408-424.

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

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

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

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

Appendix

Table A1: Labor Supply versus SWB models: Alternative Modelling of Unobserved Heterogeneity

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)	-0.0000388	(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)	-0.0017	(0.00288)	(0.00183)	(0.00184)
x over 40	-0.00127	0.001	0.000231	0.000444	0.000445
	(0.00207)	-0.000622	(0.000791)	(0.000628)	(0.000629)
x high educ.	-0.0216***	0.0000662	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)	-0.00055	(0.00264)	(0.000756)	(0.000756)
Log-likelihood	-12,909.25				
R-Squared	0.136	0.243	0.053	0.237	0.240
#Obs	5,501	5501	5,501	5,501	5,501

Notes: Labor supply model estimated by MNL. Subjective well-being models estimated using OLS with the following additive controls: age and age squared, male, education, marital status, presence of children aged 0-2, household size, ethnicity, migrant, health status, home ownership, personality traits, region and year dummies.

*, **, *** indicate 1%, 5% and 10% significance levels. Standard errors in parenthesis

Table A2: Labor Supply versus SWB models: Alternative Functional Forms

Coefficients	Quadratic Baseline		Linear		Log-linear	
	Labour Supply	Income-Leisure Concentrated Satisfaction	Labour Supply	Income-Leisure Concentrated Satisfaction	Labour Supply	Income-Leisure Concentrated Satisfaction
$Income^2$	-1.87e-05***	-4.82e-07***				
	(8.30e-07)	(9.12e-08)				
Income	0.0282***	0.00120***	0.00228***	0.000605***	5.582***	0.105***
	(0.00106)	(0.000135)	(0.000274)	(7.45e-05)	(0.143)	(0.0160)
$Leisure^2$	-0.00160***	-7.27e-05*				
	(5.69e-05)	(3.88e-05)				
Leisure	0.263***	0.00975**	0.0246***	0.000420	5.797***	0.0511
	(0.00761)	(0.00463)	(0.00199)	(0.00110)	(0.165)	(0.0512)
x male	-0.0404***	0.00266	-0.0278***	0.00320*	-1.079***	0.117*
	(0.00237)	(0.00170)	(0.00156)	(0.00163)	(0.0988)	(0.0696)
x over 40	-0.00127	0.00100	-0.000468	0.000880	0.115	0.00252
	(0.00207)	(0.000622)	(0.00147)	(0.000623)	(0.0969)	(0.00907)
x high educ.	-0.0216***	6.62e-05	-0.0152***	0.000400	-0.627***	0.00632
	(0.00310)	(0.000673)	(0.00208)	(0.000672)	(0.119)	(0.00827)
x young kid	0.0975***	0.00274	0.0716***	0.00203	4.380***	0.0843
	(0.00801)	(0.00403)	(0.00649)	(0.00403)	(0.487)	(0.247)
x london	0.00860**	0.00699***	-0.000415	0.00621***	0.496**	0.0567
	(0.00434)	(0.00236)	(0.00297)	(0.00237)	(0.197)	(0.0563)
x non-white	-0.0226***	-0.00738*	-0.00842	-0.00703*	-0.833**	-0.251
	(0.00758)	(0.00422)	(0.00535)	(0.00423)	(0.326)	(0.211)
x migrant	0.00136	-0.00155	7.30e-05	-0.00222	-0.00107	0.109
	(0.00645)	(0.00372)	(0.00445)	(0.00372)	(0.297)	(0.195)
x conscientious	-0.0101***	-0.00148**	-0.00505***	-0.00151***	-0.430***	-0.0183**
	(0.00206)	(0.000584)	(0.00146)	(0.000586)	(0.0973)	(0.00845)
x neurotic	0.00322	0.000544	0.00189	0.000529	0.126	0.0118
	(0.00206)	(0.000550)	(0.00147)	(0.000552)	(0.0978)	(0.00804)
Log-likelihood	-12909.25		-10828.49		-9447.16	
R-Squared	0.136	0.243	0.046	0.238	0.067	0.234
#Obs	5,501	5,501	5,501	5,501	5,501	5,501

Notes: Labor supply model estimated by MNL. Subjective well-being models estimated using OLS with the following additive controls: age and age squared, male, education, marital status, presence of children aged 0-2, household size, ethnicity, migrant, health status, home ownership, personality traits, region and year dummies.

*, **, *** indicate 1%, 5% and 10% significance levels. Standard errors in parenthesis