OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 0305-9049
doi: 10.1111/obes.12446

A Multiple Cohort Study of the Gender Gradient of Life Satisfaction during Adolescence: Longitudinal Evidence from Great Britain*

GEORGIOS MARIOS CHRYSANTHOU

School of Mathematics and Statistics, University of Sheffield, Sheffield S3 7RH, UK (e-mail: G.Chrysanthou@sheffield.ac.uk)

Abstract

This study is unique in exploiting 12 youth cohorts (aged 11–15) from the British Household Panel Survey (BHPS) and the UK Household Longitudinal Study (UKHLS) spanning 1996–2017 to investigate the gender gradient of adolescent life satisfaction. We find robust evidence of a cross-cohort gender gap particularly at the extremes of the adolescent life satisfaction distribution. Male adolescents are significantly more likely to report complete life satisfaction (by around 6%–14%) and females to report dissatisfaction (by around 3%–7%) indicating a higher female depression propensity. An intra-household gender gap is found for female adolescents raised with opposite sex siblings. Previous period life satisfaction is the strongest determinant of prospective higher self-reported male satisfaction levels.

I. Introduction

Adolescence is a critical transitional stage of psychological and physical development spanning the period from puberty to legal adulthood. During this transformational period of maturation, emotional and behavioural difficulties may be frequent with consequent adverse effects on future adult outcomes, human capital formation, cognition and mental health (see Heckman, 2006; Cunha, Heckman and Schennach, 2010; Layard *et al*., 2014; Van den Berg *et al*., 2014).

The onset of anxiety and depressive disorders peaks during adolescence and early adulthood, with girls facing significantly higher risks and twice the lifetime rates of depression and the majority of anxiety disorders (Angold and Worthman, 1993; Van de Velde, Bracke and Levecque, 2010; Altemus, Sarvaiya and Epperson, 2014, p. 320). Higher adult female life satisfaction is reported by some economic studies such as Blanchflower and Oswald (2004), UK and US data (1972–98), Blanchflower and Oswald (2008), European data (2001–02), and Oswald and Wu (2011) using US

JEL Classification numbers: C25. I31. J13.

*I gratefully acknowledge partial financial support from the Spanish Ministerio de Economía y Competitividad through the research project ECO2016-77200-P.

data (2005–08). At the same time, Blanchflower and Oswald (2008) find higher female anxiety, depression (UK data, 2004–07) and mental distress (European data) as do Oswald and Wu (2011) using US data. Stevenson and Wolfers (2009) using US GSS data (1972–06) and Eurobarometer data (1970–02) conclude that female happiness fell relative to male. These studies use pooled adult cross-sections not following the same individuals longitudinally and can only provide estimates of the mean life satisfaction trend across time (see Bond and Lang, 2019). Chrysanthou and Vasilakis (2019) find a positive male life satisfaction gradient using a single cohort (2009–13) from the UKHLS youth sample. It is therefore of great interest to investigate the gender gradient of life satisfaction employing multiple cohorts of longitudinal adolescent samples spanning two decades.

This study is unique in that it exploits all possible youth cohort formations from the BHPS spanning 1996–2008 and the UKHLS spanning 2011–17. This gives 12 cohorts of adolescents aged 11–15 employed to investigate the gender gradient and socioeconomic determinants of life satisfaction. The investigation provides a quantitative comparison of the gender gradient across different cohorts and levels of life satisfaction. Potential mechanisms driving gender differences in self-reported life satisfaction are explored by considering intra-household gender differentials and allowing control variables to play distinct predictive roles per gender. Furthermore, the life satisfaction gender gradient trend across the duration of each of the two datasets is also analysed.

Due to numerous discrepancies in brain structure, stress responsivity, hormonal/genetic sexual dimorphisms and social influences, identifying the causes of gender differences relevant to subjective wellbeing, affective illness and mental health is challenging (Altemus *et al*., 2014, p.320; Rainville, Tsyglakova and Hodes, 2018, p. 79). Genetic, epigenetic and biomarker data have not been generally collected in longitudinal surveys nor in the BHPS and UKHLS youth samples used in this study. Given the unavailability of pertinent biomedical data, controlling for individual-specific latent characteristics is essential if one wishes to isolate gender differences in self-reported life satisfaction.[1]

Longitudinal data permit controlling for latent individual time-invariant characteristics (unobserved heterogeneity) such as personality/genetics preconditioning life satisfaction (e.g. pessimists/less attractive individuals might be more prone to report lower levels of satisfaction – see Jones and Schurer, 2011; Angelini *et al*., 2014). On the other hand, past adverse events or personal experiences may trigger a self-perpetuating cycle of low life satisfaction levels (state dependence) via negative thoughts and emotions. To disentangle state dependence from unobserved heterogeneity we control for the initial reported values of life satisfaction (initial conditions).

Employing both the BHPS and the UKHLS, we provide robust cross-cohort evidence that female adolescents are more likely to report dissatisfaction and less likely to report complete life satisfaction. The gender gradient holds for all cohorts spanning 1996–2017 and is robust to a series of robustness checks using alternative model specifications. Additionally, an intra-household gender gap is also present for females

[1]The UKHLS collects biomedical measures from adults.

growing up with male siblings. This could indicate parental ingrained gender stereotypes preconditioning child nurturing. Furthermore, there is significant state dependence in life satisfaction. Consequently, early-life interventions are necessary to curb increases in depression prevalence. We find a gender differential impact of past period life satisfaction: past satisfaction levels are the strongest determinant of prospective self-reported adolescent male satisfaction levels. However, to investigate whether this is due to gender-specific norms producing distinct interpretations of life satisfaction scales a vignette methodology would be appropriate (Hsee and Tang, 2007; Angelini *et al*., 2014). The BHPS and UKHLS do not include vignettes and youth respondents only rate their personal level of life satisfaction. Controlling for household income per capita, living in a high income region reduces self-reported life satisfaction as in Oswald and Wu (2011).

The paper is organized as follows. Section II discusses the sample selection mechanism and variable-related issues. Section III analyses the cross-cohort longitudinal evolution of life satisfaction. Section IV introduces the estimation methodology, section V reports and analyses the estimation results and section VI concludes. An extensive online Appendix provides descriptive statistics and alternative estimations.

II. Sample selection and variables

We employ the British Household Panel Survey (BHPS) and the Understanding Society (UK Household Longitudinal Study, UKHLS) datasets to construct compact unbalanced panels of successive cohorts of youth respondents aged 11–15, living in Great Britain during the period 1996–2017. The BHPS is a nationally representative survey introduced in 1991 and was initially addressed to adult members. The first wave panel consisted of some 5,500 households and 10,300 individual respondents selected from 250 areas of Great Britain. Since 1994 children aged 11–15 years are also asked to complete a self-completion Youth Questionnaire. The UKHLS was introduced in 2009 (a year after the final BHPS wave in 2008) and interviewed around 40,000 households. All young people aged 10–15 living in UKHLS participating households, complete their own self-completion Youth Questionnaire. For comparability with the BHPS and to ensure that the cohort initiation period coincides with the onset of puberty, we consider youth respondents aged 11–15 in the UKHLS and omit those aged 10 in the corresponding cohort initiation period.[2]

We homogenize the samples such that individuals are 11–13 years old in the corresponding cohort initiation period. The average age of onset of puberty is at 11 for girls and 12 for boys and inclusion of respondents aged 14 at the year of cohort commencement can condition the estimates. Also note that the biological components of pubertal development can exacerbate the negative experiences (mood variability and

[2]The first half of the age range studied corresponds to preadolescence and the other half to adolescence though this is an approximate framework of young developmental periods and varies across individuals. Slightly abusing the terminology, we henceforth refer to our investigation as an adolescent life satisfaction study. Note that, a small number of youth respondents turn 16 in their final youth sample membership year.

negative emotions) of pubertally advanced adolescents (see Hamilton *et al.*, 2014; Copeland *et al.*, 2019).[3,4]

We construct compact unbalanced panels of youth respondents present in the first year of the respective cohort (to facilitate initial conditions estimation) that have no missing values (to model persistence) in any of the covariates included in the estimations.[5] We therefore analyse contiguous sequences of non-missing data including individuals pertaining to the sample at the starting period of each of the 12 cohorts, that is, individuals can exit a given cohort but cannot re-enter. Such a sample selection mechanism is employed for instance by Arulampalam, Booth and Taylor (2000), Contoyannis, Jones and Rice (2004), and suggested by Skrondal and Rabe-Hesketh (2014). An individual can belong to a maximum of three adjacent 'pseudo-cohorts' since cohort sample membership requires a total of three consecutive observations including the cohort initiation period (permitting inclusion of both dynamics and initial conditions). Balancing the panels is prohibitive in terms of sample attrition and can produce inconsistent parameter estimates (see Albarran, Carrasco and Carro, 2019).

The initial year self-reported life satisfaction is included to capture the impact of early life events. However, this initial year can be when the individual is 11, 12 or 13 years old depending on the first year an individual is included in a specific cohort. For the initial period to act as a valid proxy for early life events/shocks for all sample members, we use the self-reported value of life satisfaction at age 11 as the initial value for all individuals. To this extent, for individuals aged 12 or 13 at the initiation period of a given cohort, we use the life satisfaction level reported 1 or 2 years before correspondingly such that their initial value is the reported outcome at age 11. Therefore, the first BHPS cohort commences in 1996 and uses 1994 and 1995 to compute initial life satisfaction values for 13 and 12-year-olds in 1996. Likewise, the first UKHLS cohort starts in 2011 and uses 2009 and 2010 to compute initial values for 13- and 12-year-olds in 2011.

Youth sample membership endures for a maximum of 5 years in the BHPS and by construction is reduced from 6 to 5 in the UKHLS. Subsequently youth respondents may enter the adult samples. We can define 9 BHPS 5-year length cohorts spanning

[3]Hamilton *et al.* (2014) conclude that premature pubertal maturation predicts depressive symptoms for adolescents with more negative cognitive styles and adolescent girls with poor emotional clarity. Puberty describes the physical changes to sexual maturation, whereas the psychosocial and cultural maturation is strictly speaking denoted by the term adolescent development. Individual pubertal timetables are variable and are primarily affected by heredity, although environmental epigenetic factors can also be important. The female pubertal landmark is menarche (the onset of menstruation), see https://www.nhs.uk/live-well/sexual-health/stages-of-puberty-what-happens-to-boys-and-girls/. Pubertal status cannot be appropriately described by a single parameter and Tanner's widely used stages of morphological puberty require separate ratings on a number of dimensions – see Angold and Worthman, 1993, pp. 149–50.

[4]The sample sizes including individuals aged 14 at any of the 12 cohorts' initiation period are approximately 2% higher. Restricting sample membership such that individuals are 11 years old at each cohort initiation year introduces substantial sample attrition averaging 59% across cohorts.

[5]Life satisfaction is not reported for 44/11,111 and 205/21,278 observations in the BHPS 1994–2008 and UKHLS 2009–17 youth samples respectively. Missing values in the BHPS: household income (39), talking to mum/dad (20/39), family meals (26), close friends number (114), fights (18), smoking (76) and region (12). Missing values in the UKHLS: household income (18), family meals (117), close friends number (592), smoking (186) and region (6).

1996–2008 and 3 UKHLS 5-year cohorts spanning 2011–17. As the BHPS and UKHLS are not cohort studies and follow the same representative sample of individuals (the panel) over a period of years, the cohorts constructed in this study should be viewed as *ad hoc* ‘pseudo-cohorts’. Specifically, the first BHPS ‘pseudo-cohort’ (1996 cohort) in an unbalanced panel spanning the period 1996–2000 and the last BHPS cohort (2004 cohort) spans 2004–08 (see Figures 1–9). Similarly, the first UKHLS ‘pseudo-cohort’ (2011 cohort) is an unbalanced panel spanning 2011–15 and the final UKHLS cohort (2013 cohort) spans 2013–17 (see Figures 10–12).

Matching individual youth respondents to the household level data files we obtain the household income, household size, the number of children in the household (aged 15 or under) and the government office region variables (the last is included in the UKHLS youth questionnaires). Matching the merged youth and household files to the corresponding parental interview files is prohibitive in terms of sample attrition if one wishes to undertake a longitudinal analysis. The existing set of explanatory variables

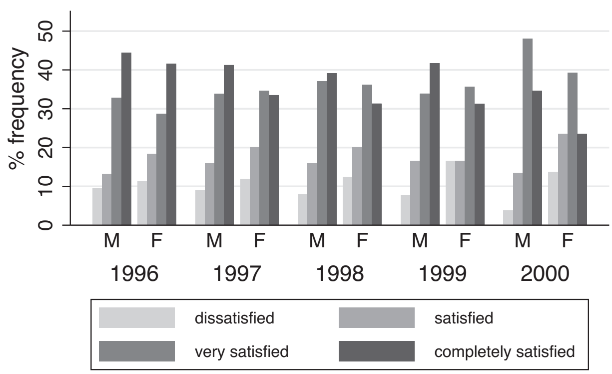

Figure 1. Adolescent life satisfaction by gender (Age 11–15) BHPS, compact unbalanced panel 1996–2000

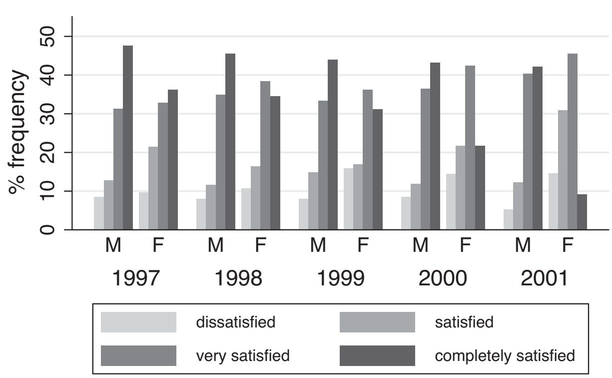

Figure 2. Adolescent life satisfaction by gender (Age 11–15) BHPS, compact unbalanced panel 1997–2001

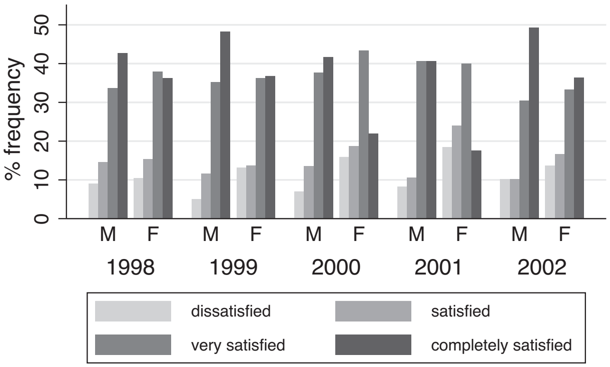

Figure 3. Adolescent life satisfaction by gender (Age 11–15) BHPS, compact unbalanced panel 1998–2002

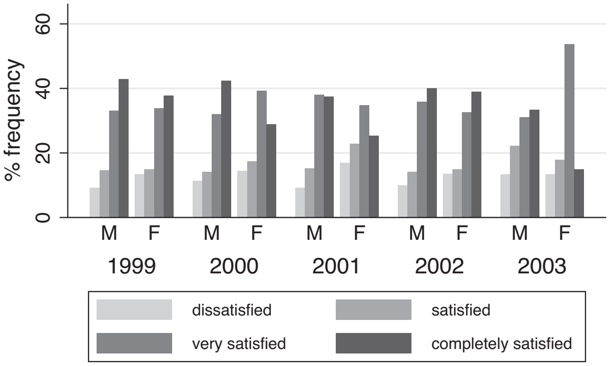

Figure 4. Adolescent life satisfaction by gender (Age 11–15) BHPS, compact unbalanced panel 1999–2003

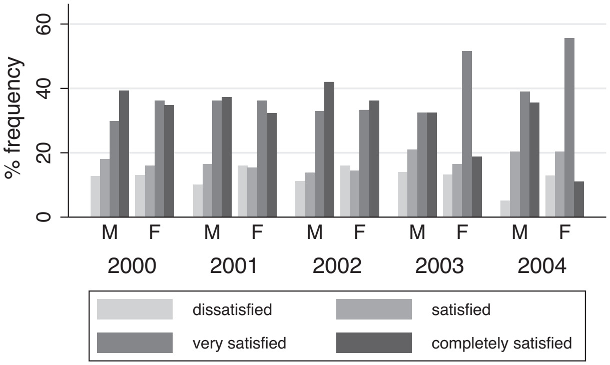

Figure 5. Adolescent life satisfaction by gender (Age 11–15) BHPS, compact unbalanced panel 2000–2004

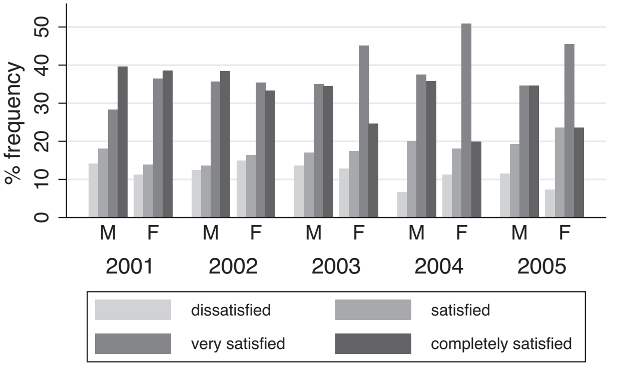

Figure 6. Adolescent life satisfaction by gender (Age 11–15) BHPS, compact unbalanced panel 2001–2005

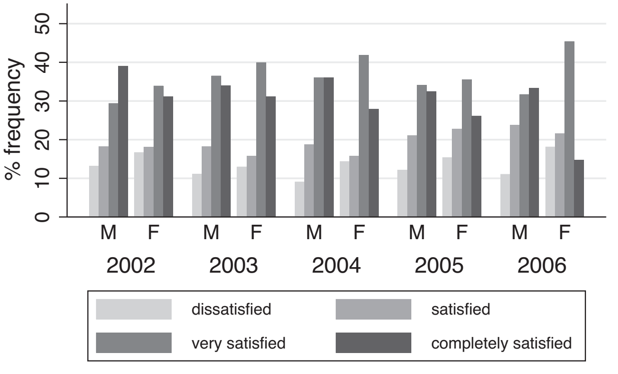

Figure 7. Adolescent life satisfaction by gender (Age 11–15) BHPS, compact unbalanced panel 2002–2006

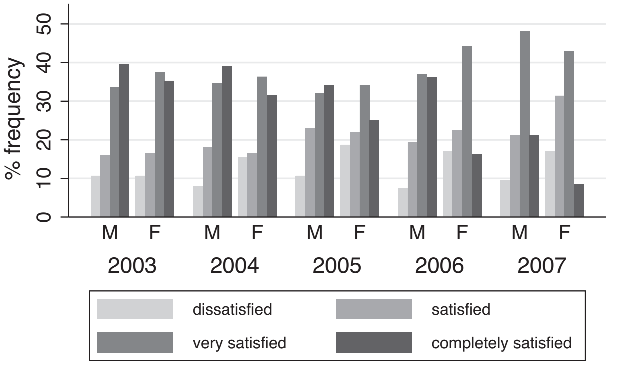

Figure 8. Adolescent life satisfaction by gender (Age 11–15) BHPS, compact unbalanced panel 2003–2007

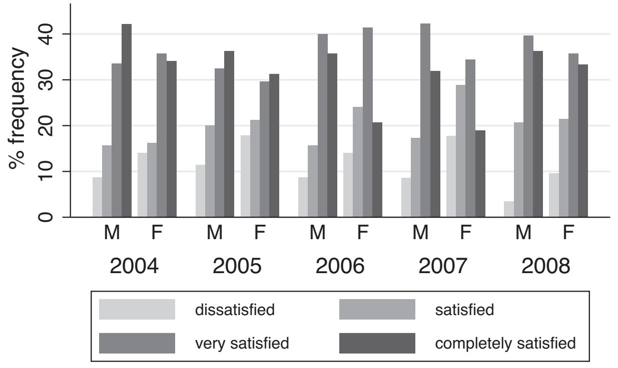

Figure 9. Adolescent life satisfaction by gender (Age 11–15) BHPS, compact unbalanced panel 2004–2008

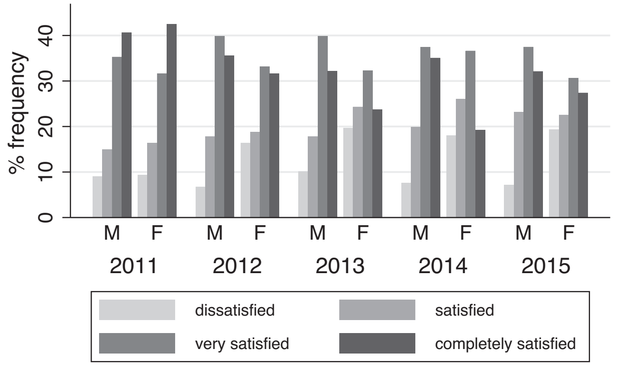

Figure 10. Adolescent life satisfaction by gender (Age 11–15) UKHLS, compact unbalanced panel 2011–2015

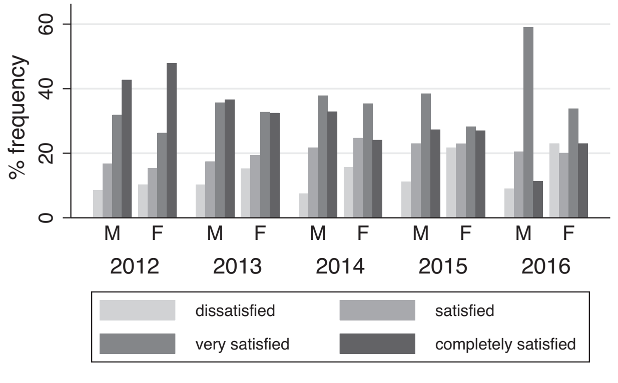

Figure 11. Adolescent life satisfaction by gender (Age 11–15) UKHLS, compact unbalanced panel 2012–2016

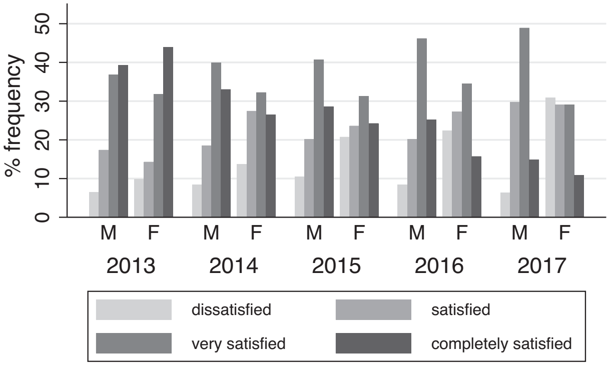

Figure 12. Adolescent life satisfaction by gender (Age 11–15) UKHLS, compact unbalanced panel, 2013–2017

seems to adequately control for the family environment.[6] Parental conversation frequencies (Not talking to Mum/Dad) and the frequency of fights are reported biennially in the UKHLS and can thus be only included in the BHPS cohort regressions. Finally, note that puberty is a period of rapid physical growth and psychological changes, culminating in sexual maturity. Weight and height could be used to construct of a body mass index and proxy physical changes and attractiveness. However, youth respondent weight and height are only reported in 2004, 2006 and 2008 in the BHPS and in 2010, 2012, 2014 and 2016 in the UKHLS.[7,8]

III. Longitudinal evolution of adolescent life satisfaction by gender

The BHPS and UKHLS Youth questionnaires contain a life satisfaction variable each year during 1994–2008 and 2009–17. The respective question is ‘Which best describes how you feel about your life as a whole?’. Responses take values 1–7 with 1 corresponding to ‘completely happy’ and 7 to ‘completely unhappy’. We collapse responses into a four-point scale variable by grouping the less frequently reported values 4–7 of lower satisfaction levels to aid identification. We then reverse code the response values to become increasing in life satisfaction such that 1 corresponds to ‘dissatisfied’, 2 to ‘satisfied’, 3 to ‘very satisfied’ and 4 to ‘completely satisfied’. The original seven-point scale life-satisfaction

[6]For example, adding the number employed in the household from the household respondent files does not significantly affect the estimates and seems to be collinear with the household income per capita control (see Tables 34 and 35, online Appendix).

[7]We employ the BHPS main original sample and the Great Britain (England, Scotland and Wales) general population UKHLS sample noting that both exclude Northern Ireland. The Scotland and Wales extension samples added in the BHPS in 1999 and the Ethnic Minority Boost UKHLS sample are also excluded since they were obtained using distinct sample selection mechanisms. Ethnic group membership is not asked of youth respondents in the BHPS. In the UKHLS youth questionnaires ethnicity is reported in the odd waves biennially. In additional regressions (see Table 36, online Appendix), we find that white background has a positive statistically significant impact on life satisfaction in the 2013–17 cohort only while, it has a minimal impact on the significant gender gradient (0.326 as opposed to 0.315 in Table 2).

[8]To calculate the log of real household income per capita, we divide household income by household size and use the CPI deflator from the Office of National Statistics (Consumer Price Indices – CPI index: all items: 2005 = 100 (BHPS), 2015 = 100 (UKHLS)).

responses for each cohort are plotted in histograms in Figures (A1–A12, online Appendix) and the respective frequencies for the effective estimation sample per cohort are given in Tables (9–20, online Appendix). Summary statistics for all variables are given in the Descriptive Statistics Tables (7 and 8, online Appendix).

The longitudinal evolution of the four-point scale self-reported life satisfaction by gender is depicted in the five-year cohort-specific histograms in Figures 1–12 where the first and last figures correspond to the 1996–2000 and 2013–17 cohorts respectively. A clear pattern is generally evident across the annual observations of all 12 cohorts: the frequencies of male adolescents reporting complete life satisfaction are higher than their female counterparts while in contrast, the frequencies of female adolescents reporting life dissatisfaction are greater as opposed to males.[9]

Hence, the raw data clearly indicate a cross-cohort gender gradient at the two extremes of the self-reported life satisfaction distribution. This is in accordance with a large body of interdisciplinary literature – see Altemus *et al*. (2014) for a review on gender differences in anxiety, depression, affective disorders and mental health; Bergman and Scott (2001) for evidence on greater female unhappiness using BHPS 1994-97 youth questionnaires; Chrysanthou and Vasilakis (2019) for evidence on higher adolescent female emotional symptoms and lower life satisfaction (UKLHS 2009–13); Van de Velde *et al*. (2010) for European cross-national evidence of the adult (18–75) gender depression gap (using the Center for Epidemiologic Studies Depression Scale, CES-D).

IV. Estimation methodology

The reduced-form model for self-reported life satisfaction is a dynamic random effects (RE) ordered probit given by

$$y_{it}^{*} = \boldsymbol{x}_{it}\boldsymbol{\beta} + \gamma y_{it-1} + \varepsilon_i + \eta_{it}; \quad i = 1, \ \ldots, \quad N; \quad t = 2, \ \ldots, \ T_i \tag{1}$$

where y_{it}^{*} is a latent ordered response variable capturing satisfaction level propensity, $\boldsymbol{x}_{it}$ is a vector of contemporaneous explanatory variables (including gender) for the ith individual in the tth time period and the vector $\boldsymbol{\delta} = (\boldsymbol{\beta}, \gamma)$ represents the set of the unknown parameters to be estimated. The composite error term $v_{it} = \varepsilon_i + \eta_{it}$ captures the unobserved heterogeneity underlying individual life satisfaction and is decomposed into an individual-specific time-invariant component $\{\varepsilon_i\}_{i=1,2,\ldots,N}$ and an individual time-specific effect η_{it} assumed to be serially uncorrelated and normally distributed $\eta_{it} \sim N(0, \sigma_\eta^2)$ independently of ε_i. Since y_{it} is binary a normalization is necessary and the usual normalization $\sigma_\eta^2 = 1$ is adopted.

Including y_{it-1} in equation (1) raises the initial conditions problem (Heckman, 1981a, b) which is subsequently addressed. Self-reported life satisfaction status $y_{it} = j$ for $j \in \{1, \ldots, J\}$ is observed if latent life satisfaction falls in an interval between μ_{j-1} and μ_j:

[9]Some notable exceptions exist. At the lower end of the distribution in the 2001 cohort (Figure 6) where for 3 out of 5 years male dissatisfaction frequencies are higher. At the top of the distribution, initial period complete satisfaction is higher for females in the 2011, 2012 and 2013 cohorts in (Figures 10–12).

$$y_{it} = j \text{ if } \mu_{j-1} < y_{it}^{*} \leq \mu_j \tag{2}$$

where $\mu_0 = -\infty$, $\mu_j \leq \mu_{j+1}$ and $\mu_J = \infty$. Under the normality assumption of η_{it}, the probability p_{itj} of observing outcome j for response y_{it}, conditional on the set of cutpoints $\boldsymbol{\mu} = \{\mu_1, \mu_2, \ldots, \mu_{J-1}\}$, $\mathbf{x_{it}}$ and ε_i is

$$Pr(y_{it} = j | \boldsymbol{\mu}, \boldsymbol{x}_{it}, \varepsilon_i) = \Phi\left(\mu_j - \boldsymbol{x}_{it}\boldsymbol{\beta} - \boldsymbol{y}_{it-1}\boldsymbol{\gamma} - \varepsilon_i\right) - \Phi\left(\mu_{j-1} - \boldsymbol{x}_{it}\boldsymbol{\beta} - \boldsymbol{y}_{it-1}\boldsymbol{\gamma} - \varepsilon_i\right) \tag{3}$$

where Φ is the standard normal *cdf* and $\mathbf{y_{it-1}}\boldsymbol{y}_{it-1}$ is a vector of J–1 lagged indicators, $\mathbf{1}[y_{it-1} = j]$, $j = 2, \ldots, J$.

Note that $\boldsymbol{x}_{it}$ does not include a constant which is absorbed into the cutpoints provided that we cannot separately identify a global intercept and the cutpoints $\boldsymbol{\mu}$, i.e. only $\left(\mu_j - \varepsilon_i\right)$ is identified. In addition, in a random effects framework we cannot disentangle ε_i from individual-specific cutpoint shifts. Integrating out $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$, we obtain the sample log likelihood function

$$lnL(\boldsymbol{\beta}, \boldsymbol{\mu}, \sigma_\varepsilon^2) = \sum_{i=1}^{N} ln \int_{-\infty}^{+\infty} \frac{exp(-\frac{\varepsilon_i^2}{2\sigma_\varepsilon^2})}{\sqrt{2\pi}\sigma_\varepsilon} \left\{ \prod_{t=2}^{T} Pr(y_{it} = j | \boldsymbol{\mu}, \boldsymbol{x}_{it}, \varepsilon_i) \right\} d\varepsilon_i. \tag{4}$$

Given the time-invariance of gender, the key variable of interest, we are unable to perform conditional maximum likelihood (CMLE) fixed effects (FE) estimation for dichotomizations of life satisfaction as this eliminates time-invariant covariates along with ε_i. Consistency of the RE estimator relies on the orthogonality assumption between ε_i and $\boldsymbol{x}_{it}$. To the extent that gender is exogenously determined for individual adolescents, the RE estimator provides a consistent estimate of the gender gradient of life satisfaction but, caution should be taken at interpreting other time-invariant covariates' coefficients (see section Initial conditions: disentangling the true dependence of life satisfaction). We refer to the estimates obtained using equation (4) as panel RE estimates.

Initial conditions: disentangling the true dependence of life satisfaction

The presence of ε_i in (3) invalidates the assumption of exogeneity of self-reported life satisfaction in the initiation period of each cohort (y_{i1}) since the beginning of the sample is unlikely to coincide with the initiation of the stochastic process determining satisfaction level propensity. State dependence and individual heterogeneity offer 'diametrically opposite' explanations of persistence in life satisfaction outcomes (see Hsiao, 2003, p. 216). Considering otherwise identical adolescents, those experiencing lower levels of life satisfaction in the past (e.g. due to exogenous adverse events) could amend their behaviour which in turn determines future satisfaction levels: this is an entirely behavioural effect that can trigger a self-perpetuating cycle.

Alternatively, adolescents may differ in specific unobservables affecting satisfaction level propensity, while at the same time not being influenced by previous satisfaction levels (e.g. latent personality/behavioural traits such as sensitivity, dominance, emotional stability, self-reliance, social boldness, genetic factors, sexual preferences and attractiveness). If such unobservables are temporally correlated, and not accounted for, past satisfaction levels will appear as the most prominent determinant of future life satisfaction since they act as proxy for the temporally persistent unobservables. This denotes spurious as opposed to true (structural) state dependence (Heckman, 1981a, b).

Wooldridge (2005) proposes specifying a distribution of ε_i conditional on the initial condition, as opposed to Heckman's (1981b) proposal to obtain the joint distribution of all outcomes of the endogenous variables. We use Wooldridge's (2005) solution to the initial conditions problem due to is computational simplicity. Employing Mundlak's (1978) correlated random effects (CRE) estimator, we induce a correlation between ε_i and the time means of the time-varying covariates taking the form of $\varepsilon_i = \bar{\boldsymbol{x}}_i \boldsymbol{a} + \lambda_i$, where $\lambda_i \sim iidN(0, \sigma_\lambda^2)$ and is independent of $(\boldsymbol{x}_{it}, \eta_{it})$ for all (i,t) in (1).[10] The model for the unobserved individual effect, λ_i, in its simplest form is

$$\lambda_i = \vartheta_0 + \vartheta_1 y_{i1} + \zeta_i \tag{5}$$

where ζ_i is $N(0, \sigma_\zeta^2)$ and independent of the initial condition, the covariates and η_{it}. Since we cannot separately identify ϑ_0 from $\boldsymbol{\mu}$, we adopt the usual normalization setting $\vartheta_0 = 0$.

The ordered choice log likelihood function in (4) is modified such that the explanatory variables at time t are $\mathbf{q}_{it} \equiv (\boldsymbol{x}_{it}, \boldsymbol{y}_{it-1}, \boldsymbol{y}_{i1}, \bar{\boldsymbol{x}}_i)$ where $\boldsymbol{y}_{it-1}$ and $\boldsymbol{y}_{i1}$ denote, respectively, the vectors of the J–1 lagged, $\mathbf{1}[y_{it-1} = j]$, and initial conditions set of indicators, $\mathbf{1}[y_{i1} = j], j = 2, \ldots, J$. Finally, $\bar{\boldsymbol{x}}_i = (T_i - 1)^{-1} \sum_{t=2}^{T_i} \boldsymbol{x}_{it}$ as suggested by Rabe-Hesketh and Skrondal (2013).[11] Including time-constant covariates in $\boldsymbol{x}_{it}$ only increases the explanatory power as it is not possible to separately identify their partial effects from their partial correlation with the unobserved effect. Due to minimal within variation, we cannot include individual-specific time means of the regional control for London/South East/South West/East of England (as opposed to living in the rest of Great Britain). We do not add within means of household income and the number of dependent children in the household as they are outcomes of parental socioeconomic attributes, choices and behaviour. As within cohorts individuals age jointly, the models only include time effects since age is collinear with the time dummies and the threshold cutpoints. The estimations in Tables 1 and 2 reveal a negative time trend indicating a downward age-related adjustment in life satisfaction (also visible in Figures 1–12) noting that statistical significance varies.

[10]Arulampalam and Stewart (2009) show that, none of the Heckman (1981b) and Wooldridge (2005) solutions dominates the other and given the Mundlak (1978) CRE device is used the estimators provide similar results.

[11]In terms of relative bias and RMSE, this version performs similarly to the specification of the conditional distribution of the unobserved effect used in Wooldridge (2005) except in the case of an AR(1) process assumed for $\boldsymbol{x}_{it}$ with short panels (see Rabe-Hesketh and Skrondal, 2013).

TABLE 1

Adolescent life satisfaction, cohorts 1996–2001, heteroskedastic CRE ordered probits

	(1) 1996–2000	(2) 1997–2001	(3) 1998–2002	(4) 1999–2003	(5) 2000–2004	(6) 2001–2005
Completely Satisfied (t − 1)	1.1089***	1.2874***	0.7648***	0.7230***	0.2348*	0.1141
	(0.1875)	(0.1754)	(0.1382)	(0.1115)	(0.1280)	(0.1431)
Very Satisfied (t − 1)	0.5397***	0.7539***	0.3362***	0.3928***	0.0414	−0.0704
	(0.1752)	(0.1672)	(0.1242)	(0.1014)	(0.1248)	(0.1381)
Satisfied (t−1)	0.2075	0.5056***	0.1365	0.2720**	0.0148	0.0069
	(0.1703)	(0.1780)	(0.1341)	(0.1194)	(0.1440)	(0.1533)
Completely Satisfied (t = 1)	0.3692*	0.2172	0.1237	0.0308	0.2950**	0.2206
	(0.2044)	(0.1625)	(0.1241)	(0.1053)	(0.1433)	(0.1395)
Very Satisfied (t = 1)	0.1863	0.0568	0.1076	0.0462	0.2513*	0.1533
	(0.2055)	(0.1681)	(0.1231)	(0.1074)	(0.1425)	(0.1396)
Satisfied (t = 1)	0.0803	0.0080	0.0008	0.0333	−0.0495	−0.1960
	(0.1980)	(0.1807)	(0.1384)	(0.1373)	(0.1605)	(0.1516)
Male	0.2291***	0.3515***	0.3793***	0.2575***	0.2980***	0.2453***
	(0.0798)	(0.0902)	(0.0679)	(0.0669)	(0.0825)	(0.0846)
Not Talking to Mum	−0.2395**	−0.2629**	−0.1455	−0.3189***	−0.1876*	−0.1886
	(0.1195)	(0.1211)	(0.1011)	(0.0962)	(0.1107)	(0.1236)
Not Talking to Dad	−0.0400	−0.0452	−0.0412	0.0198	0.0062	−0.0987
	(0.1019)	(0.1034)	(0.0928)	(0.0910)	(0.0939)	(0.0988)
Number of Children in Household	0.0340	−0.0145	−0.0159	−0.0287	−0.0039	−0.0171
	(0.0437)	(0.0420)	(0.0379)	(0.0378)	(0.0390)	(0.0439)
> 2 Family Meals Weekly	−0.0010	0.1050	0.0937	−0.0396	0.0154	−0.0662
	(0.0996)	(0.1073)	(0.0871)	(0.0858)	(0.0907)	(0.1009)
Close Friends Number	0.0393***	0.0130	0.0035	−0.0094	0.0001	0.0116*
	(0.0105)	(0.0090)	(0.0056)	(0.0071)	(0.0055)	(0.0062)

(*Continued*)

TABLE 1

(*Continued*)

	(1) 1996–2000	(2) 1997–2001	(3) 1998–2002	(4) 1999–2003	(5) 2000–2004	(6) 2001–2005
Smokes/Smoked	−0.1157	−0.1595	−0.2089**	0.0715	0.0192	−0.0065
	(0.1105)	(0.1086)	(0.0858)	(0.0871)	(0.1008)	(0.1031)
Frequency of Fights with Someone (past month)	−0.1205*	−0.1491*	−0.0809	−0.0996**	−0.1366**	−0.0911
	(0.0668)	(0.0786)	(0.0550)	(0.0459)	(0.0538)	(0.0599)
Ln(Real Household Income p.capita)	0.0539	0.0555	0.0105	0.0429	0.0568	−0.0069
	(0.0663)	(0.0728)	(0.0627)	(0.0622)	(0.0647)	(0.0662)
London, S.East, S.West, East England	−0.0650	−0.1334	−0.0904	−0.1048	−0.1101	−0.0788
	(0.0785)	(0.0816)	(0.0672)	(0.0691)	(0.0769)	(0.0796)
$T_i = 3$	−0.0120	−0.4043	−0.1731	−0.3702	0.2197	0.0184
	(0.3118)	(0.3394)	(0.3198)	(0.2522)	(0.3265)	(0.3309)
$T_i = 4$	−0.1743	−0.4419	0.0496	−0.0964	−0.0069	0.0812
	(0.2969)	(0.3192)	(0.2387)	(0.2528)	(0.2993)	(0.2436)
$t_1 + 2$	0.0396	−0.0442	−0.2096***	−0.1164*	0.0456	−0.0818
	(0.0872)	(0.0914)	(0.0758)	(0.0649)	(0.0931)	(0.0963)
$t_1 + 3$	−0.0506	−0.2346**	−0.2067**	0.0606	−0.2455***	−0.0777
	(0.1023)	(0.1096)	(0.0843)	(0.0995)	(0.0933)	(0.1049)
$t_1 + 4$	−0.2552**	−0.3002**	0.0495	−0.2516**	−0.2141*	−0.1467
	(0.1179)	(0.1359)	(0.1496)	(0.1202)	(0.1162)	(0.1317)
$Variance function : ln(\sigma_\zeta)$						
$T_i = 3$	0.1180	0.1307	−0.1433	−0.3132***	0.1564*	0.1013
	(0.0968)	(0.1005)	(0.1052)	(0.0958)	(0.0909)	(0.0943)
$T_i = 4$	0.0347	0.0954	−0.1266	−0.0476	0.1440*	0.0181
	(0.0913)	(0.0983)	(0.0883)	(0.0791)	(0.0809)	(0.0922)

(*Continued*)

TABLE 1

(*Continued*)

	(1) 1996–2000	(2) 1997–2001	(3) 1998–2002	(4) 1999–2003	(5) 2000–2004	(6) 2001–2005
μ_1	−0.4264	−0.6055	−0.9037**	−0.8389**	−0.9921***	−1.4112***
	(0.3994)	(0.4161)	(0.3704)	(0.3459)	(0.3660)	(0.3997)
μ_2	0.4111	0.1710	−0.2997	−0.2340	−0.3231	−0.7405*
	(0.3947)	(0.4116)	(0.3642)	(0.3387)	(0.3590)	(0.3912)
μ_3	1.5698***	1.4068***	0.7599**	0.7224**	0.7995**	0.3651
	(0.4015)	(0.4204)	(0.3594)	(0.3357)	(0.3572)	(0.3844)
Log-Likelihood	−1233.575	−1211.615	−1332.946	−1397.073	−1438.076	−1368.990
Number of Observations	1081	1073	1155	1143	1145	1087
Wald (Global Significance)	188.123	151.641	165.397	144.415	80.969	56.955
LR: $ln(\sigma_\zeta)$	1.548	1.867	2.943	11.441	4.238	1.367
P-value: $ln(\sigma_\zeta)$	0.461	0.393	0.230	0.003	0.120	0.505

Notes: ***, **, * is statistical significance at 1%, 5% and 10% respectively. Source: (BHPS: Cohorts 1996–2001). Compact Unbalanced Panels. Individual cluster-robust standard errors (in parentheses). t_1 is the cohort-specific initial period. Estimations include ($T_i = 3$, $T_i = 4$) interactions with the individual-specific means of (Not Talking to Mum/Dad, > 2 Family Meals Weekly, Close Friends, Smokes/Smoked, Frequency of Fights).

TABLE 2

Adolescent life satisfaction, cohorts 2002–13, heteroskedastic CRE ordered probits

	(1) *2002–2006*	*(2)* *2003–2007*	*(3)* *2004–2008*	*(4)* *2011–2015*	*(5)* *2012–2016*	*(6)* *2013–2017*
Completely Satisfied (t−1)	1.2705***	1.2783***	1.5168***	1.5835***	1.7496***	1.4665***
	(0.1387)	(0.1816)	(0.1851)	(0.1673)	(0.1753)	(0.1542)
Very Satisfied (t−1)	0.6992***	0.6077***	0.8524***	0.8879***	0.9209***	0.8229***
	(0.1272)	(0.1677)	(0.1492)	(0.1346)	(0.1335)	(0.1266)
Satisfied (t−1)	0.2993**	0.2929*	0.4662***	0.3480***	0.4816***	0.3102**
	(0.1228)	(0.1630)	(0.1432)	(0.1315)	(0.1295)	(0.1270)
Completely Satisfied (t = 1)	−0.2104*	−0.1076	0.0126	0.2665**	0.2987**	0.4646***
	(0.1198)	(0.1658)	(0.1721)	(0.1353)	(0.1471)	(0.1502)
Very Satisfied (t = 1)	−0.2410*	−0.1725	−0.0961	0.1362	0.2461*	0.4542***
	(0.1248)	(0.1632)	(0.1670)	(0.1247)	(0.1445)	(0.1463)
Satisfied (t = 1)	−0.3175**	−0.2910*	−0.2403	0.0713	0.0644	0.3204**
	(0.1317)	(0.1628)	(0.1645)	(0.1284)	(0.1414)	(0.1433)
Male	0.2002***	0.3013***	0.2710***	0.3188***	0.2241***	0.3146***
	(0.0768)	(0.0878)	(0.0840)	(0.0654)	(0.0660)	(0.0709)
Not Talking to Mum	−0.2598**	−0.2637**	−0.2799**			
	(0.1026)	(0.1169)	(0.1239)			
Not Talking to Dad	−0.1050	−0.1704	−0.0447			
	(0.0877)	(0.1137)	(0.1152)			
Number of Children in Household	−0.0045	0.0193	0.0256	0.0736**	0.0710*	0.0349
	(0.0374)	(0.0486)	(0.0448)	(0.0349)	(0.0400)	(0.0413)
> 2 Family Meals Weekly	0.1362	0.0896	0.1913*	0.0294	0.3485***	0.2214**
	(0.0840)	(0.0984)	(0.0998)	(0.0968)	(0.1014)	(0.1056)

(*Continued*)

TABLE 2

(*Continued*)

	(1) *2002–2006*	*(2)* *2003–2007*	*(3)* *2004–2008*	*(4)* *2011–2015*	*(5)* *2012–2016*	*(6)* *2013–2017*
Close Friends Number	0.0058	0.0036	0.0018	0.0131	0.0086	0.0099**
	(0.0056)	(0.0068)	(0.0073)	(0.0080)	(0.0084)	(0.0047)
Smokes/Smoked	−0.2361**	−0.1064	−0.2395*	−0.3971***	−0.0282	−0.1287
	(0.0974)	(0.1150)	(0.1244)	(0.1421)	(0.1844)	(0.2078)
Frequency of Fights with Someone (past month)	−0.1613***	−0.2175***	−0.0498			
	(0.0580)	(0.0571)	(0.0694)			
Ln(Real Household Income p.capita)	0.0637	0.0322	0.0447	0.1245*	0.0849	0.0356
	(0.0576)	(0.0717)	(0.0658)	(0.0681)	(0.0660)	(0.0688)
London, S.East, S.West, East England	−0.1732**	−0.1626*	−0.2256***	−0.1908***	−0.1500**	−0.0437
	(0.0746)	(0.0903)	(0.0812)	(0.0649)	(0.0692)	(0.0703)
$T_i = 3$	−0.0696	−0.2572	0.0621	−0.7041***	−0.1058	−0.1285
	(0.2178)	(0.3434)	(0.3564)	(0.1933)	(0.2181)	(0.2059)
$T_i = 4$	0.0526	−0.3469	0.0816	−0.5110**	−0.3037	−0.0237
	(0.2479)	(0.3001)	(0.2476)	(0.2190)	(0.2325)	(0.2161)
$t_1 + 2$	0.0387	−0.1918**	0.1217	−0.0984	−0.0007	0.0046
	(0.0759)	(0.0978)	(0.0860)	(0.0687)	(0.0710)	(0.0743)
$t_1 + 3$	−0.1505*	−0.0344	−0.0839	−0.1158	−0.1047	−0.0357
	(0.0900)	(0.0947)	(0.0982)	(0.0803)	(0.0880)	(0.0859)
$t_1 + 4$	−0.1523	−0.3888***	0.1036	−0.1342	−0.1344	−0.0937
	(0.1045)	(0.1265)	(0.1382)	(0.1213)	(0.1183)	(0.1237)
$Variance function: ln(\sigma_\zeta)$						
$T_i = 3$	−0.1462	0.1546	0.1523	0.1593**	0.1995**	0.1761**
	(0.0912)	(0.1099)	(0.0934)	(0.0719)	(0.0797)	(0.0787)

(*Continued*)

TABLE 2

(*Continued*)

	(1) 2002–2006	(2) 2003–2007	(3) 2004–2008	(4) 2011–2015	(5) 2012–2016	(6) 2013–2017
$T_i = 4$	−0.0178	0.1414	−0.0618	0.1263*	0.1045	0.0510
	(0.0836)	(0.0892)	(0.0931)	(0.0741)	(0.0834)	(0.0782)
μ_1	−0.8629***	−1.1900***	−0.4837	−0.0059	0.3512	0.4568*
	(0.3316)	(0.4502)	(0.3994)	(0.2694)	(0.2477)	(0.2615)
μ_2	−0.1091	−0.2833	0.4210	0.9483***	1.3302***	1.4358***
	(0.3347)	(0.4485)	(0.4012)	(0.2844)	(0.2554)	(0.2668)
μ_3	1.0206***	0.9545**	1.6294***	2.2370***	2.6186***	2.6944***
	(0.3412)	(0.4567)	(0.4187)	(0.3126)	(0.2810)	(0.2895)
Log-Likelihood	−1457.456	−1283.889	−1179.194	−2118.784	−1913.023	−1784.732
Number of Observations	1247	1083	1034	1840	1668	1508
Wald (Global Significance)	220.142	139.785	166.328	234.342	214.498	204.697
LR: $ln(\sigma_\zeta)$	2.664	2.958	5.443	5.132	6.364	5.555
P-value: $ln(\sigma_\zeta)$	0.264	0.228	0.066	0.077	0.042	0.062

Notes: ***, **, * is statistical significance at 1%, 5% and 10% respectively. Source: (BHPS: Cohorts 2002–04), (UKHLS: Cohorts 2011–13). Compact Unbalanced Panels. Individual cluster-robust standard errors (in parentheses). t_1 is the cohort-specific initial period. Estimations include ($T_i = 3$, $T_i = 4$) interactions with the individual-specific means of (Not Talking to Mum/Dad, > 2 Family Meals Weekly, Close Friends, Smokes/Smoked, Frequency of Fights).

Correlated random effects (CRE) ordered probits with unbalanced panels

Given the unbalancedness of the panels, we follow Wooldridge (2019) to allow unobserved heterogeneity to be correlated with time-varying observed covariates and unbalanced panel sample selection. The compact unbalanced panels employed have a minimum of three consecutive observations (to account for dynamics and initial conditions) and a maximum of five consecutive observations corresponding to the maximum duration of youth sample membership. The series of selection indicators for each i is denoted by $\{s_{i3}, \ldots, s_{iT}\}$ where $s_{it} = 1$ if individual i is included in estimation in time period t (i.e. *iff* $(\mathbf{x_{it}}, y_{it})$ is fully observed). The number of time periods observed for individual i in the sample is $T_i = \left(2 + \sum_{r=3}^{T} s_{ir}\right)$, $T = 5$.

Maintaining strict exogeneity of selection (s_{it} at time period t unrelated to η_{it}), s_{it} at time t can have an arbitrary correlation with the unobserved heterogeneity (ζ_i) and the observed covariates. Following Wooldridge (2019), we consider a linear function of the within means with different coefficients per number of periods: $E(\zeta_i | s_{it}, s_{it}\mathbf{x_{it}}) = \sum_{r=3}^{T} \psi_r \mathbf{1}[T_i = r] + \sum_{r=3}^{T} \mathbf{1}[T_i = r]\bar{\mathbf{x}}_{\mathbf{i}}\boldsymbol{\xi}_{\mathbf{r}}$. The sample log likelihood function becomes

$$lnL(\boldsymbol{\beta}, \boldsymbol{\mu}, \omega_r) = \sum_{i=1}^{N} ln \prod_{t=2}^{T} \left[\Phi\left(\frac{\mu_j - \boldsymbol{x}_{it}\boldsymbol{\beta} - \boldsymbol{y}_{it-1}\boldsymbol{\gamma} - \boldsymbol{y}_{i1}\boldsymbol{\vartheta}_1 - \sum_{r=3}^{T} \psi_r \mathbf{1}[T_i = r] - \sum_{r=3}^{T} \mathbf{1}[T_i = r]\bar{\boldsymbol{x}}_i\boldsymbol{\xi}_{\mathrm{r}}}{exp\left(\sum_{r=3}^{T} \mathbf{1}[T_i = r]\omega_r\right)^{1/2}} \right) - \Phi\left(\frac{\mu_{j-1} - \boldsymbol{x}_{it}\boldsymbol{\beta} - \boldsymbol{y}_{it-1}\boldsymbol{\gamma} - \boldsymbol{y}_{i1}\boldsymbol{\vartheta}_1 - \sum_{r=3}^{T} \psi_r \mathbf{1}[T_i = r] - \sum_{r=3}^{T} \mathbf{1}[T_i = r]\bar{\boldsymbol{x}}_i\boldsymbol{\xi}_{\mathrm{r}}}{exp\left(\sum_{r=3}^{T} \mathbf{1}[T_i = r]\omega_r\right)^{1/2}} \right) \right] \quad (6)$$

where the denominator is the square root of the variance of ζ_i (standard error) and corresponds to unity in the balanced case where each deviation ω_r from the base group ($T_i = T$) is equal to zero. The likelihood in equation (6) can therefore be estimated using pooled heteroskedastic ordered probit software. An alternative specification, adding interactions $\mathbf{1}[T_i = r]\bar{\boldsymbol{x}}_i$ to the variance function in the denominator of equation (6), gives similar estimates (see online Appendix, Table 23).

Average partial effects (APEs)

Given the nonlinear nature of the models, the estimated parameters are only informative regarding the direction and relative impact of the explanatory variables. To obtain a clear quantitative interpretation we calculate APEs. We focus on the gender gradient and persistence in self-reported life satisfaction.

The pooled heteroskedastic ordered probit APEs, are estimated by taking first differences of the expected value of expression (6) with respect to gender (in $\boldsymbol{x}_{it}$) and $\boldsymbol{y}_{it-1}$ (see Wooldridge, 2019). While the heteroskedastic probit estimator parameterizes the unobserved heterogeneity, pooled estimation sets $\zeta_i = 0$. Therefore, the pooled heteroskedastic APEs and their corresponding standard errors (computed via the Delta-method) can be obtained using standard software such as Stata without additional programming.

Regarding the panel RE ordered probit APEs, we estimate first differences of the expected value of expression (7) for gender and $\boldsymbol{y}_{it-1}$ with respect to the distribution of $(\boldsymbol{y}_{i1}, \bar{\boldsymbol{x}}_i)$. Computation requires making an assumption about the unobserved heterogeneity ζ_i. One approach is to average the partial effects over the distribution of unobserved heterogeneity and use population averaged parameters $\boldsymbol{b}_{\xi} = \boldsymbol{b} / \sqrt{(1 + \sigma_{\zeta}^2)}$, where $\boldsymbol{b}$ denotes the vector of estimated parameters (see Wooldridge, 2005).

Life satisfaction thresholds might vary across individuals due to unobserved time-invariant characteristics such personality or cognitive ability. Differences in reporting behaviour that are a function of unobserved individual attributes can influence evaluation and perception of life satisfaction. For example, at a given level of the latent life satisfaction index pessimists may consistently report lower satisfaction levels (see Jones and Schurer, 2011; Angelini *et al.*, 2014). To accommodate this, computation of panel RE ordered probit partial effects can take into account the unobserved effect ζ_i by drawing from its estimated distribution (see Jones and Schurer, 2011):

$$\Phi\left(\mu_j - \boldsymbol{x}_{it}\boldsymbol{\beta} - \boldsymbol{y}_{it-1}\boldsymbol{\gamma} - \boldsymbol{y}_{i1}\boldsymbol{\vartheta}_1 - \bar{\boldsymbol{x}}_i\boldsymbol{a} - \zeta_i\right) - \Phi\left(\mu_{j-1} - \boldsymbol{x}_{it}\boldsymbol{\beta} - \boldsymbol{y}_{it-1}\boldsymbol{\gamma} - \boldsymbol{y}_{i1}\boldsymbol{\vartheta}_1 - \bar{\boldsymbol{x}}_i\boldsymbol{a} - \zeta_i\right) \quad (7)$$

We approximate $\hat{\zeta}_i$ by

$$\hat{\zeta}_i \simeq \Phi^{-1}\left(\bar{y}_{i,j}\right) - \left(\bar{\boldsymbol{x}}_{it}\hat{\boldsymbol{\beta}} + \bar{\boldsymbol{y}}_{it-1}\hat{\boldsymbol{\gamma}} + \bar{\boldsymbol{y}}_{i1}\hat{\boldsymbol{\vartheta}}_1 + \bar{\boldsymbol{x}}_i\hat{\boldsymbol{a}}\right) \quad (8)$$

where Φ^{-1} denotes the inverse of the standard normal *cdf*, $\bar{y}_{i,j}$ is the individual-specific sample average obtained by dichotomizing life satisfaction at threshold j and $(\bar{\boldsymbol{x}}_{it}, \bar{\boldsymbol{y}}_{t-1}, \bar{\boldsymbol{y}}_{i1}, \bar{\boldsymbol{x}}_i)$ are individual-specific sample averages of the covariates included in the CRE models. Since the RE APEs are functions of the estimated parameters and subject to sampling variability, we compute bootstrapped standard errors using 1,500 bootstrap replications by resampling with replacement accounting for individual-level clustering.

V. Estimation results

The estimations indicate a notable positive male gender gradient in self-reported life satisfaction across all 12 cohorts. The estimated coefficients on the male gender dummy are positive and globally statistically significant. We focus on the CRE heteroskedastic ordered probit estimates provided in Tables 1 and 2 since they explicitly model panel unbalancedness, parameterize the unobserved individual effect, incorporate dynamics and initial conditions and provide partial effects for each level of self-reported life satisfaction. The positive male gender gradient result is robust to a series of robustness checks using panel CRE ordered probits (Tables 21 and 22), flexible heteroskedastic CRE ordered models (Table 23), as well as, specifications employing the original seven-point scale life satisfaction variable estimated via heteroskedastic CRE ordered probits (Tables 28 and 29) and linear generalized least squares CRE (Tables 30 and 31). Finally, the gender gradient is robust to considering an alternative five-point scale by adding the neither satisfied/dissatisfied category and grouping together the lower three satisfaction levels (Tables 32 and 33) – see online Appendix for all alternative estimations (in Tables 21–38).[12]

The gender APEs for the baseline heteroskedadstic CRE models accounting for unbalancedness are given in Table 3. The estimated gender APEs provide a quantitative comparison of the gender gradient across cohorts and life satisfaction thresholds. A clear pattern is apparent regarding all cohorts: male adolescents are ceteris paribus less likely to report lower levels of life satisfaction and more likely to report complete satisfaction. The estimated male gender APEs are monotonically increasing in life satisfaction and globally statistically significant excluding the third threshold (Very Satisfied, Life Satisfaction $\in$ (2,3]). Hence, the probability that a male adolescent reports a higher level of life satisfaction, as opposed to a female with otherwise identical observable and unobservable attributes, is higher the higher the level of life satisfaction. The flexible heteroskedastic CRE variant APEs, adding $\mathbf{1}[T_i = r]\bar{\boldsymbol{x}}_i$ to the variance function and the panel CRE variant averaging the APEs across the distribution of the unobserved heterogeneity conduce to similar conclusions (see respective Tables 41–42 and 43–44, online Appendix).[13,14]

Additional APEs have been computed for the panel CRE variant by drawing from the estimated distribution of the unobserved heterogeneity (in Table 4) using equation (8). Incorporating the approximated unobserved effect generally seems to inflate the gender partial effects at the bottom and deflate the impact of gender at the top of the life satisfaction distribution. The corresponding APEs in Table 4 generally increase the negative male probability to report life dissatisfaction (Dissatisfied, Life Satisfaction = 1). Furthermore, unlike the heteroskedastic/averaged over the unobserved heterogeneity distribution APEs, Table 4 reveals that the probability of males reporting being very satisfied (Life Satisfaction $\in$ (2,3]) is higher than reporting complete satisfaction in the cohorts spanning 1996–2001, 2004–17 noting that the respective complete satisfaction gender APEs generally become statistically insignificant.

According to the baseline APEs (in Table 3) in all cohorts spanning 1994–2017 male adolescents are less likely to report life dissatisfaction by approximately 3.4%–6.8% and mere satisfaction by around (2.4-5). On the other hand, male adolescents are more probable to report complete life satisfaction by approximately 5.7%–13.6%

[12]Note that regarding the 1997 and 1998 cohort estimates using the original 7-point scale satisfaction the initial period observations for the very unhappy category are zero- see (Table 28, online Appendix).

[13]Tables 39–48, online Appendix provide all variants' estimated APEs for gender and lagged life satisfaction.

[14]The alternative 5-point scale heteroskedastic CRE estimates give very similar APEs for the top three categories and the lowest category gender impact from the 4-point scale is split between the lowest two categories in the 5-point scale estimates (see Tables 47 and 48, online Appendix).

TABLE 3

Adolescent life satisfaction, cohorts 1996–2013, gender (male) APE, heteroskedastic CRE ordered probits

	(1) *1996–2000*	*(2)* *1997–2001*	*(3)* *1998–2002*	*(4)* *1999–2003*	*(5)* *2000–2004*	*(6)* *2001–2005*
Dissatisfied, Life Satisfaction = 1	−0.0337***	−0.0507***	−0.0677***	−0.0528***	−0.0543***	−0.0457***
	(0.0120)	(0.0128)	(0.0132)	(0.0139)	(0.0147)	(0.0154)
Satisfied, Life Satisfaction ∈ (1,2]	−0.0292***	−0.0411***	−0.0495***	−0.0357***	−0.0350***	−0.0330***
	(0.0102)	(0.0103)	(0.00932)	(0.00980)	(0.00943)	(0.0118)
Very Satisfied, Life Satisfaction ∈ (2,3]	−0.00691**	−0.0147***	−0.0188***	−0.00596*	−0.00534	−0.00286
	(0.00311)	(0.00489)	(0.00493)	(0.00353)	(0.00351)	(0.00326)
Completely Satisfied, Life Satisfaction > 3	0.0698***	0.107***	0.136***	0.0945***	0.0947***	0.0816***
	(0.0241)	(0.0259)	(0.0240)	(0.0251)	(0.0253)	(0.0282)
Number of Observations	1081	1073	1155	1143	1145	1087
	(1) *2002–2006*	*(2)* *2003–2007*	*(3)* *2004–2008*	*(4)* *2011–2015*	*(5)* *2012–2016*	*(6)* *2013–2017*
Dissatisfied, Life Satisfaction = 1	−0.0365***	−0.0494***	−0.0433***	−0.0495***	−0.0347***	−0.0541***
	(0.0138)	(0.0142)	(0.0137)	(0.0103)	(0.0105)	(0.0117)
Satisfied, Life Satisfaction ∈ (1,2]	−0.0260***	−0.0361***	−0.0342***	−0.0355***	−0.0244***	−0.0382***
	(0.0100)	(0.0107)	(0.0109)	(0.00746)	(0.00735)	(0.00848)
Very Satisfied, Life Satisfaction ∈ (2,3]	0.000784	0.00453	0.00103	0.00240	0.00230	0.0133***
	(0.00189)	(0.00292)	(0.00261)	(0.00209)	(0.00162)	(0.00396)
Completely Satisfied, Life Satisfaction > 3	0.0617***	0.0810***	0.0765***	0.0826***	0.0568***	0.0790***
	(0.0235)	(0.0229)	(0.0235)	(0.0166)	(0.0168)	(0.0167)
Number of Observations	1247	1083	1034	1840	1668	1508

Notes: ***, **, * is statistical significance at 1%, 5%, 10%, respectively. Source: (BHPS: Cohorts 1996-2004), (UKHLS: Cohorts 2011–13). Compact Unbalanced Panels. Delta-method standard errors (in parentheses).

TABLE 4

Adolescent life satisfaction, cohorts 1996–2013, gender (male) APE: approximating the unobserved effect by drawing from its estimated distribution, panel CRE ordered probits

	(1) *1996–2000*	*(2)* *1997–2001*	*(3)* *1998–2002*	*(4)* *1999–2003*	*(5)* *2000–2004*	*(6)* *2001–2005*
Dissatisfied, Life Satisfaction = 1	−0.0872**	−0.1312***	−0.1257***	−0.0879***	−0.0662**	−0.0478*
	(0.0353)	(0.0435)	(0.0396)	(0.0291)	(0.0272)	(0.0247)
Satisfied, Life Satisfaction ∈ (1,2]	0.0172	0.0079	−0.0164	−0.0091	−0.0228*	−0.0352***
	(0.0246)	(0.0368)	(0.0298)	(0.0203)	(0.0128)	(0.0130)
Very Satisfied, Life Satisfaction ∈ (2,3]	0.0495**	0.0768**	0.0461	0.0316	0.0020	−0.0212
	(0.0236)	(0.0362)	(0.0457)	(0.0268)	(0.0237)	(0.0243)
Completely Satisfied, Life Satisfaction > 3	0.0307	0.0597*	0.0977***	0.0637**	0.0739**	0.0864***
	(0.0214)	(0.0347)	(0.0352)	(0.0304)	(0.0295)	(0.0322)
Number of Observations	1081	1073	1155	1143	1145	1087
	(1) *2002–2006*	*(2)* *2003–2007*	*(3)* *2004–2008*	*(4)* *2011–2015*	*(5)* *2012–2016*	*(6)* *2013–2017*
Dissatisfied, Life Satisfaction = 1	−0.0420	−0.0590*	−0.1317***	−0.1327***	−0.1115***	−0.1198***
	(0.0266)	(0.0345)	(0.0502)	(0.0330)	(0.0346)	(0.0301)
Satisfied, Life Satisfaction ∈ (1,2]	−0.0367	−0.0558**	−0.0171	0.0864***	0.0641***	0.0654***
	(0.0231)	(0.0248)	(0.0504)	(0.0190)	(0.0225)	(0.0202)
Very Satisfied, Life Satisfaction ∈ (2,3]	0.0159	0.0221	0.0950**	0.0493**	0.0501**	0.0520**
	(0.0328)	(0.0458)	(0.0437)	(0.0218)	(0.0219)	(0.0205)
Completely Satisfied, Life Satisfaction > 3	0.0524	0.0863**	0.0534	0.0034	0.0045	0.0057
	(0.0347)	(0.0362)	(0.0403)	(0.0029)	(0.0040)	(0.0049)
Number of Observations	1247	1083	1034	1840	1668	1508

Notes: ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels respectively. Source: (BHPS: Cohorts 1996–2004), (UKHLS: Cohorts 2011–13). Compact Unbalanced Panels. Bootstrapped standard errors (in parentheses) accounting for individual-level clustering (1,500 replications). Estimated model parameters given in Tables 21 and 22, online Appendix.

across the 12 cohorts studied. The gender APEs for the third threshold (very satisfied) are generally low and insignificant.

Therefore, the gender gradient in adolescent life satisfaction is an increasing function of satisfaction levels and it is greater at the two extremes of the life satisfaction distribution. This is in line with Contoyannis and Li (2017) concluding that the (CES-D) depression gender gap is larger at more severe depression levels (using the US CNLSY79 survey, ages 15–25).

To formally test for state dependence, we estimate dynamic models including dummy variables representing one-period lags of the categories of life satisfaction. There is a gradient across the estimated coefficients of previous period life satisfaction status (see Tables 1–2) with the top category (completely satisfied(t-1)) entering with the highest coefficient magnitudes and strongest statistical significance (the base category is 'dissatisfied'). This gradient pattern is reflected in the respective APEs noting that the lagged variables' impact is generally highest at the upper and lower extreme values of life satisfaction in terms of both magnitude and statistical significance (see Tables 39 and 40, online Appendix). A notable exception is the 2001 cohort where lagged life satisfaction controls and the corresponding APEs are statistically insignificant.

Similarly, there is a positive gradient in the estimated coefficients of initial period self-reported life satisfaction though, statistical significance is more prominent regarding the top category (completely satisfied ($t = 1$)) in the panel CRE estimations (Tables 21 and 22, online Appendix) and to a lesser extent in the baseline heteroskedastic CRE models (see Tables 1 and 2). This indicates a positive correlation between initial period observations and the unobserved heterogeneity.

The quality of family environment and peer interaction relate to self-reported life satisfaction in a predictable manner. In particular, lower maternal conversation frequency and fighting with someone (involving physical violence) generally have negative statistically significant coefficients. Smoking/having smoked enters with negative significant coefficients in some cohort estimations.[15]

The estimates reveal a North–South divide in self-reported life satisfaction. Holding per head household income (and all other covariates) fixed, living in a higher income per capita English region (London, South East, South West, East England) reduces adolescent life satisfaction. The regional control is negative across all 12 cohorts though statistical significance is restricted to the five cohorts spanning 2002–16 (see Tables 1 and 2). Oswald and Wu (2011) conclude similarly using US adult data invoking compensating differentials theoretical predictions (i.e. higher nominal remuneration in richer regions, but generally worse remaining factors such as property prices and traffic congestion). Controlling for household income per capita, the regional dummy provides information about the remaining intrinsic regional

[15]We exclude arguing frequencies with parents as they are collinear with parental conversation frequencies and statistically insignificant. Frequency of fights with someone in the past month take on values 1 (none) to 5 (two times or more). Parental conversation frequencies (Not talking to Mum/Dad) are derived from questions 'How often do you talk to your mother/father, about things that matter to you?'. The last two variables take the value of one if the response was 'hardly ever, don't have a mother/father' and zero otherwise (most days, more than once a week, less than once a week).

disamenities for which compensating higher pay must be offered (Oswald and Wu, 2011, p. 1122).[16]

Nature versus nurture?

Though there is no support of sexually dimorphic human brains (see Joel *et al*., 2015) patrilineal society nurturing has been found to affect gender differences in spatial abilities (Hoffman, Gneezy and List, 2011). Ingrained stereotypical beliefs regarding sex discrepancies in terms of affective and cognitive abilities can predetermine human behaviour and parenting styles. An important question is whether the gender life satisfaction gap is still observed when considering female adolescents growing up in the same household with opposite sex siblings.

A Fawcett Society 2017 report states ‘Young women told us that they are exposed to gender norms from a wide range of sources: the media pressuring them to look or act a certain way, family members in their own home expecting them to clean and cook for male family members, parents requiring them to come home earlier at night than their male siblings or even warning them not to become too rich/powerful in case their success deters a potential husband’.[17]

Using the household identification, we can determine whether an adolescent in the sample resides in a household with a sibling of the opposite sex. Accordingly, an interaction variable between female gender and having at least one opposite sex sibling in the household (Female x Brother in Household) is added to the covariates. If the estimated gender gap persists and an intra-household gender gap is additionally revealed, this could be at least partially attributed to differential parenting styles shaped by gender stereotypes (e.g. girls inadvertently receiving indirect albeit salient hints of limited future opportunities).

The respective estimates, in Table 5, produce a statistically insignificant impact for the ‘Female x Brother in Household’ interaction (except in the 1997 cohort) noting that the number of female adolescents with male siblings identified in the samples in small (1.9%–4.4% of the observations across cohorts).[18] Quite crucially, the male gender gradient remains positive, statistically significant and is only marginally reduced in magnitude. Consequently, there is an intra-household adolescent life satisfaction gender gap. This hints, at least to some extent, to gender-differential parenting styles and nurturing preconditioned by stereotypical beliefs. However, given our inability to

[16]We use the first-level Classification of Territorial Units for Statistics (NUTS) of the EU to group together the wealthiest regions in terms of GVA per capita. During the period analysed London and the South East are above the national average GVA per capita while the East of England and South West follow. Scotland is close to the average but, we aggregate high income English regions due to geographical proximity – see https://www.ons.gov.uk/economy/grossvalueaddedgva/bulletins/regionalgrossvalueaddedincomeapproach/ December 2016. Using ‘Scotland, London, South East, South West, East of England’ instead, produces similar estimates changing the statistical significance of the 2003–07 cohort only (see Tables 35 and 36, online Appendix).

[17]see https://www.fawcettsociety.org.uk/sounds-familiar

[18]The BHPS and UKHLS youth questionnaires do not contain information about siblings and thus, siblings can only be matched via the household wave-specific identifier if they fall within the 11–15 and 10–15 age range in the two datasets, correspondingly.

TABLE 5

Adolescent life satisfaction, cohorts 1996–2013, heteroskedastic CRE ordered probits

	(1) *1996–2000*	*(2)* *1997–2001*	*(3)* *1998–2002*	*(4)* *1999–2003*	*(5)* *2000–2004*	*(6)* *2001–2005*
Male	0.2136***	0.3289***	0.3801***	0.2602***	0.2848***	0.2523***
	(0.0808)	(0.0917)	(0.0689)	(0.0670)	(0.0823)	(0.0857)
Female x Brother in Household	−0.1916	−0.3600**	0.0178	0.0385	−0.1638	0.1364
	(0.1866)	(0.1640)	(0.1783)	(0.1839)	(0.2193)	(0.2096)
Log-Likelihood	−1233.044	−1210.250	−1332.943	−1397.045	−1437.714	−1368.802
Number of Observations	1081	1073	1155	1143	1145	1087
Wald (Global Significance)	189.122	158.944	165.511	144.426	80.836	56.824
LR: $ln(\sigma_\zeta)$	1.573	1.519	2.938	11.189	4.360	1.406
P-value: $ln(\sigma_\zeta)$	0.456	0.468	0.230	0.004	0.113	0.495
	(1) *2002–2006*	*(2)* *2003–2007*	*(3)* *2004–2008*	*(4)* *2011–2015*	*(5)* *2012–2016*	*(6)* *2013–2017*
Male	0.2018***	0.2934***	0.2573***	0.3394***	0.2266***	0.3192***
	(0.0773)	(0.0904)	(0.0854)	(0.0673)	(0.0672)	(0.0719)
Female x Brother in Household	0.0289	−0.0948	−0.1959	0.3291	0.0334	0.0806
	(0.2087)	(0.1975)	(0.1980)	(0.2121)	(0.2081)	(0.1726)
Log-Likelihood	−1457.442	−1283.780	−1178.739	−2117.036	−1913.004	−1784.634
Number of Observations	1247	1083	1034	1840	1668	1508
Wald (Global Significance)	222.432	139.861	166.385	228.957	215.024	205.327
LR: $ln(\sigma_\zeta)$	2.670	2.962	5.423	5.252	6.305	5.559
P-value: $ln(\sigma_\zeta)$	0.263	0.227	0.066	0.072	0.043	0.062

Notes: ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels respectively. Source: (BHPS: Cohorts 1996–2004), (UKHLS: Cohorts 2011–13). Compact Unbalanced Panels. Individual cluster-robust standard errors (in parentheses). Estimations additionally include all remaining covariates in Tables 1 and 2.

control for the onset of female menstruation, the paucity of parental controls and the unavailability of biomedical variables this is just an assertion.

The gender-differential impact of previous period life satisfaction

Control variables have been assumed to play the same predictive role for both genders. To explore which covariates underlie the gender adolescent life-satisfaction gap we interact all controls with the male gender dummy. We begin by interacting all explanatory variables with the male dummy. This excludes year dummy interactions as these are collinear with the gender dummy. The corresponding results (see online Appendix, Tables 26 and 27) reveal that the interaction between the male gender dummy and past-period self-reported life satisfaction level is the most significant interaction across all cohorts (excluding the 2000 and 2001 cohorts). In order to isolate this impact, Table 6 reports the estimates obtained by interacting lagged satisfaction levels by gender and retaining all other covariates in their original form. This outcome indicates that previous year life satisfaction level is a strong determinant of prospective year self-reported adolescent male satisfaction. One channel could be lower male adolescent emotional intelligence producing a retarded lagged response in perceiving and understanding emotions relating to life satisfaction. Emotional intelligence (EI) involves the ability to perceive and accurately express emotion, to use emotion to facilitate reasoning, to comprehend emotions, and to manage emotions for emotional growth. Lower male EI has been documented in some studies (see Brackett, Mayer and Warner, 2004; Cabello *et al*., 2016). However, the BHPS and UKHLS youth questionnaires do not include an EI test therefore, limiting our ability to investigate this channel.

The gender-differential impact of past period satisfaction may reflect differences in the reporting life satisfaction scales (termed as differential item functioning). Scale differences may reflect gender-specific characteristics such that the two genders use systematically different norms concerning the interpretation of completely happy, neither happy/unhappy or completely unhappy (Hsee and Tang, 2007). Scale differences due to individual-specific time-invariant characteristics can be accounted for by computing APEs incorporating the approximated unobserved effect by drawing from its estimated distribution (see equations 7 and 8). To eliminate potential bias from differential item functioning and account for differences in gender reporting styles, a vignette methodology would be preferable (see Angelini *et al*., 2014). This approach requires a second additional set of questions using anchoring vignettes prompting individuals to evaluate the life satisfaction level of one or more hypothetical individuals described in particular states. Unfortunately, this cannot be explored since the BHPS and UKHLS do not include vignettes and youth respondents only rate their personal satisfaction.

The mean life satisfaction trend across time: do alternative cardinalizations reverse the gender gradient?

Ordered probit estimation assumes a single cardinalization under which life satisfaction is normally distributed with equal variance for both genders. We investigate the latter

TABLE 6

Adolescent life satisfaction, cohorts 1996–2013, gender interacted lags, heteroskedastic CRE ordered probits

	(1) *1996–2000*	*(2)* *1997–2001*	*(3)* *1998–2002*	*(4)* *1999–2003*	*(5)* *2000–2004*	*(6)* *2001–2005*
Male x Completely Satisfied (t−1)	0.5657***	0.8935***	0.5209***	0.8571***	0.2972	0.2930
	(0.2066)	(0.2390)	(0.1953)	(0.1693)	(0.1818)	(0.2421)
Male x Very Satisfied (t−1)	0.2577	0.3777	0.1215	0.4879***	−0.0094	−0.0365
	(0.2143)	(0.2331)	(0.1869)	(0.1526)	(0.1646)	(0.2329)
Male x Satisfied (t−1)	−0.1622	0.0514	−0.0309	0.3850**	0.1099	0.0411
	(0.2168)	(0.2379)	(0.2051)	(0.1855)	(0.2028)	(0.2775)
Male	−0.0472	−0.0945	0.1975	−0.2518*	0.2131	0.2186
	(0.1994)	(0.2087)	(0.1787)	(0.1447)	(0.1604)	(0.2289)
$Variance function: ln(\sigma_\zeta)$						
Male	−0.0459	0.0684	0.0080	0.0373	0.0693	0.1778**
	(0.0831)	(0.0896)	(0.0792)	(0.0765)	(0.0742)	(0.0763)
$T_i = 3$	0.1100	0.1108	−0.1268	−0.2818***	0.1627*	0.1140
	(0.0972)	(0.1038)	(0.1025)	(0.0977)	(0.0909)	(0.0929)
$T_i = 4$	0.0735	0.0659	−0.1206	−0.0318	0.1598*	0.0094
	(0.0931)	(0.1002)	(0.0901)	(0.0803)	(0.0819)	(0.0939)
Log-Likelihood	−1262.217	−1238.667	−1351.352	−1406.876	−1437.906	−1364.536
Number of Observations	1081	1073	1155	1143	1145	1087
Wald (Global Significance)	104.134	88.055	95.688	91.550	67.190	57.140
LR: $ln(\sigma_\zeta)$	1.715	1.688	2.503	9.240	5.204	6.940
P-value: $ln(\sigma_\zeta)$	0.634	0.640	0.475	0.026	0.157	0.074

	(1) *2002–2006*	*(2)* *2003–2007*	*(3)* *2004–2008*	*(4)* *2011–2015*	*(5)* *2012–2016*	*(6)* *2013–2017*
Male x Completely Satisfied (t−1)	0.8221***	0.8033***	1.1393***	0.6427***	0.7444***	0.7126***
	(0.1844)	(0.2234)	(0.2071)	(0.1920)	(0.1897)	(0.1995)
Male x Very Satisfied (t−1)	0.4636***	0.2904	0.4648***	0.3060*	0.3916**	0.3279*
	(0.1745)	(0.2251)	(0.1770)	(0.1764)	(0.1676)	(0.1726)

(*Continued*)

TABLE 6
(*Continued*)

	(1) 2002–2006	(2) 2003–2007	(3) 2004–2008	(4) 2011–2015	(5) 2012–2016	(6) 2013–2017
Male x Satisfied(t-1)	−0.0356	−0.0829	0.2503	−0.1025	0.1418	0.0087
	(0.1703)	(0.2178)	(0.1958)	(0.1822)	(0.1845)	(0.1891)
Male	−0.2262	−0.0262	−0.3128*	0.0074	−0.2078	−0.0011
	(0.1786)	(0.2116)	(0.1857)	(0.1781)	(0.1704)	(0.1752)
$Variance function : ln(\sigma_{\zeta})$						
Male	−0.0075	−0.0393	−0.0828	−0.1247**	−0.1648**	−0.1441**
	(0.0806)	(0.0836)	(0.0896)	(0.0624)	(0.0695)	(0.0700)
$T_i = 3$	−0.1968**	0.1513	0.1571*	0.0953	0.1595*	0.1666**
	(0.0928)	(0.1099)	(0.0943)	(0.0754)	(0.0830)	(0.0837)
$T_i = 4$	−0.0541	0.2174**	−0.1004	0.1199	0.0526	0.0843
	(0.0879)	(0.0932)	(0.0929)	(0.0801)	(0.0879)	(0.0857)
Log-Likelihood	−1510.648	−1317.195	−1212.355	−2204.495	−2004.391	−1851.750
Number of Observations	1247	1083	1034	1840	1668	1508
Wald (Global Significance)	119.957	92.395	103.402	118.074	133.091	97.233
LR: $ln(\sigma_{\zeta})$	4.741	6.103	8.660	5.942	8.515	8.846
P-value: $ln(\sigma_{\zeta})$	0.192	0.107	0.034	0.114	0.036	0.031

Notes: ***, **, and * indicate statistical significance at the 1%, 5% and 10% levels respectively. Source: (BHPS: Cohorts 1996–2004), (UKHLS: Cohorts 2011–13). Compact Unbalanced Panels. Individual cluster-robust standard errors (in parentheses). Estimations additionally include all remaining covariates in Tables 1 and 2.

assumption by adding gender to the variance function of the heteroskedastic CRE ordered probits and also estimate a variant including gender only in the variance function. The male gender dummy only enters the 2001 cohort variance function with a positive statistically significant effect and does not generally have a significant impact on the estimated male gender coefficient estimate (see Tables 24 and 25, online Appendix). However, the compact unbalanced panel per pseudo-cohort design does not permit analysing the life satisfaction gender gradient trend across the duration of the two datasets.

We use heteroskedastic ordered probits containing gender only to estimate the mean and standard deviation of life satisfaction by gender for each year assuming normality and test whether the standard deviation is constant (male adolescents in the initial youth respondent survey wave, 1994 in the BHPS and 2009 in the UKHLS, are the control group) – see Bond and Lang (2019). The LR test for variance equality is $\chi^2(29) = 35.34$ with a $P = 0.1934$ in the BHPS (1994–2008) and $\chi^2(17) = 45.45$ with a $P = 0.0002$ in the UKHLS (2009–17). We therefore reject equality of the variances in the UKHLS but not, in the BHPS data set.

In Figures (13–16), we plot the means and standard deviations of male and female life satisfaction per survey wave for the BHPS (1994–2008) and UKHLS (2009–17)

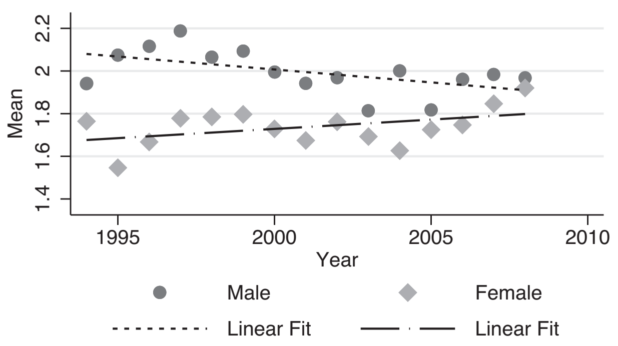

Figure 13. Mean life satisfaction over time under normality

Figure 14. Standard deviation of life satisfaction over time under normality

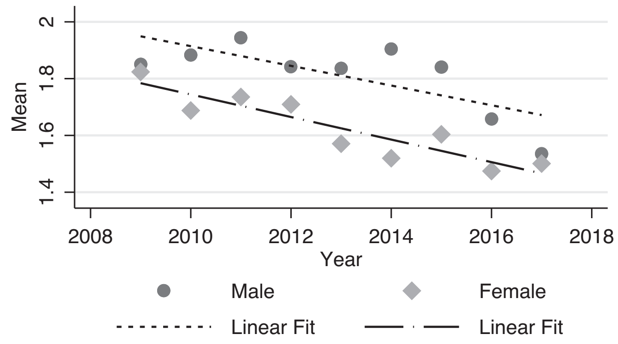

Figure 15. Mean life satisfaction over time under normality

Figure 16. Standard deviation of life satisfaction over time under normality

data sets respectively. The gender life satisfaction gap is evident in (Figures 13 and 15) where mean male satisfaction rates are consistently above female satisfaction rates across all survey waves analysed. The yearly mean of male life satisfaction falls and the female rises such that they come closer to convergence by the final years of the BHPS while, a steady decline in mean satisfaction levels is observed for both genders in the UKHLS. The standard deviation declines across time in the BHPS for both genders though, the linear trend is steeper in the female case (see Figure 14). A notable divergence among the increasing female standard deviation and the slightly diminishing male standard deviation occurs across time in the UKHLS (see Figure 16).

We consider alternative cardinalizations (monotonic transformations) of the ordinal life satisfaction variable, that can reverse the conclusions reached by the normal (see Bond and Lang, 2019). If the male mean and standard deviations of life satisfaction across time are greater than the corresponding female location and scale parameters ($LS\mu_m > LS\mu_f$ and $LS\sigma_m > LS\sigma_f$) as occurs in the BHPS, the gender gradient can be reversed by a concave transformation such as the left-skewed log-normal as in Figure 17. Conversely, in the UKHLS $LS\mu_m > LS\mu_f$ and $LS\sigma_m < LS\sigma_f$ such that the ordering can be preserved by all concave transformations but, not convex transformations as the right-skewed log-normal that reverses the gender gradient as in Figure 18. Note that this approach has three important limitations. First, the

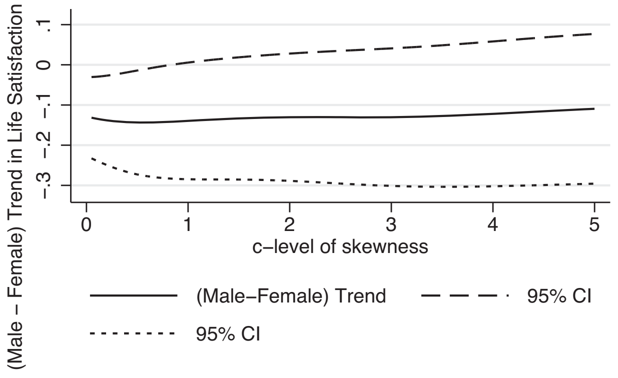

Figure 17. Male-female life satisfaction difference trend for left-skewed log-normal

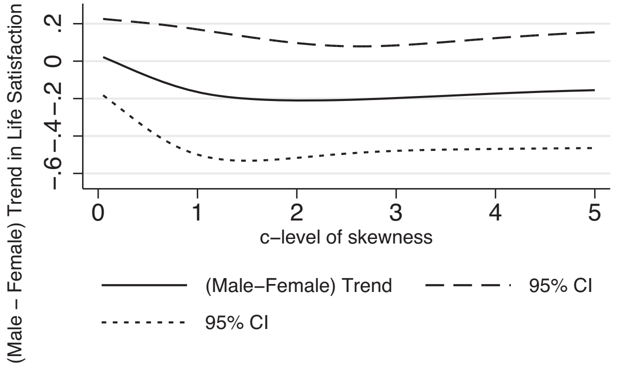

Figure 18. Male-female life satisfaction difference trend for right-skewed log-normal

transformations considered assume time invariance of the skewness of the distribution, measured by the skewness parameter $c = |2(\mathrm{LS}\mu_m - \mathrm{LS}\mu_f)/(\mathrm{LS}\sigma_f^2 - \mathrm{LS}\sigma_m^2)|$. Second, it treats all successive pseudo-cohorts of adolescents as a single dataset of pooled cross-sections and assumes that male and female adolescents report their life satisfaction in the same manner across cohorts and over time. Finally, while a comparison of the means can be informative regarding the gender gradient across time, the APEs reported in Tables 3 and 4 illustrate that considering the middle and ignoring the differences at the two extremes of the life satisfaction distribution can lead to misleading conclusions.

VI. Summary and conclusions

We investigate the cross-cohort gender gradient and socioeconomic determinants of adolescent self-reported life satisfaction in Great Britain. The study is unique in exploiting all possible formations of youth cohorts of individuals aged 11–15 from the BHPS and the UKHLS datasets spanning 1996–2017. We provide robust evidence of a cross-cohort gender gap in life satisfaction. Female adolescents are significantly more likely to report life dissatisfaction and male adolescents more probable to report complete satisfaction. This may indicate a higher adolescent female depression risk, if lower satisfaction predicts depression and the gap is not a product of gender-differential reporting, noting that the extant evidence is inconclusive (e.g. Moksnes *et al*., 2016; Gigantesco *et al*., 2019). Establishing a life satisfaction and depression link would require estimation of joint determination models for depressive symptoms and life satisfaction.

There is a notable positive male gender gradient in self-reported life satisfaction across all 12 cohorts analysed. The estimated coefficients on the male gender gradient are positive and globally statistically significant. To establish robustness, several specifications are estimated: heteroskedastic CRE ordered probits accounting for unbalancedness (considering various variance functions for panel unbalancedness), models using the original 7-point scale, an alternative 5-point scale of the self-reported life satisfaction variable (as opposed to the 4-point scale in the remaining models), panel CRE ordered probits and CRE linear GLS treating life satisfaction as a cardinal measure. The positive male gender gradient is robust and universally statistically significant across all alternative specifications.

Additionally accounting for female adolescents growing up with male siblings, we find evidence of an intra-household gender gap while the overall female life satisfaction gap still persists. This could indicate ingrained gender stereotypes preconditioning child nurturing but, data paucity and lack of biomedical controls preclude testing.

There is significant state dependence in life satisfaction outcomes, even after controlling for unobserved individual heterogeneity. The positive correlation between initial period observations and the unobserved heterogeneity hints to a preconditioning of life satisfaction by time-invariant individual attributes such as neurobiological/genetic factors that we cannot control for due to data unavailability. Hence, initial conditions and past period satisfaction levels predetermine life satisfaction propensity and future satisfaction levels respectively. Interacting all covariates with gender we find a gender differential impact of past period life satisfaction: previous period satisfaction levels are the strongest determinant of successive period self-reported adolescent male satisfaction levels. We posit that this could be either due to emotional intelligence discrepancies among the two genders or differential item functioning. In the latter case, gender-specific norms may produce different interpretations of life satisfaction scales meaning that, a vignette methodology would be appropriate (Hsee and Tang, 2007; Angelini *et al*., 2014). Unfortunately, the BHPS and UKHLS youth questionnaires do not include emotional intelligence tests and do not report vignettes since respondents only rate their personal life satisfaction.

The quality of family environment and peer interaction are positively associated with life satisfaction. Finally, the estimates indicate a North–South divide in adolescent life satisfaction. Controlling for household income, residing in a higher income region (South/East of England) is negatively related to self-reported adolescent life satisfaction – this is in line with Oswald and Wu (2011) and compensating-differentials theoretical predictions.

To analyse the life satisfaction gender gradient trend across the duration of each of the two datasets, we treat all successive pseudo-cohorts as a single dataset of pooled cross-sections. Mean male adolescent life satisfaction under normality is higher across time in both the BHPS and UKHLS. We employ alternative time-invariant cardinalizations that can reverse this conclusion (see Bond and Lang, 2019). However, the APEs estimated in our cohort-specific analysis clearly indicate distinct gender propensities at the two extremes of the life satisfaction distribution such that, considering the middle and ignoring the extremes can provide misleading conclusions.

In summary, the likelihood that a female adolescent reports a lower level of life satisfaction, compared to a male with otherwise identical observable and unobservable characteristics, is higher the lower the level of life satisfaction. The estimated APEs from the baseline heteroskedastic CRE models indicate that male adolescents are approximately 6%–14% more likely to report complete life satisfaction and 3%–7% less likely to report dissatisfaction. Incorporating the unobserved effect using its estimated distribution, seems to inflate the gender impact at the bottom and deflate it at the top of the life satisfaction distribution.

High previous period satisfaction levels are the strongest determinant of high prospective adolescent life satisfaction. Repeatedly reporting life dissatisfaction is indicative of either current, or future high propensity towards, depression which is a psychiatric/mental health disorder. Consequently, early-life interventions might be necessary to reduce adult depression prevalence.

Interventions addressing the adolescent gender life satisfaction gap, and targeting adolescents from both genders persistently reporting low satisfaction, can lead to substantive later life outcome improvements. Life cycle skill formation is a dynamic process with early inputs preconditioning future productivity (Heckman, 2006, p. 1900). Heckman (2012) notes that 'health economists should consider the costs and benefits of preventing rather than treating' and the present study offers a clear policy direction. Prevention efforts should be targeted towards raising parental, school teacher and public awareness regarding the prevalence of persistent life dissatisfaction and the higher adolescent female vulnerability to depression. Finally, the data limitations faced here are yet another reminder of the need to incorporate biomedical information, personality tests and vignettes alongside self-reported subjective well-being measures in multi-purpose panels with a social science orientation.

Final Manuscript Received: April 2021

References

Albarran, P., Carrasco, R. and Carro, J. M. (2019). 'Estimation of dynamic nonlinear random effects models with unbalanced panels', *Oxford Bulletin of Economics and Statistics*, Vol. 81, pp. 1424–1441.

Altemus, M., Sarvaiya, N. and Epperson, C. N. B. (2014). 'Sex differences in anxiety and depression clinical perspectives', *Frontiers in Neuroendocrinology*, Vol. 35, pp. 320–330.

Angold, A. and Worthman, C. W. (1993). 'Puberty onset of gender differences in rates of depression: a developmental, epidemiologic and neuroendocrine perspective', *Journal of Affective Disorders*, Vol. 29, pp. 145–158.

Angelini, V., Cavapozzi, D., Corazzini, L. and Paccagnella, O. (2014). 'Do danes and italians rate life satisfaction in the same way? using vignettes to correct for individual-specific scale biases', *Oxford Bulletin of Economics and Statistics*, Vol. 76, pp. 643–666.

Arulampalam, W. and Stewart, M. B. (2009). 'Simplified implementation of the heckman estimator of the dynamic probit model and a comparison with alternative estimators', *Oxford Bulletin of Economics and Statistics*, Vol. 71, pp. 659–681.

Arulampalam, W., Booth, A. L. and Taylor, M. P. (2000). 'Unemployment persistence', *Oxford Economic Papers*, Vol. 52, pp. 24–50.

Bergman, M. M. and Scott, J. (2001). 'Young adolescents' wellbeing and health-risk behaviours: gender and socio-economic differences', *Journal of Adolescence*, Vol. 24, pp. 183–197.

Blanchflower, D. G. and Oswald, A. J. (2004). 'Well-being over time in Britain and the USA', *Journal of Public Economics*, Vol. 88, pp. 1359–1386.

Blanchflower, D. G. and Oswald, A. J. (2008). 'Is well-being u-shaped over the life cycle?', *Social Science & Medicine*, Vol. 66, pp. 1733–1749.

Bond, T. N. and Lang, K. (2019). 'The sad truth about happiness scales', *Journal of Political Economy*, Vol. 127, pp. 1629–1640.

Brackett A. M., Mayer D. J. and Warner M. R. (2004). 'Emotional intelligence and its relation to everyday behaviour', *Personality and Individual Differences*, Vol. 36, pp. 1387–1402.

Cabello, R., Sorrel, M. A., Fernandez-Pinto, I., Extremera, N. and Fernandez-Berrocal, P. (2016). 'Age and gender differences in ability emotional intelligence in adults: a cross-sectional study', *Developmental Psychology*, Vol. 52, pp. 1486–1492.

Copeland W. E., Worthman C., Shanahan L., Costello E. J. and Angold, A.(2019) 'Early pubertal timing and testosterone associated with higher levels of adolescent depression in girls', *Journal of the American Academy of Child and Adolescent Psychiatry*, Vol. 58, pp. 1197–1206.

Chrysanthou, G. M. and Vasilakis, C. (2019). The Impact of Bullying Victimisation on Mental Wellbeing, IZA Discussion Paper No. 12206.

Contoyannis, P., Jones, M. A. and Rice, N. (2004). 'The dynamics of health in the british household panel survey', *Journal of Applied Econometrics*, Vol. 19, pp. 473–503.

Contoyannis, P. and Li, J. (2017). 'The dynamics of adolescent depression: an instrumental variable quantile regression with fixed effects approach', *Journal of the Royal Statistical Society Series A, Royal Statistical Society*, Vol. 180, pp. 907–922.

Cunha, F., Heckman, J. J. and Schennach, S. M. (2010). 'Estimating the technology of cognitive and noncognitive skill formation', *Econometrica*, Vol. 78, pp. 883–931.

Gigantesco, A., Fagnani, C., Toccaceli, V., Stazi, M. A., Lucidi, F., Violani, C. and Picardi, A. (2019). 'The relationship between satisfaction with life and depression symptoms by gender', *Frontiers in Psychiatry*, Vol. 10, 419, pp. 1–9.

Hamilton, J. L., Hamlat, E. J., Stange J. P., Abramson, L. Y. and Alloy, L. B. (2014). 'Pubertal timing and vulnerabilities to depression in early adolescence: differential pathways to depressive symptoms by sex', *Journal of Adolescence*, Vol. 37(2), pp. 165–174.

Heckman, J. J. (1981a). 'Statistical models for discrete panel data', in Manski, C. F. and McFadden, D. L., (eds), *Structural Analysis of Discrete Data and Econometric Applications*, Cambridge: The MIT Press, pp. 114–178.

Heckman, J. J. (1981b). 'The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process', in Manski, C. F. and McFadden, D. L., (eds), *Structural Analysis of Discrete Data and Econometric Applications*, Cambridge: The MIT Press, pp. 179–195.

Heckman, J. J. (2006). 'Skill formation and the economics of investing in disadvantaged children', *Science*, 312, pp. 1900–1902.

Heckman, J. J. (2012). 'The developmental origins of health', *Health Economics*, Vol. 21, pp. 24–29.

Hoffman, M., Gneezy, U. and List, J. A. (2011). 'Nurture affects gender differences in spatial abilities', *PNAS*, Vol. 108, pp. 14786–14788.

Hsee, C. K. and Tang, J. N. (2007). 'Sun and water: on a module-based measurement of happiness', *Emotion*, Vol. 7, pp. 213–218.

Hsiao, C. (2003). '*Analysis of Panel Data*', *Econometric Society Monographs*, Cambridge University Press, Cambridge, 2nd ed.

Joel, D., Berman, Z., Tavor, I., Wexler, N., Gaber, O., Stein, Y., Shefi, N., Pool, J., Urchs, S., Margulies, D. S., Liem, F., Hanggi, J., Jancke, L. and Assaf, Y. (2015). 'Sex beyond the genitalia: the human brain mosaic', *PNAS*, Vol. 12, pp. 15468–15473.

Jones, A. M. and Schurer, S. (2011). 'How does heterogeneity shape the socioeconomic gradient in health satisfaction?', *Journal of Applied Econometrics*, Vol. 26, pp. 549–714.

Layard, R., Clark, A. E., Cornaglia, F., Powdthavee, N. and Vernoit, J.(2014). 'What predicts a successful life? a life-course model of well-being', *Economic Journal*, Vol. 124, pp. 720–738.

Moksnes, U. K., Lohre, A., Lillefjell, M., Byrne, D. G. and Haugan, G.(2016). 'The association between school stress, life satisfaction and depressive symptoms in adolescents: life satisfaction as a potential mediator', *Social Indicators Research*, Vol. 125, pp. 339–357.

Mundlak, Y. (1978). 'On the pooling of time series and cross section data', *Econometrica*, Vol. 46, pp. 69–85.

Oswald, A. J. and Wu, S. (2011). 'Well-being across America', *The Review of Economics and Statistics*, Vol. 93, pp. 1118–1134.

Rabe-Hesketh, S. and Skrondal, A. (2013). 'Avoiding biased versions of wooldridge's simple solution to the initial conditions problem', *Economics Letters*, Vol. 120, pp. 346–349.

Rainville, J. R., Tsyglakova, M. and Hodes, G. E. (2018). 'Deciphering sex differences in the immune system and depression', *Frontiers in Neuroendocrinology*, Vol. 50, pp. 67–90.

Skrondal, A. and Rabe-Hesketh, S. (2014). 'Handling initial conditions and endogenous covariates in dynamic/transition models for binary data with unobserved heterogeneity', *Journal of the Royal Statistical Society Series C, Royal Statistical Society*, Vol. 63, pp. 211–237.

Stevenson, B. and Wolfers, J. (2009). 'The paradox of declining female happiness', *American Economic Journal: Economic Policy*, Vol. 1, pp. 190–225.

Van de Velde, S., Bracke, P. and Levecque, K. (2010). 'Gender differences in depression in 23 European countries. cross-national variation in the gender gap in depression', *Social Science & Medicine*, Vol. 71, pp. 305–313.

Van den Berg, G., Lundborg, P., Nystedt, P. and Rooth, D. (2014). 'Critical periods during childhood and adolescence', *Journal of the European Economic Association*, Vol. 12, pp. 1521–1557.

Wooldridge, J. M. (2005). 'Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved effects', *Journal of Applied Econometrics*, Vol. 20, pp. 39–54.

Wooldridge, J. M. (2019). 'Correlated random effects models with unbalanced panels', *Journal of Econometrics*, Vol. 211, pp. 137–150.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Supplementary Material. A Multiple Cohort Study of the Gender Gradient of Life Satisfaction during Adolescence: Longitudinal Evidence from Great Britain.