Day-of-the-Week Effects in Subjective Well-Being: Does Selectivity Matter?

Tumen, Semih and Zeydanli, Tugba

Central Bank of the Republic of Turkey, Paris School of Economics

7 October 2013

Day-of-the-Week Effects in Subjective Well-Being: Does Selectivity Matter?*

Semih Tumen †
Central Bank of the Republic of Turkey

Tugba Zeydanli ‡
Paris School of Economics
and
Universidade Nova de Lisboa

October 7, 2013

Abstract

Individuals tend to self-report higher well-being levels on certain days of the weeks than they do on the remaining days, controlling for observables. Using the 2008 release of the British Household Panel Survey, we test whether this empirical observation suffers from selection bias. In other words, we examine if subjective well-being is correlated with unobserved characteristics that lead the individuals to take the interview on specific days of the week. We focus on two distinct well-being measures: job satisfaction and happiness. We provide convincing evidence for both of these measures that the interviews are not randomly distributed across the days of the week. In other words, individuals with certain unobserved characteristics tend to take the interviews selectively. We conclude that a considerable part of the day-of-the-week patterns can be explained by a standard "non-random sorting on unobservables" argument rather than "mood fluctuations". This means that the day-of-the-week estimates reported in the literature are likely to be biased and should be treated cautiously.

JEL codes: C25; D60; J28.

Keywords: Day-of-the-week effects; subjective well-being; self-selection; treatment effects; BHPS.

*We thank Andrew Clark, Laszlo Goerke, Alex Michalos (the editor), Claudia Senik, four anonymous referees, and the participants of the EDEEM Jamboree 2013 Meeting (CORE) in Louvain-la-Neuve for very helpful comments and insightful suggestions. Tugba Zeydanli gratefully acknowledges financial support from the European Doctorate in Economics – Erasmus Mundus. The views expressed here are of our own and do not necessarily reflect those of the Central Bank of the Republic of Turkey. All errors are ours.

†semih.tumen@tcmb.gov.tr. Research and Monetary Policy Department, Central Bank of the Republic of Turkey, Istiklal Cad. No:10, 06100 Ulus, Ankara, Turkey.

‡Tugba.Zeydanli@malix.univ-paris1.fr. CES-Centre d'Economie de la Sorbonne – Maison des Sciences Eco. 106-112 boulevard de l'Hôpital 75647 Paris cedex 13. Nova School of Business and Economics–Campus de Campolide-Campolide, Lisboa.

1 Introduction

There is a vast literature documenting significant day-of-the-week effects in subjective well-being. Empirical studies find that individuals tend to report lower levels of happiness on Sundays and/or Mondays, whereas they tend to report higher job satisfaction levels on Fridays and/or Saturdays than the other days of the week. Recent breakthrough studies confirming the empirical relevance of the day-of-the-week effects in this literature include Taylor (2006), Akay and Martinsson (2009), and Helliwell and Wang (2011).[1] These are the widely-agreed day-of-the-week patterns extracted from the main micro-level datasets including large-scale ones such as the British Household Panel Survey (BHPS), the German Socio-Economic Panel Survey (GSOEP), and Gallup/Healthways polls as well as several small-scale surveys. The observed patterns are often attributed to the "circaseptum rhytms" (i.e., seven day cycles) hypothesis studied in the behavioral psychology literature [Larsen and Kasimatis (1990, 1991), Croft and Walker (2001)]. Overall, this literature suggests that well-being is subject to mood fluctuations and has a highly state-dependent nature.

These findings have important implications for economic modeling. The abstract concept of "utility" is at the heart of modern economics, but the main problem with this concept is that there is no direct measure of utility. Instead, several proxies are used to measure utility. In particular, self-reported well-being is a widely-agreed proxy on various aspects of individual utility. For example, Frey and Stutzer (2002) and Clark, Frijters, and Shields (2008) argue that self-reported happiness scores can be used as a general measure to proxy individual-level utility and they provide detailed reviews of the related literature. Similarly, Clark and Oswald (1996) argue that the self-reported job satisfaction score is a direct measure of individual-level utility derived from the current job. The results reported in the day-of-the-week effects literature imply that utility—as it is proxied by the subjective well-being scores—depends on

[1] Specifically, Taylor (2006) uses the BHPS data and documents that those who are interviewed on Fridays report higher levels of job satisfaction and lower levels of mental stress than those interviewed in the middle of the week. Akay and Martinsson (2009) test the same hypothesis using the GSOEP data and the result yields a "blue" Sunday. Helliwell and Wang (2011) utilize the Gallup/Healthways U.S. daily poll to examine the differences in the dynamics of two key measures of subjective well-being: emotions and life evaluation. They find no day-of-the-week effect for life evaluations, but significantly more happiness, enjoyment, and laughter; while significantly less worry, sadness, and anger on weekend than on weekdays. Earlier studies on this topic include Rossi and Rossi (1977), Stone, Hedges, Neale, and Satin (1985), Kennedy-Moore, Greenberg, Newman, and Stone (1992), and Egloff, Tausch, Kohlmann, and Krohne (1995). See Csikszentmihalyi and Hunter (2003) and Pettengill (2003) for literature surveys.

the events and circumstances that affect individuals even for only a very short period of time. In other words, this literature suggests that utility is not necessarily stable and it is subject to mood fluctuations.[2] The main principle behind this argument is that individuals assess their well-being at any given moment over time [Kahneman, Diener, and Schwarz (1999)]. However, this is in stark contrast with the neoclassical tradition—in particular, the Beckerian tradition—assuming stable preferences that do not often change over time and across states [Becker (1976)]. Although the stable preferences assumption is no longer a rigid requirement of neoclassical analysis[3], there is still considerable emphasis on preferences that do not quickly change over time or across states—otherwise, every economic phenomenon could be explained by quickly changing preferences, which would easily be labelled as a tautological statement.

In this paper, we ask if the observed day-of-the-week effects in subjective well-being suffer from selection bias. We focus on two well-being categories: happiness and job satisfaction. Sundays and/or Mondays are often regarded as "blue," so individuals are, on average, unhappy on these days. Fridays and/or Saturdays are the days in which self-reported job satisfaction is, on average, the highest. The selectivity question is a sensible one, because it may well be the case that individuals who are interviewed on Fridays or Saturdays are mostly the ones who enjoy working hard during the week and more relaxed days like Fridays or Saturdays are the only available time for them to take the survey. Similarly, it may be the case that individuals who are interviewed on Sundays represent an over sample of those doing housework and, thus, tend to report lower happiness levels. Alternatively, individuals who are not working hard throughout the week can prefer to take the survey on Sundays instead of resting. On Mondays, responding the survey could be a good reason for procrastination due to the overload of beginning of new week. These types of individuals can be unsatisfied with their jobs or their lives in general. If there is selectivity, then this would weaken the argument that individual-level "mood" regularly fluctuates over the days of the week. Instead, the existence of selectivity would suggest that the changes in self-reported well-being scores over the week likely come from the changes in the composition of interviewees over the week based on their unobserved

[2] That individual-level well-being significantly varies across the days of the week is an extreme version of short-term state dependency.

[3] See Pollak (2003).

characteristics.

To summarize, the main hypothesis we test in this paper is the following: the day-of-the-week estimates reported in the empirical literature may be contaminated with selection bias. Whether this hypothesis is rejected or not will be important for economic modeling. If the selection bias is significant and, as a result, the day-of-the-week effects disappear after selection correction, then this will cast doubt on the relevance of the "mood fluctuations" argument. Thus, the shadow hypothesis we test is the relevance of "the neoclassical stable preferences assumption" against "preferences subject to mood fluctuations." We employ a standard Heckman (1979) selection-correction procedure to test the existence of selection bias. In other words, we formally examine if subjective well-being is correlated with unobserved characteristics that lead the individuals to take the interview on some specific days of the week.

It will perhaps be useful to briefly outline what we find. We find significant positive selection both for job satisfaction and happiness measures. Specifically, we find that the ones interviewed on Fridays or Saturdays tend to report higher job satisfaction than a random sample drawn from the population of employed workers with a comparable set of observed characteristics would report. For happiness, we find that those interviewed on Sundays or Mondays tend to report lower happiness levels than a random sample drawn from the population of employed workers with a comparable set of observed characteristics would report. We move one step further and calculate various treatment effects using the techniques summarized by Heckman and Vytlacil (2007a,b), which enable us to attribute causal meanings to our estimates.

We conclude that the day-of-the-week effects reported in the literature are likely to be biased and, therefore, should be treated cautiously. Our interpretation of this result is that there is a considerable individual-level unobserved heterogeneity determining well-being scores, and the compositional changes in interviewees in terms of these heterogeneous factors drive most of the observed differences is self-reported well-being across the days of the week. Our findings suggest that the magnitude of the selection bias originating from these compositional shifts is so large that there is only little room for the "mood fluctuations" argument.

The plan of the paper is as follows. Section 2 summarizes our data and describes the statistical procedures we employ. Section 3 discusses the results in depth. Section 4 concludes. The Technical Appendix, at the end of the paper, formally presents the details of our statistical model, formulates the selection-correction procedure as well as the treatment effect parameters, and provides the related mathematical derivations.

2 Data and Methods

2.1 Data Description and Summary Statistics

We use the latest (2008) release of the British Household Panel Survey (BHPS) in our analysis. The BHPS provides information on individual, household, and job/employer related characteristics from 1991 to 2008 in the Great Britain, Scotland, Wales, and Northern Ireland. It yearly follows a nationally representative sample of households, interviews every adult member of sampled households, and assigns a unique identification number for each interviewer. The date of interview is recorded as day-month-year; the day-of-the-week on which an interview is conducted. Eighteen waves of data are available. Due to changes in the measurement instrument in Wave 1, the subjective well-being scores are higher in Wave 1 than those in other waves [Rose (1999)]. We accordingly drop Wave 1 from our analysis and use the data from Wave 2 to Wave 18. Our analysis focuses on the working population only, because the day-of-the-week patterns are more prevalent for the employed.

The individual-level job satisfaction in the BHPS dataset is reported based on a seven-point scale ranging from 1 (not satisfied at all) to 7 (completely satisfied). On each interview, the employed workers are asked to rate their job satisfaction levels regarding the promotion prospects, total income, relationship with boss, job security, able to use their initiatives in the work, the actual work itself, and hours worked. The last question about job satisfaction is “Overall, how satisfied or dissatisfied are you with your present job?”, which is again measured on the 1–7 scale and named the “overall job satisfaction.” This is a direct measure of individuals’ utility derived from their current job [Clark and Oswald (1996)]. We use this overall measure in our analysis.

Happiness/psychological well-being is derived from the General Health Questionnaire (GHQ) in the BHPS. The GHQ is widely used in the United Kingdom as a self-completion assessment measure of minor psychiatric morbidity [Goldberg and Williams (1988), McCabe, Thomas, Brazier, and Coleman (1996)]. This is a reliable indicator of mental distress [Argyle (2001)] and has been used extensively in the medical literature [Goldberg (1972), Goldberg (1978)]. The GHQ measures whether a respondent suffers from a health problem related to anxiety or depression (available at all waves) and overall life satisfaction scores (from Wave 6 onward). The following questions have been asked in the GHQ to measure happiness/psychological well-being. Have you recently:

1. Been able to concentrate on whatever you are doing?
2. Lost much sleep over worry?
3. Felt that you are playing a useful part in things?
4. Felt capable of making decisions about things?
5. Felt constantly under strain?
6. Felt you couldn't overcome your difficulties?
7. Been able to enjoy your normal day to day activities?
8. Been able to face up to your problems?
9. Been feeling unhappy and depressed?
10. Been losing confidence in yourself?
11. Been thinking of yourself as a worthless person?
12. Been feeling reasonably happy all things considered?

Answers are coded on a four-point scale: from "Disagree strongly" (coded 1) to "Agree strongly" (coded 4). The questions 1, 3, 4, 7, 8, and 12 are coded in the reverse order. The literature typically brings these scores together to provide an aggregate GHQ mental distress score. This final aggregate measure ranges from 12 to 48 [Taylor (2006), Hu, Stewart-Brown, Twigg, and Weich (2007)]. Low scores correspond to low levels of stress/depression (i.e., high feelings of well-being and happiness). We focus on this general happiness score in our empirical analysis. It will perhaps be useful to check internal consistency and test-retest reliability of this measure within our sample. To test internal consistency, we calculate the Cronbach's alpha, which is 0.88 for the general happiness measure and between 0.85–0.89 for each of the twelve items listed above. This suggests that the GHQ measures we use are internally consistent. The test-retest reliability scores—which we calculate both through the canonical correlation coefficient and intra-class correlation coefficient based on a mixed-effects linear model—range in the interval 0.52–0.75, which means that the GHQ measure has a reasonably reliable correlation between the test and retest for an annual survey. All the coefficients are significant at 1 percent level.

For the individual- and job-related characteristics, we follow the recent studies using the BHPS and control for gender, age, age-squared, education levels, preferences over working hours, types of contract, size of establishment, promotion opportunities, union membership, and health status [Taylor (2006)]. We collapse the education-levels into seven broad groups as follows: *higher degree* refers to postgraduate education, *first degree* refers to college education, *'A'-level*, *'O'-level* and *other higher qualification* refer to high school graduates of different types (consistent with the UK education system), *vocational qualification* refers to teaching, nursing, commercial, apprenticeship, and the certificate of secondary education (CSE), and, finally, the ones with *no qualification*. We also construct a dummy variable ("Income") for earnings. It is equal to 1 if the worker earns more than the median level of earnings in her reference group (in the corresponding wave) and it is equal to 0 otherwise. The reference groups are simply the region-industry combinations, in which the individuals can potentially interact [see Tumen and Zeydanli (2012b,a) for more details on the construction of the reference groups]. We construct such a variable to control for the group-level analogue of the Easterlin paradox, named after a series of work by Richard Easterlin including Easterlin (1974, 1995, 2001).

Table (1) presents the summary statistics of the data that we use in our analysis. In order to be included into our sample, the respondent must be employed and have reported an overall job satisfaction score or a general happiness score. The mean age of the respondents is around 39. Among the 69,000 observations, 50% are male, 56% are married, 33% are never married, 2.9% have higher degree, 12.3% have first degree, another 13.2% have 'A'-level degree, 21.2% have 'O'-level degree, 26.2% have other higher qualifications, 11.6% have vocational qualifications, and the remaining 12.7% have no qualifications. 2.9% and 1.7% have temporary and fixed term contracts, respectively. 17% work in the public sector and 69.1% work in a company of size 200 workers or smaller. 23.9% are union members. 8% prefer to work more hours and 31.4% prefer to work fewer hours. 25% report their health to be very good, whereas 15.9% report to be satisfactory. 52.4% earn above the median monthly income in their respective reference groups. The mean overall job satisfaction rate is 5.38 out of 7, with a standard deviation of 1.296. The mean general happiness score is 22.75 out of 48, with a standard deviation of 5.073. Notice that there are two dummy variables in the table labeled Fri/Sat and Sun/Mon. Fri/Sat is equal to 1 if the interview is conducted on a Friday or Saturday and 0 otherwise. Sun/Mon is equal to 1 if the interview is conducted on a Sunday or Monday and 0 otherwise. 18.9% of the workers in our sample are interviewed on Friday or Saturday, while 24.7% on Sunday or Monday. All the means and the standard deviations reported in Table (1) are calculated using the BHPS frequency weights.

2.2 Empirical Methods

Our empirical analysis features a selection-correction procedure that aims to remove the potential biases in the observed day-of-the-week patterns. Our starting point is the idea that the days on which the interviews are conducted may not be random; that is, certain types of individuals may "choose" to take the survey on certain days of the week. These selectivity issues may substantially affect the estimates if they are not appropriately addressed. Building on the Roy model [see Roy (1951) and Heckman and Honore (1990)], we develop a random-utility framework to model each respondent's choice of the interview day. The technical details of the model along with the mathematical derivations are provided in the Technical Appendix.

Before presenting the results, it will perhaps be useful to motivate our empirical strategy briefly. As we mention above, we hypothesize that the interview days may be selectively—in a systematic way—determined by the respondents. For example, those who are interviewed on Fridays and Saturdays might be the ones who enjoy working hard during the working days and who are able to file their responses only on their off days. This type of individuals might be the ones who are more likely to report higher well-being levels on average. Similarly, the individuals who are interviewed on Sundays may represent those doing housework and, thus, tend to report lower well-being scores. Taking the survey on Mondays might be a signal of low job/life satisfaction. If the selectivity argument is empirically valid, then the day-of-the-week effects reported in the literature might be biased. In particular, the existence of selectivity might suggest that changes in self-reported well-being scores over the week likely come from the changes in the composition of the interviewee types over the week (i.e., unsatisfied types tend to take the test on Sundays/Mondays, whereas the satisfied types tend to take it on Fridays/Saturdays) rather than mood fluctuations.

To perform this task, we run a first-stage probit regression to estimate each respondent's choice of the interview day. Then, we calculate inverse Mills ratios using the regression output from the first stage. At the second stage, we regress the subjective well-being measures on observed covariates and these inverse Mills ratios. Incorporating the inverse Mills ratios serves the purpose of removing the potential selection biases. This is the standard Heckman correction procedure.[4] We then use the parameter estimates to construct several treatment effects parameters, mainly the Average Treatment Effect (ATE), Treatment on the Treated (TT), and Treatment on the Untreated (TUT). The treatment effect estimates are often appealed in the causal inference literature and we will use these estimates to interpret the selectivity patterns we detect. The Technical Appendix provides the details of these statistical procedures.

One important point—that deserves specific attention—is the exclusion restriction that is needed to guarantee econometric identification in the selection-correction procedure we describe above. The next subsection presents the details of the exclusion restriction we use in

[4]See Heckman (1979) for the original paper.

our analysis.

2.3 The Exclusion Restriction

There are two traditional ways through which the Heckman selection-correction method can be applied. The first one is the existence of an additional variable in the selection equation, which does not affect the outcome of interest. This is called the "exclusion restriction" (or "instrument") and it secures identification of bias-corrected estimates. The second one is to use the nonlinearity inherent in the inverse Mills ratio. In this case, identification solely comes from the normality assumption. The latter is disadvantageous for two reasons; (i) self-selection may not be originating from a normally distributed process [Little and Rubin (1987)] and (ii) the inverse Mills ratio may still be highly collinear with the other regressors in the outcome equation [Leung and Yu (1996)]. A potential disadvantage of the exclusion restriction approach is that there is no natural guide to specify a variable that affects the choice but does not affect the outcome; moreover, a wrongful implementation of the restriction may be harmful [Manning, Duan, and Rogers (1987)]. Nevertheless, the main consensus is that, using an appropriate exclusion restriction, if there exists one, will secure a more convincing identification of the selection-corrected estimates. Fortunately, the BHPS dataset allows us to construct a sensible exclusion restriction for our analysis.

We use the interviewer ID as an exclusion restriction. More precisely, we rely on the identifying assumption that who the interviewer is a determinant of when the interviewee takes the survey, but it is not a determinant of the survey outcome (i.e., happiness and job satisfaction). Figures (1) and (2) plot the tendencies of the interviewers in terms of timing. To be concrete, Figure (1) presents the distribution of the interviewers' probabilities of conducting the survey on a Friday or Saturday. Similarly, Figure (2) describes the distribution of the interviewers' probabilities of conducting the survey on a Sunday or Monday. For example, a value of 0.4 read on the horizontal axis should be interpreted as a 40% of the interviews conducted by that particular interviewer are on a Friday or Saturday. Clearly, some interviewers are more likely to conduct the interview of certain days.

The validity of this exclusion restriction is justified in the empirical analysis presented by Taylor (2006). He shows that the interviewer ID likely affects the day on which the interview is conducted; but, it does not affect the outcome (i.e., the subjective well-being score). We follow this suggestion and use interviewer ID as an exclusion restriction in our selection-correction exercise.

To construct the variable that we use as the exclusion restriction, we determine the mean values in these two distributions. We generate a binary variable taking the value 1 if the interviewer's probability is greater than the mean and 0 otherwise. This new dummy variable characterizes if the interviewer is more likely to conduct the interview on a Friday or Saturday (Sunday or Monday) than the average tendency in the job satisfaction (happiness) analysis. The mean tendency to conduct the interview on a Friday or Saturday is 0.189 for job satisfaction and the corresponding mean tendency to conduct the interview on a Sunday or Monday for happiness is around 0.24. Table (5) documents that this binary variable (i.e., interviewer dummy) is a relevant determinant of the day of interview. Intuitively, who the interviewer is should not be a systematic determinant of well-being. As a result, we use this dummy variable as an exclusion restriction in our selection-correction exercise.

3 Results and Discussion

In this section, we document the empirical results and provide an extensive discussion of the main implications of our analysis. We start with a simple observation. Fridays and Saturdays are the days on which the self-reported job satisfaction scores are higher, on average, than the scores reported on the other days. Moreover, Sundays and Mondays are the days on which the self-reported happiness scores are lower, on average, than the scores reported on the remaining days of the week. These raw patterns are best observed from the results of an OLS regression of the associated well-being score on the day dummies. Tables (2) and (3) document these patterns.

The second step is to see whether including observed characteristics into these regressions changes these results or not. We include a comprehensive set of regressors for both worker- and job-related characteristics. The worker-related regressors include age as a quadratic polynomial and dummy variables for gender, marital status, education, health, region, and the year of interview. The job-related regressors include dummy variables for job contractual status, permanency of job, promotion opportunities, union membership status, public/private sector job, firm size, preference for work hours, relative income, and industry.[5] We perform two separate regressions for job satisfaction and happiness controlling for these variables as well as the day-of-the-week dummies. We find that the results of the simple regressions described above are reinstated; that is, on average, job satisfaction scores are higher for those interviewed on Fridays or Saturdays and happiness scores are lower for those interviewed on Sundays or Mondays. These results are in line with the day-of-the-week patterns documented by the main papers in the related literature.[6] See Table (4) for the results.

We investigate if there is any sorting on unobservables that can potentially bias these day of the week patterns. If self-selection is in effect, then individuals with certain unobserved characteristics tend to take the survey on certain days of the week. For example, those who take the survey on a Friday or Saturday may be the ones who are the most satisfied with their jobs. These individuals may have a strong motivation to work hard during the week and the only available time for them to respond may be a Friday afternoon or a Saturday. In this example, "motivation" is the unobserved variable. There may also be other unobserved factors which are also relevant for this example. For happiness, those who take the survey on a Sunday or Monday may be the ones who are the most unhappy ones with their jobs or lives in general due to some unobserved factors. These individuals may be, say, the least conscientious[7] ones, therefore they are the ones who are more likely to express their unhappiness at the end of a weekend vacation or at the beginning of a busy week. One can easily extend these examples.

If selection is a concern, then the differences in days, in terms of subjective well-being outcomes, may be driven by these unobserved individual-level heterogeneity components. In other

[5] Table (1) presents the basic summary statistics for these variables as well as the outcome variables.

[6] See, for example, Taylor (2006) and Helliwell and Wang (2011).

[7] Conscientiousness is one of the big-five personality traits that constitute an individual's non-cognitive skills. See Borghans, Duckworth, Heckman, and ter Weel (2008) for an extensive description of these concepts.

words, the day of the week patterns can be explained by non-random sorting on unobservables if self-selection is strong. To test this hypothesis, we perform a simple selection-correction procedure motivated by a combination of the Roy model with a standard random utility specification described in Section 3. As we discuss in Section 3.3, the interviewer ID is used to construct an exclusion restriction to secure identification. Table (5) documents the results our first-stage probit estimations. Tables (6) and (7) report the second-stage regressions, in which we use the inverse Mills ratios constructed from the first-stage as regressors.

Our results reveal that selectivity is very strong; that is, individuals sort into the days of the week based on their unobserved characteristics that affect the outcomes. Back of the envelope calculations (i.e., averaging the inverse Mills ratio and multiplying this average with the estimated coefficient) yield the result that almost all of the difference between Friday/Saturday ($D = 1$) and the remaining days ($D = 0$) disappear after controlling for selectivity. Similar calculations show that, after controlling for selectivity, Sunday and Monday ($D = 1$) are actually happier days than the other days of the week ($D = 0$).

We perform a further investigation of these selectivity patterns using the treatment effect parameters described in Section 3.2. Tables (8), (9), (10) and Figures (3), (4) document these estimates. The existence of selectivity is confirmed from the result that TT > ATE > TUT. The treatment on the treated parameter is quite high, strongly supporting the "non-random sorting on unobservables" idea.

We further show that, for job satisfaction, the estimated treatment effects are higher among males, non-married workers, workers with permanent jobs, public sector workers, workers in large firms, union members, workers with good health, workers who prefer to work less, workers with higher relative income, workers with higher education, and middle-aged workers. These patterns are important, because the existence of significant selectivity signals that the OLS estimates of the coefficients of other observed covariates are biased. There is a consensus in the empirical job satisfaction literature using BHPS—see, e.g., Taylor (2006)—that, on average, females, married workers, and workers with low education levels are more satisfied with their jobs. Our estimates show that these results are biased. For example, married workers are known to be more satisfied jobwise. Our raw OLS estimates reported in Table (4) reads a coefficient of 0.066 for marriage. The selection corrected estimates yield coefficients of 0.025 for the $D = 1$ sector and 0.075 for the $D = 0$ sector. In terms of our results, this suggests that workers whose unobserved characteristics lead to relatively lower job satisfaction ($D = 0$) tend to be married and this generates a higher coefficient in the $D = 0$ group versus a much lower coefficient in the $D = 1$ group. The signs of the coefficients have not changed after correcting for selection, but the magnitudes have become much weaker. Another example is for the job satisfaction patterns across age groups [see Figures (3) and (4)]. The literature reports that—see, e.g., Clark, Oswald, and Warr (1996)—there is a U-shaped relationship between job satisfaction and age.[8] Our findings reveal that the U-shaped relationship is preserved, but the magnitudes get weaker after correcting for selectivity. All of these patterns are also observed for happiness along similar lines.

The main practical implication of this study is that the observed day-of-the-week effects are mostly due to compositional shifts rather than behavioral changes. We show that the compositional effects are driven by heterogeneity in unobserved factors that diffuse into individuals' choice of the interview date. This result does not mean that psychological factors have no effect on well-being. It rather suggests that the observed day-of-the-week patterns should not be interpreted as direct evidence of the link between "mood" and well-being. Uncovering the details of the unobserved factors driving compositional shifts is an interesting topic for future research, but it is out of the scope of this paper.

We conduct our analysis with the BHPS, which is a representative dataset for the United Kingdom. This means that both the observed day-of-the-week patterns and the results of the selection-correction exercise should be interpreted taking the British cultural norms as the benchmark. Depending on the country, norms, and even religious beliefs, the observed day-of-the-week patterns may change. For example, Monday is a major "blue" day in the United Kingdom, while Sunday is shown to be "blue" in Germany [Akay and Martinsson (2009)]. In

[8]See, for example, Blanchflower and Oswald (2008) for similar findings for happiness.

North America, on the other hand, Sunday is often reported as a happy day [Helliwell and Wang (2011)]. The results may change further in, say, Muslim or Jewish societies. Although the observed day-of-the-week patterns tend to change across cultures, we believe that our analysis has broader implications that can be relevant for other countries, regions, and cultures. In some sense, our results imply that the cross-cultural differences in the observed day-of-the-week patterns will tend to disappear after correcting for selectivity. But, it may well the case that the unobserved characteristics that lead to selectivity can also be based on cultural differences, social norms, differences in working hours, worker motivation, expectations, etc. Further empirical research is needed to test the validity of these cross-cultural concerns.

The BHPS is not the only dataset on which our procedures can be implemented. The same empirical exercise can be performed for other countries, where micro-level subjective well-being datasets are available with proper "date of interview" information. For example, a similar analysis may also be conducted using the German Socio-Economic Panel (GSOEP) dataset. Another dataset that can be used is the Gallup/Healthways U.S. daily poll. However, the same analysis cannot be carried out with datasets like World Values Survey, Euro-barometer, and International Social Survey Programme, because the date of interview is mostly missing in these datasets.

A potential limitation of our analysis is related to the instrument we use in the selection-correction exercise. As we discuss in the Technical Appendix, the instrument—or the exclusion restriction—has to be correlated with the choice of the interview day, but it has to be uncorrelated with the outcome, i.e., the subjective well-being score. In other words, the instrument has to be placed into the choice equation, but excluded from the outcome equation to guarantee identification. To satisfy these requirements, we use the interviewer ID number as our exclusion restriction. We argue that who the interviewer is can affect the interview day, because some interviewers may be more likely to work at weekends than the others. However, we also argue that the interviewer ID has very little or no effect on the interviewee's responses. The limitation may apply at this point: if the interviewee's response is systematically affected by the interviewer ID, then this logic would not work. We perform several robustness checks to question the relevance of this concern. We find that some interviewers are indeed much more inclined to conduct interviews at weekends than the others. We also find that the interviewer dummies are mostly insignificant in the regression of the well-being score on all the explanatory variables and the interviewer ID's. This provides suggestive evidence that the interviewer ID might be a valid instrument.

4 Concluding Remarks

In this paper, we investigate whether the day-of-the-week effects reported in the empirical subjective well-being literature suffer from selectivity bias. We use the BHPS dataset to answer this question. Our answer is yes; that is, we show that the observed day-of-the-week patterns can be regarded as a by product of non-random sorting of individuals into the days of the week. More precisely, we show that individuals who take the BHPS interview on a Friday or Saturday—the days on which the self-reported job satisfaction score is the highest—are selectively different in terms of their unobserved characteristics from the ones interviewed on the remaining days. Similarly, the individuals who take the BHPS interview on a Sunday or Monday—the days on which the self-reported happiness score is the lowest—are selectively different in terms of their unobserved characteristics from the ones interviewed on the remaining days. We also discuss the potential channels through which the self-selection process operates.

The previous literature argues that, everything else constant, the individual well-being is lower in certain days of the week than the remaining days. This is generally interpreted as an evidence supporting the view that individuals assess their well-being at any given moment over time. Subjective well-being measures are often used to proxy individual-level utility (or preferences), which is the main building block of the theory of economic decisions. Thus, if well-being is an "objective" motivating economic choices, then the decisions made on Sundays would be different than those made on, say, Wednesdays. This implies that behavioral changes can mostly be attributed to psychological factors. However, this is strictly against the neoclassical economic theory, which is built on the basic idea that preferences should not change often (i.e., they are stable). Our findings provide evidence that the existence of weekly cycles in individual well-being may not be as relevant as the literature documents. Our results reveal that interpreting the observed day-to-day differences in the average subjective well-being scores as mood fluctuations might be incorrect. We *do not* say that preferences are not affected by psychological motives. We say that ruling out the neoclassical economic theory based on the uncorrected day-of-the-week patterns might produce misleading results.

We provide an alternative explanation for the observed day-of-the-week patterns in subjective well-being scores: the composition of survey respondents in terms of their unobserved characteristics changes across the days of the week on a non-random basis. We argue that these compositional shifts have a potential to be falsely interpreted as mood fluctuations. That said, we do not totally rule out the state-dependent nature of utility. Utility may be changing across states if these states reflect some fundamental feature of individual utility; such as employment status, marital status, etc. We rather argue that day-to-day shifts in agents' valuation of economic objects do not have strong empirical basis, when selectivity is controlled for.

References

Akay, A. and P. Martinsson (2009): "Sundays are Blue: Aren't They? The Day-of-the-Week Effect on Subjective Well-Being and Socio-Economic Status," IZA Discussion Paper No. 4563.

Argyle, M. (2001): *The Psychology of Happiness*, London: Routledge.

Becker, G. S. (1976): *The Economic Approach to Human Behavior*, Chicago, IL: University of Chicago Press.

Blanchflower, D. and A. J. Oswald (2008): "Is Well-Being U-shaped over the Life Cycle?" *Social Science & Medicine*, 66, 1733–1749.

Borghans, L., L. A. Duckworth, J. J. Heckman, and B. ter Weel (2008): "The Economics and Psychology of Personality Traits," *Journal of Human Resources*, 43, 972–1059.

Clark, A. E., P. Frijters, and M. Shields (2008): "Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles," *Journal of Economic Literature*, 46, 95–144.

Clark, A. E. and A. J. Oswald (1996): "Satisfaction and Comparison Income," *Journal of Public Economics*, 61, 359–381.

Clark, A. E., A. J. Oswald, and P. Warr (1996): "Is Job Satisfaction U-Shaped in Age," *Journal of Occupational and Organizational Psychology*, 69, 57–81.

Croft, G. P. and A. E. Walker (2001): "Are the Monday Blues Ail in the Mind? The Role of Expectancy in the Subjective Experience of Mood," *Journal of Applied Social Psychology*, 31, 1133–1145.

Csikszentmihalyi, M. and J. Hunter (2003): "Happiness in Everyday Life: The Uses of Experience Sampling," *Journal of Happiness Studies*, 4, 185–199.

Easterlin, R. A. (1974): "Does Economic Growth Improve the Human Lot?" in *Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz*, ed. by P. A. David and M. W. Reder, New York, NY: Academic Press.

——— (1995): "Will Raising the Incomes of All Increase the Happiness of All?" *Journal of Economic Behavior & Organization*, 27, 35–48.

——— (2001): "Income and Happiness: Towards a Unified Theory," *Economic Journal*, 111, 465–484.

Egloff, B., A. Tausch, C. Kohlmann, and H. Krohne (1995): "Relationships Between Time of Day, Day of the Week, and Positive Mood: Exploring the Role of the Mood Measure," *Motivation and Emotion*, 19, 99–110.

Frey, B. S. and A. Stutzer (2002): "What Can Economists Learn from Happiness Research?" *Journal of Economic Literature*, 40, 402–435.

Goldberg, D. P. (1972): *The Detection of Psychiatric Illness by Questionnaire*, Oxford, UK: Oxford University Press.

——— (1978): *Manual of the General Health Questionnaire*, Windsor, UK: NFER.

Goldberg, D. P. and P. Williams (1988): *A User's Guide to the General Health Questionnaire*, Windsor, UK: NFER.

Heckman, J. J. (1979): "Sample Selection Bias as a Specification Error," *Econometrica*, 47, 153–161.

Heckman, J. J. and B. E. Honore (1990): "The Empirical Content of the Roy Model," *Econometrica*, 58, 1121–1149.

Heckman, J. J. and R. Robb (1985): "Alternative Methods for Evaluating the Impact of Interventions: An Overview," *Journal of Econometrics*, 30, 239–267.

Heckman, J. J. and E. J. Vytlacil (2007a): "Econometric Evaluation of Social Programs, Part I: Causal Models, Structural Models, and Econometric Policy Evaluation," in *Handbook* *of Econometrics*, ed. by J. J. Heckman and E. E. Leamer, New York, NY: Elsevier, vol. 6, chap. 70, 4779–4874.

——— (2007b): "Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast their Effects in New Environments," in *Handbook of Econometrics*, ed. by J. J. Heckman and E. E. Leamer, New York, NY: Elsevier, vol. 6, chap. 71, 4875–5143.

Helliwell, J. F. and S. Wang (2011): "Weekends and Subjective Well-Being," NBER Working Paper No. 17180.

Hu, Y. J., S. Stewart-Brown, L. Twigg, and S. Weich (2007): "Can the 12-Item General Health Questionnaire be Used to Measure Positive Mental Health?" *Psychological Medicine*, 37, 1005–1013.

Kahneman, D., E. Diener, and N. Schwarz (1999): *Well-Being: The Foundation of Hedonic Psychology*, New York, NY: Russell Sage Foundation.

Kennedy-Moore, E., M. A. Greenberg, M. G. Newman, and A. A. Stone (1992): "The Relationship Between Daily Events and Mood: The Mood Measure may Matter," *Motivation and Emotion*, 162, 143–155.

Larsen, R. J. and M. Kasimatis (1990): "Individual Differences in Entrainment of Mood to the Weekly Calendar," *Journal of Personality and Social Psychology*, 58, 164–171.

——— (1991): "Day-to-Day Physical Symptoms: Individual Differences in the Occurrence, Duration, and Emotional Concomitants of Minor Daily Illnesses," *Journal of Personality*, 59, 387–423.

Leung, S. F. and S. Yu (1996): "On the Choice Between Sample Selection and Two-Part Models," *Journal of Econometrics*, 72, 197–229.

Little, R. J. A. and D. B. Rubin (1987): *Statistical Analysis with Missing Data*, New York, NY: John Wiley & Sons.

Manning, G. M., N. Duan, and W. H. Rogers (1987): "Monte Carlo Evidence on the Choice between Sample Selection and Two-Part Models," *Journal of Econometrics*, 35, 59–82.

McCabe, C. J., K. J. Thomas, J. E. Brazier, and P. Coleman (1996): "Measuring the Mental Health Status of a Population: A Comparison of the GHQ-12 and the SF-36," *British Journal of Psychiatry*, 169, 516–521.

Pettengill, G. N. (2003): "A Survey of the Monday Effect Literature," *Quarterly Journal of Business and Economics*, 42, 3–27.

Pollak, R. A. (2003): "Gary Becker's Contributions to Family and Household Economics," *Review of Economics of the Household*, 1, 111–141.

Rose, M. (1999): "Explaining and Forecasting Job Satisfaction: The Contribution of Occupational Profiling," Unpublished manuscript, University of Bath.

Rossi, A. S. and P. E. Rossi (1977): "Body Time and Social Time: Mood Patterns by Menstrual Cycle Phase and Day of Week," *Social Science Research*, 6, 273–308.

Roy, A. D. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3, 135–146.

Stone, A. A., S. M. Hedges, J. M. Neale, and S. Satin (1985): "Prospective and Cross-Sectional Mood Reports Offer No Evidence of a "Blue Monday" Phenomenon," *Journal of Personality and Social Psychology*, 1, 129–134.

Taylor, M. P. (2006): "Tell Me Why I Don't Like Mondays: Investigating Day of the Week Effects on Job Satisfaction and Psychological Well-Being," *Journal of the Royal Statistical Society (Series A)*, 169, 127–142.

Tumen, S. and T. Zeydanli (2012a): "Is Happiness Contagious? Separating Spillover Externalities from the Group-Level Social Context," Unpublished manuscript, Paris School of Economics.

——— (2012b): "Social Interactions in Job Satisfaction," Unpublished manuscript, Paris School of Economics.

A Technical Appendix

A.1 The Random Utility Model

The econometric framework we use is a standard random-utility specification in combination with a version of the two-sector Roy model [Roy (1951)].[9] Suppose that the survey respondents can choose whether to take the survey on certain days of the week ($D = 1$) versus the remaining days ($D = 0$). For our job satisfaction analysis, $D = 1$ refers to taking the survey on a Friday or Saturday and $D = 0$ refers to taking it on the remaining days of the week. Similarly, for happiness, $D = 1$ refers to taking the survey on a Sunday or Monday and $D = 0$ refers to taking it on the remaining days of the week. For simplicity, we mention throughout this section only $D = 1$ or $D = 0$ without a further reference to the days associated with these choices.

The equations motivating the individuals' choice of $D = 1$ versus $D = 0$ can be written as follows:

$$U_0 = \boldsymbol{Z}\boldsymbol{\alpha}_0' + \nu_0, \tag{A.1}$$

$$U_1 = \boldsymbol{Z}\boldsymbol{\alpha}_1' + \nu_1, \tag{A.2}$$

where $\boldsymbol{Z}$ is a row-vector of observed covariates. This is the standard additive random-utility specification, where $\boldsymbol{\alpha}_0'$ and $\boldsymbol{\alpha}_1'$ are the deterministic components, and ν_0 and ν_1 are the random components.

To rationalize the choice of D, we assume an index function

$$I = U_1 - U_0, \tag{A.3}$$

which can be rewritten, after plugging in the random utility equations, as

$$I = \boldsymbol{Z}\boldsymbol{\gamma}' + \eta, \tag{A.4}$$

[9] See also Heckman and Honore (1990).

where $\boldsymbol{\gamma} = \boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_0$ and $\eta = \nu_1 - \nu_0$. The key consideration is that the econometrician observes the subjective (or self-reported) well-being response Y_1 if $I \geq 0$ and he observes Y_0 if $I < 0$. The intuition is as follows. For a moment, let's consider the job satisfaction example. The interviewee takes the interview on a Friday or Saturday ($D = 1$) rather than the other days if she receives higher utility from doing so. This higher utility (i.e., $U_1 > U_0$) is translated into the expression $I \geq 0$ and, accordingly, Y_1 is observed. The utilities are not observed; but, what the econometrician observes are a choice and an associated well-being outcome. The observed subjective well-being outcome, in this setup, is

$$Y = (1 - D)Y_0 + DY_1, \tag{A.5}$$

which means that $Y = Y_0$ if $D = 0$ and $Y = Y_1$ if $D = 1$. Y_1 is observed when $U_1 \geq U_0$ and Y_0 is observed otherwise. The main lesson that this formulation communicates is the following. The day of the week on which the interviewee takes the interview is a matter of choice. There are both observed and unobserved factors that may be affecting this choice. Accounting for unobservables may change the results reported in the literature. This formulation aims at explicitly controlling for unobserved determinants of the day of the week.

To map this formulation to data, we formulate two outcome equations:

$$Y_0 = \boldsymbol{X}\boldsymbol{\beta}_0' + \epsilon_0, \tag{A.6}$$

$$Y_1 = \boldsymbol{X}\boldsymbol{\beta}_1' + \epsilon_1, \tag{A.7}$$

where $\boldsymbol{X}$ is a row-vector of observed covariates. We assume that $(\boldsymbol{X}, \boldsymbol{Z}) \perp\!\!\!\perp (\eta, \epsilon_1, \epsilon_0)$, where $\perp\!\!\!\perp$ denotes statistical independence. We also assume that the error terms are jointly normally distributed as $(\eta, \epsilon_1, \epsilon_0) \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is the covariance matrix and can be written as

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{\eta\eta} & \sigma_{\eta 1} & \sigma_{\eta 0} \\ \sigma_{\eta 1} & \sigma_{11} & \sigma_{10} \\ \sigma_{\eta 0} & \sigma_{10} & \sigma_{00} \end{bmatrix}. \tag{A.8}$$

Note that from $\eta = \epsilon_1 - \epsilon_0$, it is easy to show that $\sigma_{\eta\eta} = \sigma_{11} + \sigma_{00} - 2\sigma_{10}$, $\sigma_{\eta 1} = \sigma_{11} - \sigma_{10}$, and $\sigma_{\eta 0} = \sigma_{10} - \sigma_{00}$.

As we explain above, $D = \mathbb{1}(I \geq 0)$, where $\mathbb{1}$ stands for an indicator function. From data on Y, D, and $(\boldsymbol{X}, \boldsymbol{Z})$, the following quantities can be obtained:

$$\mathbb{P}[D = 1|\boldsymbol{Z}], \;\; \mathbb{E}[Y|D = 1, \boldsymbol{X}, \boldsymbol{Z}], \;\; \text{and} \;\; \mathbb{E}[Y|D = 0, \boldsymbol{X}, \boldsymbol{Z}].$$

One key issue is the distinction between $\boldsymbol{Z}$ and $\boldsymbol{X}$. For identification purposes, we assume that these two data vectors overlap except one extra variable in $\boldsymbol{Z}$; that is, $\dim(\boldsymbol{Z}) = \dim(\boldsymbol{X}) + 1$, where the notation "dim" describes the dimension of a data vector. In other words, we need an extra variable that affects the choice of the agent, but does not affect the outcome of interest. This is known in the literature as an "exclusion restriction" (or an instrument) that secures identification in selection-correction models. See Section 3.2 for a comprehensive discussion of this issue as well as the details of the specific exclusion restriction that we use in this paper.

A.2 Selection Correction

We start with the following Probit regression, which is the typical first step in a selection-correction procedure:

$$\begin{aligned}
\mathbb{P}\big[D = 1\big|\boldsymbol{Z} = \boldsymbol{z}\big] &= \mathbb{P}\big[\boldsymbol{Z}\boldsymbol{\gamma}' + \eta \geq 0\big|\boldsymbol{Z} = \boldsymbol{z}\big] \\
&= \mathbb{P}\big[\boldsymbol{z}\boldsymbol{\gamma}' + \eta \geq 0\big] \\
&= \mathbb{P}\left[\frac{\eta}{\sigma_\eta} \geq -\frac{\boldsymbol{z}\boldsymbol{\gamma}'}{\sigma_\eta}\right] \\
&= \Phi\left(\frac{\boldsymbol{z}\boldsymbol{\gamma}'}{\sigma_\eta}\right).
\end{aligned} \tag{A.9}$$

This probit equation identifies $\boldsymbol{\gamma}/\sigma_\eta$, where $\sigma_\eta = \sqrt{\sigma_{\eta\eta}}$. Now we consider the regression equations related to the two outcome equations. The first outcome equation gives

$$\begin{aligned}\mathbb{E}\big[Y\big|D=1,\boldsymbol{X}=\boldsymbol{x},\boldsymbol{Z}=\boldsymbol{z}\big] &= \mathbb{E}\big[Y_1\big|\boldsymbol{X}=\boldsymbol{x},\boldsymbol{Z}=\boldsymbol{z}\big] \\ &= \boldsymbol{x}\boldsymbol{\beta}_1' + \mathbb{E}\big[\epsilon_1\big|\boldsymbol{z}\boldsymbol{\gamma}'+\eta \geq 0\big] \\ &= \boldsymbol{x}\boldsymbol{\beta}_1' + \frac{\sigma_{\eta 1}}{\sigma_\eta}\lambda\left(-\frac{\boldsymbol{z}\boldsymbol{\gamma}'}{\sigma_\eta}\right) \qquad \text{(A.10)}\end{aligned}$$

and the second outcome equation gives

$$\begin{aligned}\mathbb{E}[Y|D=0,\boldsymbol{X}=\boldsymbol{x},\boldsymbol{Z}=\boldsymbol{z}] &= \mathbb{E}[Y_0|\boldsymbol{X}=\boldsymbol{x},\boldsymbol{Z}=\boldsymbol{z}] \\ &= \boldsymbol{x}\boldsymbol{\beta}_0' + \mathbb{E}[\epsilon_0|\boldsymbol{z}\boldsymbol{\gamma}'+\eta < 0] \\ &= \boldsymbol{x}\boldsymbol{\beta}_0' - \frac{\sigma_{\eta 0}}{\sigma_\eta}\lambda\left(\frac{\boldsymbol{z}\boldsymbol{\gamma}'}{\sigma_\eta}\right), \qquad \text{(A.11)}\end{aligned}$$

where $\lambda(\cdot)$ is the inverse Mills ratio and, as a general rule, $\lambda(c) = \phi(c)/\Phi(-c)$.

From the probit regression in (A.9), we already know the parameter $\boldsymbol{\gamma}/\sigma_\eta$. Therefore, we can form $\lambda\left(-\frac{\boldsymbol{z}\boldsymbol{\gamma}'}{\sigma_\eta}\right)$ and $\lambda\left(\frac{\boldsymbol{z}\boldsymbol{\gamma}'}{\sigma_\eta}\right)$. Equations (A.10) and (A.11) suggest that we can run regressions of Y_1 on $\boldsymbol{X}$ and $\lambda\left(-\frac{\boldsymbol{z}\boldsymbol{\gamma}'}{\sigma_\eta}\right)$, and of Y_0 on $\boldsymbol{X}$ and $\lambda\left(\frac{\boldsymbol{z}\boldsymbol{\gamma}'}{\sigma_\eta}\right)$ to identify $\boldsymbol{\beta}_0$, $\boldsymbol{\beta}_1$, $\sigma_{\eta 0}/\sigma_\eta$, and $\sigma_{\eta 1}/\sigma_\eta$.[10]

A.3 Treatment Effects

In our context, "treatment" refers to taking the interview on a Friday or Saturday for job satisfaction analysis and Sunday or Monday for happiness analysis (i.e., $D = 1$). Obtaining the treatment effect estimates would be useful for our analysis, since it will enhance our understanding of the existence, extent, and the sources of the selection structure. Calculation of the treatment effects are simple and straightforward after obtaining the bias corrected estimates described in the previous subsection. The most commonly sought treatment effect

[10] Identification of the other parameters is also possible. But, in this paper, we are not interested in the magnitudes of the rest of the parameters. See Heckman and Honore (1990) for the details. See also Heckman and Robb (1985) and Heckman and Vytlacil (2007a,b).

parameter is the Average Treatment Effect (ATE). It can simply be formulated as

$$\begin{aligned}\mathrm{ATE}(\boldsymbol{x}) &= \mathbb{E}[Y_1 - Y_0 | \boldsymbol{X} = \boldsymbol{x}] \\ &= \boldsymbol{x}(\boldsymbol{\beta}_1' - \boldsymbol{\beta}_0'). \end{aligned} \tag{A.12}$$

This can be interpreted as the effect of randomly assigning $D = 1$ to everyone with $\boldsymbol{X} = \boldsymbol{x}$. The main problem with this parameter is analogous to the central question that we address in this paper; that is, it may not reflect a true causal effect of $D = 1$ versus $D = 0$ on the subjects, because the ones who choose $D = 1$ maybe systematically different from the ones who choose $D = 0$.[11] This difference leads the evaluation of the outcome at the counterfactual states to be biased.

The other two treatment effect parameters that we calculate in this study are the treatment on the treated (TT) and the treatment on the untreated (TUT). These parameters can nicely be formulated as a function of the control functions estimated during the implementation of the selection-correction procedure [see Heckman and Vytlacil (2007a,b) for details]. The parameter TT can be formulated as

$$\begin{aligned}\mathrm{TT}(\boldsymbol{x}, p_{\boldsymbol{z}}) &= \mathbb{E}[Y_1 - Y_0 | D = 1, \boldsymbol{X} = \boldsymbol{x}, \boldsymbol{Z} = \boldsymbol{z}] \\ &= \boldsymbol{x}(\boldsymbol{\beta}_1' - \boldsymbol{\beta}_0') + \frac{\sigma_{\eta 1} - \sigma_{\eta 0}}{\sigma_\eta} \frac{\phi(\Phi^{-1}(p_{\boldsymbol{z}}))}{p_{\boldsymbol{z}}}, \end{aligned} \tag{A.13}$$

while TUT can be formulated as

$$\begin{aligned}\mathrm{TUT}(\boldsymbol{x}, p_{\boldsymbol{z}}) &= \mathbb{E}[Y_1 - Y_0 | D = 0, \boldsymbol{X} = \boldsymbol{x}, \boldsymbol{Z} = \boldsymbol{z}] \\ &= \boldsymbol{x}(\boldsymbol{\beta}_1' - \boldsymbol{\beta}_0') + \frac{\sigma_{\eta 0} - \sigma_{\eta 1}}{\sigma_\eta} \frac{\phi(\Phi^{-1}(p_{\boldsymbol{z}}))}{1 - p_{\boldsymbol{z}}}, \end{aligned} \tag{A.14}$$

where $p_{\boldsymbol{z}}$ refers to the propensity score estimated in the first stage probit regression. The average TT is the average gain for those who sort into treatment compared to what the average person would gain. It oversamples the unobserved characteristics that lead to selectivity for

[11] Remember that in our case $D = 1$ refers to taking the interview on a Friday or Saturday versus the remaining days for the job satisfaction analysis and on a Sunday or Monday versus the remaining days for the happiness analysis.

those individuals who are more likely to choose $D = 1$. In other words, it calculates the net effect between those who actually participate and those who do not, as if they had given the chance to revert their choice of $D = 0$ into $D = 1$. A symmetric definition can be provided for TUT, oversamples the unobserved characteristics that lead to selectivity for those individuals who are less likely to choose $D = 1$.

For the purposes of this paper, we are interested in the "averages" of these three treatment effect parameters. In other words, the estimates reported at the end of the paper are the parameter estimates integrated over the entire horizon of $\boldsymbol{x}$ and $\boldsymbol{z}$ in our sample. It is also possible to report the distribution of these treatment effects over the sample space. But, we report only the means to keep the paper as compact as possible.

Notice that when the coefficient of the inverse Mills ratio calculated at the second stage is zero, then the TT and TUT collapses into ATE. This is the case with no selectivity. When there is positive sorting into the treatment state (as in our case), on the other hand, the econometrician would find TT $>$ ATE $>$ TUT. Moreover, it is easy to verify that ATE is a weighted average of the TT and TUT. In Section 4, we use these formulas and calculate the treatment effect parameters for both the job satisfaction and happiness scores.

Summary Statistics				
Variable	Mean	Standard Deviation	Min.	Max.
Job satisfaction	5.383	1.296	1	7
Happiness	22.748	5.073	12	48
Male	0.497	0.500	0	1
Age	38.749	12.900	16	85
Married	0.563	0.496	0	1
Never married	0.328	0.470	0	1
Higher degree	0.029	0.168	0	1
First degree	0.123	0.328	0	1
‘A’-levels	0.132	0.339	0	1
‘O’-levels	0.212	0.409	0	1
Other higher qual.	0.262	0.440	0	1
Vocational qual.	0.116	0.320	0	1
No degree	0.127	0.332	0	1
Temporary worker	0.029	0.167	0	1
Fixed-term contract	0.017	0.129	0	1
Public sector worker	0.170	0.376	0	1
Small employer	0.691	0.462	0	1
Promotion opp.	0.405	0.491	0	1
Union member	0.239	0.426	0	1
Health very good	0.250	0.433	0	1
Health very satisfactory	0.159	0.366	0	1
Prefers to work more	0.080	0.271	0	1
Prefers to work less	0.314	0.464	0	1
Income	0.524	0.499	0	1
Fri/Sat	0.189	0.392	0	1
Sun/Mon	0.247	0.431	0	1

Table 1: **Summary statistics.** This table roughly summarizes the data we use. We focus on employed individuals in the BHPS data covering the period 1992–2008. Appropriate sampling weights are used.

Dependent variable	Job Satisfaction		Happiness	
Variable	Coefficient	(Standard Error)	Coefficient	(Standard Error)
Friday	0.0340***	(0.0020)	-0.0405***	(0.0076)
Saturday	0.0221***	(0.0022)	-0.0208**	(0.0086)
Sunday	-0.0041	(0.0027)	0.0709***	(0.0106)
Monday	0.0032**	(0.0017)	0.1008***	(0.0066)
Tuesday	0.0191***	(0.0017)	0.0683***	(0.0065)
Wednesday	-0.0072***	(0.0017)	-0.0095	(0.0065)
Thursday	Omitted		Omitted	
Constant	5.367***	(0.0012)	22.719***	(0.0048)
# of observations	68,773		68,504	
R^2	0.0246		0.0231	

Table 2: **Day orderings.** This table presents the results of an OLS regression of the subjective well-being score on the days of the week. Thursday is the omitted dummy variable; that is, the results should be read with respect to Thursday. *, **, *** indicate the 10%, 5%, and 1% significance levels, respectively. Appropriate sampling weights are used. Robust standard errors are reported.

Dependent variable	Job Satisfaction		Happiness	
Variable	Coefficient	(Standard Error)	Coefficient	(Standard Error)
Fri/Sat	0.0255***	(0.0013)	-	-
Sun/Mon	-	-	0.0879***	(0.0047)
Constant	5.371***	(0.0006)	22.726***	(0.0023)
# of observations	68,773		68,504	
R^2	0.0211		0.0202	

Table 3: **Bunching the days.** This table repeats the exercise above by regressing the job satisfaction (happiness) score on the Fri/Sat (Sun/Mon) dummy. The Fri/Sat (Sun/Mon) dummy indicates if the interview is conducted on a Friday or Saturday (Sunday or Monday). *, **, *** indicate the 10%, 5%, and 1% significance levels, respectively. Appropriate sampling weights are used. Robust standard errors are reported.

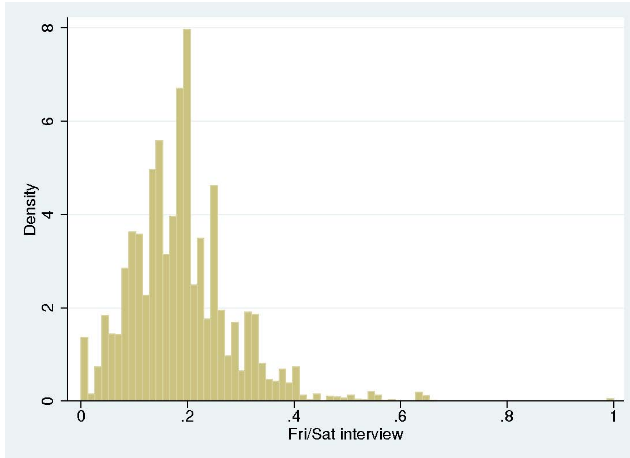

Figure 1: **Interviewer ID (job satisfaction).** This figure presents the distributional features of the interviewer ID variable that we use at the probit regression for job satisfaction score. The horizontal axis describes the probability for a specific interviewer to conduct the interview on a Friday or Saturday. For example, a value of 0.4 for interviewer j means that the interviewer j conducted 40% of his/her interviews on a Friday or Saturday.

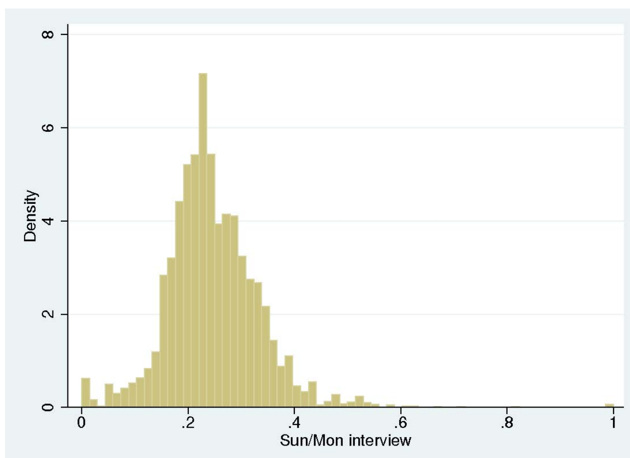

Figure 2: **Interviewer ID (happiness).** This figure presents the distributional features of the interviewer ID variable that we use at the probit regression for the general happiness score. The horizontal axis describes the probability for a specific interviewer to conduct the interview on a Sunday or Monday. For example, a value of 0.4 for interviewer j means that the interviewer j conducted 40% of his/her interviews on a Sunday or Monday.

Dependent var.	**Job Satisfaction**		**Happiness**	
Variable	Coefficient	(Standard Error)	Coefficient	(Standard Error)
Fri/Sat	0.032***	(0.0012)	-	-
Sun/Mon	-	-	0.097***	(0.0045)
Male	-0.241***	(0.0011)	-1.178***	(0.0044)
Age	-0.038***	(0.0003)	0.205***	(0.0010)
$\text{Age}^2/100$	0.052***	(0.0003)	-0.243***	(0.0012)
Married	0.066***	(0.0016)	-0.660***	(0.0072)
Never married	-0.141***	(0.0020)	-0.511***	(0.0088)
Higher degree	-0.133***	(0.0031)	0.354***	(0.0135)
First degree	-0.242***	(0.0022)	0.323***	(0.0086)
'A'-levels	-0.189***	(0.0021)	0.109***	(0.0079)
'O'-levels	-0.088***	(0.0019)	-0.021***	(0.0069)
Other higher qual.	-0.139***	(0.0019)	0.129***	(0.0070)
Vocational qual.	-0.073***	(0.0021)	-0.111***	(0.0078)
Temporary worker	-0.125***	(0.0042)	-0.026**	(0.0122)
Fixed-term contract	-0.052***	(0.0044)	-0.283***	(0.0163)
Public sector worker	-0.001	(0.0015)	-0.012*	(0.0062)
Small employer	0.149***	(0.0011)	-0.010**	(0.0043)
Promotion opp.	0.319***	(0.0011)	-0.577***	(0.0044)
Union member	-0.188***	(0.0013)	0.339***	(0.0052)
Health very good	0.225***	(0.0011)	-1.665***	(0.0043)
Health very satisfactory	-0.166***	(0.0015)	1.387***	(0.0060)
Prefers to work more	-0.241***	(0.0021)	0.685***	(0.0080)
Prefers to work less	-0.518***	(0.0011)	0.831***	(0.0044)
Income	0.063***	(0.0013)	-0.170***	(0.0049)
Year dummies	Yes		Yes	
Industry dummies	Yes		Yes	
Region dummies	Yes		Yes	
Constant	6.21***	(0.0098)	20.162***	(0.0378)
# of observations	68,773		68,504	
R^2	0.0921		0.0813	

Table 4: **Day patterns conditional on observed variation.** This table repeats the exercise in Table (3) by controlling for a comprehensive set of observed worker- and job-related characteristics. *, **, *** indicate the 10%, 5%, and 1% significance levels, respectively. Appropriate sampling weights are used. Robust standard errors are reported.

Dependent var.	Fri/Sat		Sun/Mon	
Variable	Coefficient	(Standard Error)	Coefficient	(Standard Error)
Male	0.029***	(0.0013)	0.009***	(0.0012)
Age	0.017***	(0.0003)	0.002***	(0.0003)
$\text{Age}^2/100$	-0.019***	(0.0004)	-0.002***	(0.0003)
Married	-0.045***	(0.0019)	0.008***	(0.0018)
Never married	-0.032***	(0.0023)	0.034***	(0.0023)
Higher degree	0.154***	(0.0038)	0.061***	(0.0036)
First degree	0.112***	(0.0026)	0.040***	(0.0024)
'A'-levels	0.062***	(0.0025)	-0.003	(0.0023)
'O'-levels	0.060***	(0.0022)	0.029***	(0.0020)
Other higher qual.	0.057***	(0.0022)	0.011***	(0.0020)
Vocational qual.	0.050***	(0.0024)	0.025***	(0.0022)
Temporary worker	0.042***	(0.0044)	-0.050***	(0.0034)
Fixed-term contract	-0.010**	(0.0049)	-0.088***	(0.0044)
Public sector worker	0.004**	(0.0018)	-0.039***	(0.0017)
Small employer	0.004***	(0.0013)	0.007***	(0.0012)
Promotion opp.	0.017***	(0.0013)	0.002	(0.0012)
Union member	0.015***	(0.0015)	-0.019***	(0.0015)
Health very good	0.003*	(0.0014)	0.012***	(0.0013)
Health very satisfactory	-0.013***	(0.0016)	0.026***	(0.0016)
Prefers to work more	-0.027***	(0.0023)	0.018***	(0.0021)
Prefers to work less	-0.007***	(0.0013)	0.006***	(0.0012)
Income	0.011***	(0.0015)	0.019***	(0.0014)
Interviewer dummy	0.524***	(0.0004)	0.410***	(0.0011)
Year dummies	Yes		Yes	
Industry dummies	Yes		Yes	
Region dummies	Yes		Yes	
Constant	-1.482***	(0.0118)	-0.904***	(0.0104)
# of observations	68,773		68,504	
Pseudo R^2	0.0380		0.0215	

Table 5: **Probit regression.** This table documents the results of the probit regression of the day selection of the interviewee on a set of observed characteristics and the interviewer dummy. The interviewer dummy takes the value 1 if the interviewer is more likely to conduct the interview—than the average tendency—on a Friday or Saturday (Sunday or Monday) in the job satisfaction (happiness) analysis. *, **, *** indicate the 10%, 5%, and 1% significance levels, respectively. Appropriate sampling weights are used. Robust standard errors are reported.

Dependent variable: Job Satisfaction				
	(Y_1) Fri/Sat=1		(Y_0) Fri/Sat=0	
Variable	Coefficient	(Standard Error)	Coefficient	(Standard Error)
Male	-0.204***	(0.0026)	-0.248***	(0.0013)
Age	-0.029***	(0.0007)	-0.038***	(0.0003)
$Age^2/100$	0.041***	(0.0008)	0.053***	(0.0003)
Married	0.025***	(0.0036)	0.075***	(0.0018)
Never married	-0.150***	(0.0044)	-0.141***	(0.0023)
Higher degree	-0.076***	(0.0070)	-0.144***	(0.0035)
First degree	-0.241***	(0.0050)	-0.239***	(0.0025)
'A'-levels	-0.173***	(0.0050)	-0.189***	(0.0023)
'O'-levels	-0.095***	(0.0044)	-0.083***	(0.0021)
Other higher qual.	-0.171***	(0.0044)	-0.129***	(0.0020)
Vocational qual.	-0.093***	(0.0049)	-0.067***	(0.0023)
Temporary worker	-0.196***	(0.0099)	-0.107***	(0.0046)
Fixed-term contract	-0.011	(0.0099)	-0.061***	(0.0049)
Public sector worker	0.014***	(0.0034)	-0.005***	(0.0016)
Small employer	0.122***	(0.0025)	0.156***	(0.0012)
Promotion opp.	0.340***	(0.0025)	0.313***	(0.0012)
Union member	-0.167***	(0.0030)	-0.193***	(0.0014)
Health very good	0.244***	(0.0026)	0.219***	(0.0013)
Health very satisfactory	-0.161***	(0.0033)	-0.168***	(0.0016)
Prefers to work more	-0.286***	(0.0049)	-0.229***	(0.0023)
Prefers to work less	-0.494***	(0.0026)	-0.524***	(0.0013)
Inverse Mills Ratio	0.0415***	(0.0059)	-0.0797***	(0.0054)
Year dummies	Yes		Yes	
Industry dummies	Yes		Yes	
Region dummies	Yes		Yes	
Constant	6.192***	(0.0258)	6.240***	(0.0107)
# of observations	12,901		55,872	
R^2	0.0845		0.0951	

Table 6: **Second step (job satisfaction)**. This table presents the results of the second step OLS regression of the job satisfaction score on a set of observed covariates (excluding the interviewer dummy) and the inverse Mills ratio calculated from the results of the first step probit regression, which are given in Table (5). *, **, *** indicate the 10%, 5%, and 1% significance levels, respectively. Appropriate sampling weights are used. Robust standard errors are reported.

Dependent variable: Happiness				
	(Y_1) Sun/Mon=1		(Y_0) Sun/Mon=0	
Variable	Coefficient	(Standard Error)	Coefficient	(Standard Error)
Male	-1.178***	(0.0089)	-1.173***	(0.0051)
Age	0.217***	(0.0020)	0.201***	(0.0012)
$Age^2/100$	-0.261***	(0.0024)	-0.237***	(0.0013)
Married	-0.708***	(0.0148)	-0.641***	(0.0082)
Never married	-0.546***	(0.0180)	-0.490***	(0.0101)
Higher degree	0.766***	(0.0274)	0.222***	(0.0155)
First degree	0.318***	(0.0177)	0.336***	(0.0098)
'A'-levels	0.091***	(0.0162)	0.113***	(0.0090)
'O'-levels	0.004	(0.0145)	-0.025***	(0.0079)
Other higher qual.	0.039***	(0.0145)	0.164***	(0.0080)
Vocational qual.	-0.264***	(0.0161)	-0.054***	(0.0089)
Temporary worker	-0.258***	(0.0241)	-0.218***	(0.0186)
Fixed-term contract	-0.524***	(0.0340)	-0.061***	(0.0049)
Public sector worker	0.149***	(0.0130)	-0.072***	(0.0070)
Small employer	0.081***	(0.0088)	-0.039***	(0.0050)
Promotion opp.	-0.519***	(0.0090)	-0.593***	(0.0050)
Union member	0.400***	(0.0106)	0.317***	(0.0059)
Health very good	-1.665***	(0.0088)	-1.664***	(0.0050)
Health very satisfactory	1.474***	(0.0122)	1.361***	(0.0069)
Prefers to work more	0.472***	(0.0166)	0.756***	(0.0091)
Prefers to work less	0.723***	(0.0088)	0.870***	(0.0050)
Income	-0.271***	(0.0099)	-0.137***	(0.0056)
Inverse Mills Ratio	0.102***	(0.0260)	-0.549***	(0.0240)
Year dummies	Yes		Yes	
Industry dummies	Yes		Yes	
Region dummies	Yes		Yes	
Constant	20.378***	(0.0782)	20.458***	(0.0461)
# of observations	16,972		51,532	
R^2	0.0851		0.0812	

Table 7: **Second step (happiness).** This table presents the results of the second step OLS regression of the general happiness score on a set of observed covariates (excluding the interviewer dummy) and the inverse Mills ratio calculated from the results of the first step probit regression, which are given in Table (5). *, **, *** indicate the 10%, 5%, and 1% significance levels, respectively. Appropriate sampling weights are used. Robust standard errors are reported.

Treatment Effects		
	Job Satisfaction	Happiness
Aggregate		
ATE	0.033	0.094
TT	0.211	0.936
TUT	-0.007	-0.179

Table 8: **Estimated treatment effects.** This table documents the treatment effects estimates for job satisfaction and happiness. ATE refers to the Average Treatment Effect, TT refers to the (average) Treatment on the Treated, and TUT refers to the (average) Treatment on the Untreated. Appropriate sampling weights are used.

Treatment Effects for Education Categories		
	Job Satisfaction	Happiness
Higher degree		
ATE	0.143	0.660
TT	0.311	1.473
TUT	0.098	0.369
First degree		
ATE	0.068	0.124
TT	0.240	0.958
TUT	0.025	-0.153
A-levels		
ATE	0.069	0.117
TT	0.247	0.972
TUT	0.029	-0.148
O-levels		
ATE	0.028	0.169
TT	0.205	1.110
TUT	-0.013	-0.104
Other higher qual.		
ATE	0.012	-0.013
TT	0.190	0.830
TUT	-0.028	-0.285
Vocational qual.		
ATE	0.014	-0.063
TT	0.186	0.781
TUT	-0.023	-0.334
No qual.		
ATE	0.014	0.146
TT	0.198	0.991
TUT	-0.023	-0.125

Table 9: **Estimated treatment effects for education categories.** This table documents the treatment effect estimates for job satisfaction and happiness in different education categories. Appropriate sampling weights are used.

Treatment Effects for sub-groups				
	Job Satisfaction	Happiness	Job Satisfaction	Happiness
	Male		Female	
ATE	0.054	0.055	0.012	0.132
TT	0.230	0.894	0.191	0.978
TUT	0.013	-0.220	-0.027	-0.138
	Married		Non-married	
ATE	0.015	0.080	0.058	0.110
TT	0.192	0.925	0.236	0.950
TUT	-0.025	-0.190	0.018	-0.164
	Permanent		Temporary	
ATE	0.035	0.101	-0.073	-0.150
TT	0.212	0.942	0.101	0.715
TUT	-0.005	-0.172	-0.115	-0.409
	Public sector		Private sector	
ATE	0.067	0.326	0.026	0.046
TT	0.244	1.188	0.203	0.884
TUT	0.026	0.065	-0.014	-0.229
	Small employer		Large employer	
ATE	0.016	0.124	0.071	0.026
TT	0.194	0.965	0.247	0.871
TUT	-0.024	-0.149	0.030	-0.244
	Union worker		Non-union worker	
ATE	0.073	0.206	0.020	0.058
TT	0.250	1.056	0.197	0.899
TUT	0.033	-0.062	-0.020	-0.215
	Health very good		Health satisfactory	
ATE	0.059	0.063	0.033	0.162
TT	0.236	0.900	0.212	0.994
TUT	0.018	-0.212	-0.007	-0.117
	Prefers to work more		Prefers to work less	
ATE	-0.031	-0.087	0.060	0.005
TT	0.150	0.756	0.237	0.844
TUT	-0.069	-0.359	0.019	-0.269
	Higher relative income		Lower relative income	
ATE	0.052	0.058	0.010	0.132
TT	0.226	0.897	0.191	0.979
TUT	0.010	-0.216	-0.029	-0.137

Table 10: **Estimated treatment effects for sub-groups.** This table documents the treatment effect estimates for job satisfaction and happiness in certain sub-groups determined based on worker- and job-related characteristics. Appropriate sampling weights are used.

Figure 3: **Treatment effects by age categories (job satisfaction).** This figure presents the estimated ATE, TT, and TUT categories for age groups in the job satisfaction analysis. Ten age categories are used. Appropriate sampling weights are used.

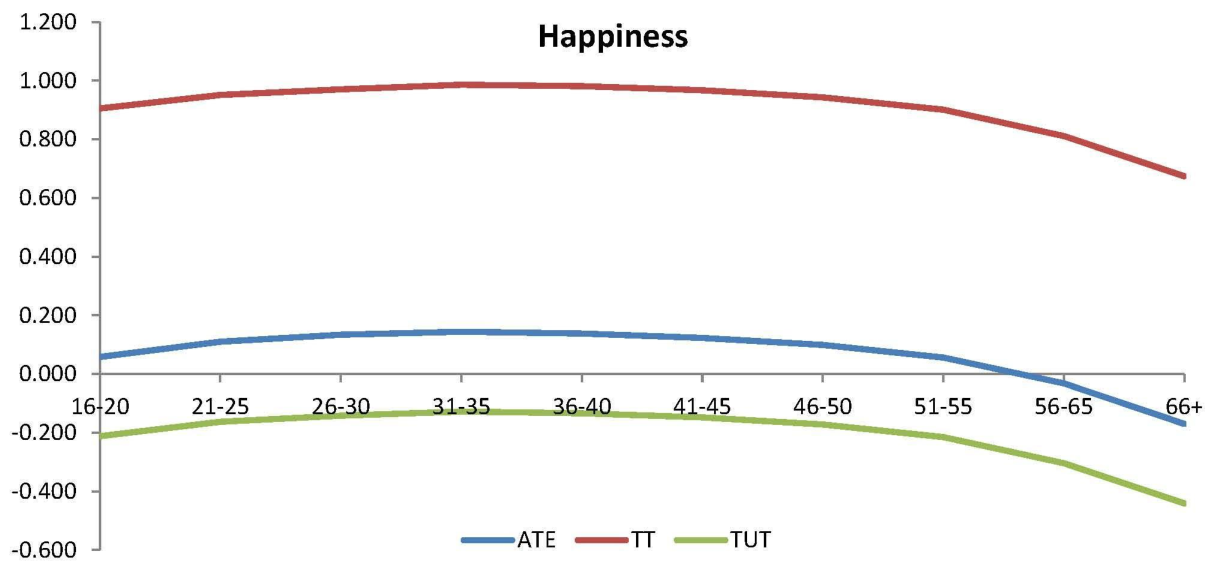

Figure 4: **Treatment effects by age categories (happiness).** This figure presents the estimated ATE, TT, and TUT categories for age groups in the happiness analysis. Ten age categories are used. Appropriate sampling weights are used.