

Boyce C (2010) Understanding fixed effects in human well-being, *Journal of Economic Psychology*, 31 (1), pp. 1-16.

This is the peer reviewed version of this article

NOTICE: this is the author's version of a work that was accepted for publication in Journal of Economic Psychology. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Journal of Economic Psychology, [VOL 31, ISS 1 (2010)] DOI: <http://dx.doi.org/10.1016/j.joep.2009.08.009>

Understanding Fixed Effects in Human Well-Being

Christopher J. Boyce
Department of Psychology, University of Warwick
Email: c.j.boyce@warwick.ac.uk

Accepted for Publication in the *Journal of Economic Psychology*
[doi:10.1016/j.joep.2009.08.009](https://doi.org/10.1016/j.joep.2009.08.009)

Abstract

In studies of subjective well-being, economists and other researchers typically use a fixed or random effect estimation to control for unobservable heterogeneity across individuals. Such individual heterogeneity, although substantially reducing the estimated effect of many characteristics, is little understood. This paper shows that personality measures can account for 20% of this heterogeneity and a further 13% can be accounted for by other observable between-person information. This paper then demonstrates that the use of personality measures, in a new technique developed by Plumper and Troeger (2007), can help researchers obtain improved estimates for important characteristics such as marital status, disability and income. The paper argues that this has important practical implications.

Keywords: Life satisfaction, individual differences, personality, fixed effect vector decomposition, GSOEP

PsycINFO codes: 3120

JEL codes: D12

Corresponding author: c.j.boyce@warwick.ac.uk

Address: Department of Psychology, University of Warwick, Coventry CV4 7AL.

Telephone: (+44) 024765 23158

Acknowledgements: For helpful discussions and advice the author is grateful to Gordon Brown, Andrew Oswald, Mark Stewart, Alex Wood and two anonymous referees. The Economic and Social Research Council (ESRC) provided research support. The usual disclaimer applies. The data used was made available by the German Socio-Economic Panel Study (SOEP) at the German Institute for Economic Research (DIW Berlin). Neither the original collectors of the data nor the Archive bears any responsibility for the analyses or interpretations presented here.

Economists wishing to evaluate how economic circumstances benefit an individual's life are increasingly turning to subjective well-being data. There are now numerous studies by economists that reveal the benefits of, for example, more income (Blanchflower & Oswald, 2004, Clark & Oswald, 1996, Ferrer-i-Carbonell, 2005, Ferrer-i-Carbonell & Frijters, 2004, Frijters et al., 2004, Luttmer, 2005, Senik, 2004) and sustained employment (Clark & Oswald, 1994, Di Tella et al., 2001, Winkelmann & Winkelmann, 1998). An important concern in subjective well-being studies is how to deal with heterogeneity between individuals that is largely considered to be unobservable. There is some uncertainty as to what individual heterogeneity consists of but the term fundamentally represents fixed factors unique to each individual that drive the individual to earn a higher income or remain in employment yet also enable them to have higher levels of well-being over the course of their life. An example would be the individual's personality. Any researcher interested in the independent effects of increasing an individual's income must somehow control for these correlated but largely fixed and unobservable heterogeneous factors.

A typical approach to overcome issues of heterogeneity is to exploit panel data and perform either a random or fixed effect estimation. Since individual heterogeneity is generally viewed as fixed across time, the observation of individuals at several time points allows researchers to statistically control for the heterogeneous factors without having to directly observe or quantify them. Inevitably, when important explanatory variables are unavailable estimation by ordinary least squares (OLS) will result in biased estimates. Ferrer-i-Carbonell & Frijters (2004) document this bias from not appropriately controlling for individual heterogeneity and observe that the positive coefficient on income reduces by about 2/3 when moving from a pooled OLS to a fixed effect estimator. The bias is large and illustrates the necessity of controlling for important yet largely unobservable and unknown characteristics. Ferrer-i-Carbonell & Frijters (2004) further suggest that individual heterogeneity, which appears to be an important source of information as to why some individuals have higher

levels of well-being than others, needs to be better understood. This paper aims to elicit an understanding of individual heterogeneity.

Aside from controlling for its presence there has been little work directed at understanding what is truly included within the set of fixed heterogeneous factors. Within the well-being literature there is no real consensus on what individual heterogeneity is and the term is often used to reflect aspects of an individual's character that researchers have difficulty measuring. Some researchers have suggested that in part there is some bias in individual responses to subjective well-being questions due to processes such as anchoring or whether the individual has an optimistic or pessimistic view of life (Clark et al, 2005; Ferrer-i-Carbonell, 2005; Winkelmann & Winkelmann, 1998). It has also been argued that an individual's health (Winkelmann & Winkelmann, 1998), their capacity to deal with adversities (Ferrer-i-Carbonell, 2005) and their ability or family back ground (Di Tella & MacCulloch, 2006) are important components of unobservable heterogeneity.

Perhaps the most widely cited and most important component of individual heterogeneity is an individual's personality. Many subjective well-being researchers have in fact made the explicit assumption that the unobservable individual heterogeneity is mainly personality traits (Booth & van Ours, 2008, Ferrer-i-Carbonell & Frijters, 2004, Frijters et al., 2004, Senik, 2004, Vendrik & Woltjer, 2007). It is unclear what is truly contained within the all encompassing term individual heterogeneity yet the assumption that personality is the main component has received little scrutiny. Perhaps unfamiliar to many economists is the fact that personality can be reliably measured and that personality research has a long history in psychology (Winter & Barenbaum, 1999). The use of such personality measures would enable researchers to improve their understanding of individual heterogeneity and determine how important personality is compared to other factors, such as the individual's health or background, as a component of individual heterogeneity. Further, an improvement to our understanding of individual heterogeneity may be of benefit in the estimation of subjective well-being equations.

The estimation strategy that seems to be favoured in subjective well-being studies to deal with individual heterogeneity is a fixed effect (FE) estimation. Based on the assumption that the unobservable heterogeneity is correlated with explanatory variables the FE model focuses solely on explaining the within-person variation. As a result the individual heterogeneity, which is considered fixed across time, contains no within-person explanatory power. In fact, all of the between-person information is not essential for estimation and is grouped together into what is referred to as the individual fixed effects and mostly ignored. However, this focus on the within-person variation is inefficient and the resultant loss of between-person information has been described by Beck and Katz (2001) as like “throwing the baby out with the bath water”. It is therefore not possible to obtain reliable estimates on characteristics that have zero or low within-person variation using an FE estimation (Plumper & Troeger, 2007).

An alternative that circumvents the problem of obtaining reliable estimates on characteristics with low within-person variation is a random effects (RE) estimation. Here, a different assumption is made about individual heterogeneity, in that the heterogeneity is uncorrelated with explanatory variables of interest. This is a strong assumption and, although the RE model is more efficient than the FE model, since it uses both within and between-person information, is likely to produce biased estimates if this core assumption is violated. Mundlak (1978) proposed a solution to this problem by allowing for correlation between the unobservable heterogeneity and some of the observable characteristics. The random effects model with a Mundlak (1978) transformation (REMT) overcomes the inefficiency problems associated with the FE model but still maintains many of the key FE assumptions. An example of the REMT’s use in subjective well-being research can be seen in Ferrer-i- Carbonell (2005).

It is clear that the correlation between the unobservable heterogeneity and observable characteristics is a key factor determining a subjective well-being researcher’s choice of estimation strategy. However, if personality measures can add to our understanding of

individual heterogeneity then this correlation problem might not be so critical. The use of personality measures in estimation may allow panel data estimation strategies to be applied to subjective well-being research that are less restrictive than the FE or REMT. It has already been demonstrated that personality measures are a useful substitute to control for individual heterogeneity when panel data is unavailable (Anand et al., 2009). This paper also exploits personality measures, but instead within a panel data framework, and suggests the fixed effect vector decomposition (FEVD) model, developed by Plumper and Troeger (2007), as an alternative to the FE and REMT models. The FEVD model is based on the standard FE model but also recognises that some of the fixed individual heterogeneity is in fact observable. The FEVD has three stages and, whilst giving the technique efficiency advantages over the FE model by maximising the use of between-person variation, the technique still controls for the truly unobservable component of individual heterogeneity. In the first of the three stages individual fixed effect residuals are estimated using the FE model. Individual fixed effect residuals, which absorb all between-person information, are then decomposed into an observable and unobservable component. Here, available characteristics that have little or zero within-person variation can be used to predict the fixed effect residual. The final stage then uses the error term from stage two, representing the true unobservable heterogeneity, and all the observable characteristics, as an explanatory variable in a pooled OLS estimation.

The FEVD allows the estimation of characteristics that have zero or only small within-person variations. In studies of subjective well-being the application of FEVD, like the REMT, seems immediately beneficial for obtaining estimates on various demographic characteristics that mostly vary between individuals and not within, for example, age, gender and geographical location, that are unobtainable using the FE model. There are, however, also other important characteristics that are known to change very little over the individual's life and the use of either the FE or REMT models to estimate the effect of such characteristics can often result in insignificant coefficients. An example would be the individual's level of education. In some studies, such as Luttmer (2005) and Ferrer-i-Carbonell (2005), the

coefficient on education is often found to be indeterminable and invites the conclusion that education does not improve the individual's life satisfaction.

In the face of estimating variables with low or zero within-person variation it might be sensible to opt for either an RE or REMT estimation. However, in addition to proposing the FEVD as an alternative to the FE model for estimating characteristics with low within-person variation, Plumper and Troeger (2007) also conduct Monte Carlo simulations to demonstrate the conditions under which the FEVD provides superior estimates to the FE model. Such specific conditions are not available for models like the RE or REMT. They show that, when a characteristic of interest has a low enough within-person variation, estimation can be preferable using FEVD. They also illustrate that what constitutes a low enough within-person variation is heavily dependent upon how much that particular variable correlates with the unobservable fixed heterogeneity. However, since the fixed heterogeneity is, by definition, unobservable, no correlation can be obtained. To aid researchers with an application of their technique Plumper and Troeger (2007) suggest that by being able to explain more of the fixed effect residual in stage two reduces the likelihood of a high correlation. In the case of subjective well-being research this would include measures of an individual's personality.

Many characteristics that are typically estimated in subjective well-being equations have fairly low within-person variations. This paper uses subjective well-being data from the German Socio-Economic Panel, which included a number of validated and reliable personality questions in a recent wave. First, it is shown that many of the conventional variables used in subjective well-being research lack within-person variation. Second, it is shown using valid and reliable personality measures that personality is the main component of individual heterogeneity and that using personality measures to increase our understanding of individual heterogeneity substantially reduces the correlation between the remaining unobservable heterogeneity and important explanatory variables. Collectively these two results make estimation possible using a FEVD. This paper argues that more reliable estimates are obtained on characteristics that have only moderately low within-person

variations and this includes the individual's marital status, disabilities and even income. The next section discusses the measurement of personality and its use within economics. Section 3 gives greater technical detail regarding the FEVD and shows how the technique can incorporate the use of personality measures and be applied to subjective well-being research. Section 4 describes the German Socio-Economic Panel data used in the analysis. The results are then given in section 5 and section 6 concludes.

2. The use of personality measures in economics

Personality research has a long history in psychology beginning with researchers such as Allport (1937) and Cattell (1946) and has since developed into a systematic analysis of individual differences. Models of personality are generally constructed around natural language and the words used to describe people (for a full development of this lexical approach see John and Srivastava (1999)). The personality literature is dominated by the Big Five Personality Inventory, which suggests at the broadest level of abstraction there are five dimensions of personality. Using factor analysis it has been shown that the large numbers of words used to describe individuals load onto five main themes: openness to experience, conscientiousness, extroversion, agreeableness and neuroticism. Although the model is atheoretical the personality measures such as the Big Five have been shown to measure what they are supposed to (Benet-Martinez & John, 1998, Borkenau & Ostendorf, 1990) making them distinct from indicators of subjective well-being.

Models like the Big Five have been used to address some areas that are of economic interest. For example, Mueller and Plug (2006), Nyhus and Pons (2005) and Groves (2005) have all looked at the effect of personality on earnings. Mueller and Plug (2006) show that some personality traits, such as openness and conscientiousness, are rewarded in the market place, whereas other traits, such as agreeableness and neuroticism, are penalised. Nyhus and Pons (2005) draw similar conclusions but also find that the degree of control an individual has is important for earnings. They further find that the financial return to personality varies

across educational groups. Groves (2005) investigates the importance of psychological traits, such as autonomy, social withdrawal and aggression in female earnings. Studies such as these may help explain why even after controlling for many factors, which includes the improved cognitive abilities that come through schooling, there are still large earning gaps. Bowles et al. (2001a) have suggested that both school and family pass on many important behavioural traits that enhance the individual's earning success. However, the use of personality traits in the determination of wages is very much in its infancy (Bowles et al., 2001b).

More recently researchers have looked at personality's relation to an individual's propensity to share knowledge with work colleagues (Matzler et al., 2008). In fact, a recent discussion paper by Borghans et al. (2008) evaluates the integration of personality into economic research more generally. From a theoretical perspective they discuss ample evidence that suggests that personality should be given greater consideration when discussing economic parameters and constraints. Borghans et al. (2008) further suggest that there could be considerable benefit to understanding how economics incentives might influence individuals with different personality traits.

In contrast to the approach by economists, a psychologist's discussion of subjective well-being will often centre on personality. Personality enables a categorisation of people and their behaviours and is therefore one of the strongest and most consistent factors predicting the individual's well-being (Diener & Lucas, 1999). Economists show greater interest in demographic factors, like age, education, income and marital status, which although explaining relatively less of the individual's well-being (Argyle, 1999) can be influenced by policy. In subjective well-being studies economists only require some way of controlling for personality. By assuming personality is fixed, an assumption that is supported by work in psychology (Costa & McCrae, 1980, Costa & McCrae, 1988), economists are able to control for personality using the models like the FE and REMT to obtain unbiased causal effects on various demographic factors. The use of personality measures, however, to better understand the fixed effects will help merge subjective well-being research of both psychologists and

economists and may help uncover the true causes of high subjective well-being (Gutierrez et al., 2005).

3. Methodology

At a given time point (t) the individual's subjective well-being (SWB) is generated as follows:

$$(1) SWB_{it} = \alpha + \sum_{d=1}^D \lambda_d \delta_t + \sum_{k=1}^K \beta_k x_{kit} + \sum_{m=1}^M \gamma_m z_{mi} + \mu_i + \varepsilon_{it}$$

There are some characteristics (δ) which vary across time periods but not individuals, for example, the economic or social conditions that affect everyone equally at a given time point. Other characteristics (x) vary across individuals and time periods and may include aspects such as the individual's income and employment status. There are other observable characteristics (z) that vary from individual to individual but not across time, such as gender. Finally, there are other important factors (μ), considered unobservable, that do not change across time periods and are referred to as individual heterogeneity.

The key assumption of the FE model is that the unobservable individual heterogeneity is correlated with the observable characteristics. The FE model eliminates the need to worry about any individual heterogeneity (μ) by focusing solely on how much individuals vary from their time-means. This is known as the within-person variation and shown in equation 2.

$$(2) SWB_{it} - \overline{SWB}_i = (\alpha - \bar{\alpha}) + (\delta_t - \bar{\delta}) + \beta_k \sum_{k=1}^K (x_{kit} - \bar{x}_{ki}) + \gamma_m \sum_{m=1}^M (z_{mi} - \bar{z}_{mi}) + (\mu_i - \bar{\mu}_i) + (\varepsilon_{it} - \bar{\varepsilon}_i)$$

where

$$\overline{SWB}_i = \frac{1}{T} \sum_{t=1}^T SWB_{it}, \quad \bar{x}_i = \frac{1}{T} \sum_{t=1}^T x_{it}, \quad \bar{z}_i = \frac{1}{T} \sum_{t=1}^T z_{it}, \quad \bar{\mu}_i = \frac{1}{T} \sum_{t=1}^T \mu_i = \mu, \quad \bar{\varepsilon}_i = \frac{1}{T} \sum_{t=1}^T \varepsilon_{it}$$

By eliminating the individual heterogeneity (μ) unbiased estimates of the x characteristics can be obtained. However, no estimates of the observable z characteristics are obtainable from such a regression. Inadvertently they will be eliminated alongside the individual heterogeneity and important information will be lost. In contrast the RE model makes better use of the information available. The RE model assumes the unobservable heterogeneity is uncorrelated with observable characteristics. This means that the time-invariant individual heterogeneity, μ , is subsumed by the error term, ε . Reasonable estimates can be obtained as long as the error structure is recognised using an estimator such as generalised least squares (GLS). However, the reliability of these estimates depends on the strength of the assumption that observable characteristics are not correlated with the individual heterogeneity. In studies of subjective well-being it would be difficult to argue convincingly that this was the case. Some researchers have therefore proposed the use of the REMT. The REMT circumvents this problem by allowing the error structure to take account of the correlation, in a similar fashion to the FE model, by including the time-mean values of the observable characteristics that are thought to be correlated with the unobservable heterogeneity. An issue with such a technique, however, is that there is no clear cut way of choosing which variables are correlated with the unobservable heterogeneity. There are some individual characteristics that it would be difficult to argue weren't correlated with the unobservable heterogeneity but ultimately the decision is down to a researcher's discretion.

The FEVD is an alternative estimation strategy proposed by Plumper and Troeger (2007) which similarly attempts to overcome the loss of information that occurs using the FE model. The advantage of this technique over the REMT is that Plumper and Troeger (2007) provide clear conditions under which the FEVD estimation is superior to an FE model. Their technique performs the FE model in its first stage in order to obtain an estimate of the fixed effect residual ($\hat{\mu}_i$). However, they note that the fixed effect residual ($\hat{\mu}_i$) obtained using the FE model is not the same as the true unobservable heterogeneity (μ) outlined in equation 1. The fixed effect residual also contains the eliminated information of characteristics contained

in z as well as the mean effects of the characteristics contained in x . An estimate of the fixed effect residual ($\hat{\mu}_i$) using the FE model effectively includes all observable and unobservable between-person information.

$$(3) \quad \hat{\mu}_i = \overline{SWB}_i - \sum_{k=1}^K \beta_k \bar{x}_{kit} - \bar{\varepsilon}_i$$

Thus, in the second stage of the FEVD technique Plumper and Troeger (2007) suggest decomposing the fixed effect residual into a part that is observable and a part that is not. It is this stage in which greater understanding of the fixed effect can be obtained by using any available between-person information, which would include personality variables. Here, it would be possible to determine what the main contributing factors to individual heterogeneity were. The decomposition takes place using z characteristics to predict the fixed effect residual obtained from stage one.

$$(4) \quad \hat{\mu}_i = \sum_{m=1}^M \gamma_m z_{mi} + \eta_i$$

This leaves the true unobservable component of μ , captured in the error term from equation 4 and denoted here as η . Next, η is used in a third stage pooled OLS regression as an explanatory variable.

$$(5) \quad SWB_{it} = \alpha + \sum_{d=1}^D \lambda_d \delta_t + \sum_{k=1}^K \beta_k x_{kit} + \sum_{m=1}^M \gamma_m z_{mi} + \theta \eta_i + \varepsilon_{it}$$

Although z variables may have been correlated with μ , they are not correlated with η . Therefore, by including the error term (η) from stage two the FEVD allows researchers to obtain reliable estimates on z characteristics.

Plumper and Troeger (2007) discuss the conditions under which characteristics can be classified as z characteristics and favourably estimated using the FEVD. There are some variables that belong strictly in either the set of x or z variables; they change all the time or not at all. However, there are other variables in which a strict categorisation is not possible. For instance, there are some characteristics that simply do not change very much. Obvious examples might include education or marital status. For a huge proportion of a population these types of characteristics may never change, whilst in others they may be changing often. Hence, for part of a given sample, a particular characteristic will behave like a z characteristic, whilst for others, the characteristic will behave like an x characteristic. In the FE model the information from the part of the sample that does not change cannot be used. Similarly, if the characteristic is treated as an x variable in the FEVD then the information is also ignored. However, it is possible to treat some characteristics as z variables if they have a low within-person variation. If the within-person variation is sufficiently small enough then the trade-off between bias and efficiency favours the efficient estimator.

In their Monte Carlo simulations Plumper and Troeger (2007) show using the root mean squared error under what conditions the gain in information at the sacrifice to bias favours estimation by FEVD. Plumper and Troeger (2007) pin-point the ratio of a particular variables between-to-within person variation as a way of distinguishing whether that variable can be better estimated using FEVD. This ratio, however, depends on how well the particular variable in question is correlated with the unobservable heterogeneity (η). For example, when the correlation is 0.5 the between-to-within person ratio must exceed 2.8 for the FEVD to be the superior estimator. When the correlation drops to 0.3 the between-to-within person ratio only needs to be 1.7. These between-to-within person ratios are fairly low and may include many characteristics that economists have so far only estimated using the FE or REMT models.

The correlation between unobservable heterogeneity and any potentially low within-person characteristic is clearly unobservable. However, as Plumper and Troeger (2007)

suggest, including additional z variables in stage two of the FEVD and obtaining a better understanding of the fixed effect is likely to decrease the size of the unobservable component (η). If the unobservable component (η) of the individual heterogeneity (μ) is reduced then so too will the likely correlation between any potentially low within-person variables and the true unobservable component. Using the truly unobservable component of individual well-being, the error, η , from stage two of FEVD, it is possible to determine empirically the approximate size of such a correlation. The use of personality measures as additional z variables in stage two of the FEVD increases our understanding of the fixed effect and therefore reduces this correlation allowing many variables to be favourably estimated using the FEVD model. This reduction in correlation through the use of personality variables may additionally make an RE estimation preferable to REMT.

4. Data

The data used to aid the understanding of the fixed factors that contribute to well-being comes from the German Socio-Economic Panel (GSOEP) survey, a representative longitudinal sample of German households. The survey asks a number of questions about each individual's life. The list of questions includes a single item life satisfaction question:

How satisfied are you with your life, all things considered?

Possible responses range from 0, indicating complete dissatisfaction, to 10, indicating complete satisfaction. The response to the life satisfaction question is assumed to be cardinal. Other questions in the GSOEP uncover various objective circumstances of an individual's life. The variables used here include: demographics, education levels, household income, household size, marital and employment status, the individual's self-rated health, whether there are children in the household and disabilities.¹

¹ A description of all variables and how they were constructed is contained in the Notes to Tables

In 2005, a series of questions designed to uncover aspects of an individual's personality were included in the GSOEP. Self-reported personality measures generally have high levels of reliability and validity. Previous research has shown, for example, that self-reported personality measures are highly stable over time (McCrae & Costa, 1990) and relate to peer ratings (McCrae & Costa, 1987). Self report measures also predict both objective behavior (Epstein, 1979) and occupational success (Hogan, 2005), have biological correlates (Ryff et al., 2006) and relate to changes in objective biological functioning (O'Cleirigh et al., 2007). Of the 31 personality questions used here, 15 are a considerably shortened version of the standard Big Five personality questionnaire, and a further 16 relate to an individual's reciprocity, locus of control and pessimism. Before being included in GSOEP the short item Big Five scale underwent extensive pre-testing and has been shown to satisfactorily replicate the standard Big five questionnaire (Gerlitz & Schupp, 2005). This scale has been used in studies such as Winkelmann and Winkelmann (2008). There are 6 questions on an individual's reciprocity. The questions on reciprocity can be separated into positive and negative reciprocity and examples of their use can be seen in Dohmen et al. (2008) and Fliessbach et al. (2007). There are 9 questions that indicate an individual's locus of control. This construct can be traced back to the work of Rotter (1966) and the same set of questions has been used in Fliessbach et al. (2007). There is one question that directly asks the degree to which an individual is pessimistic about the future.

Factor analysis confirmed that the 31 personality questions grouped meaningfully into the personality traits outlined above.² The sole purpose of using the personality measures in this paper, however, is to maximize the explanation of individual heterogeneity. In the main analysis all 31 measures are therefore included as separate predictors, rather than as the grouped personality constructs. The personality questions asked in 2005 are assumed to be reliable proxies for personality across all years of analysis. A key assumption of panel data models is that individual heterogeneity is fixed across time so this seems like a reasonable

² The Appendix contains a full list of the personality questions and a description of how the measures combine into meaningful personality constructs

assumption. However, although innate personality may be relatively stable across time an individual will not necessarily give the same response to a given question each year. It is therefore likely that personality measures will be prone to some measurement error. This represents a limitation to this study but more accurate personality measures are likely to only add to the explanatory power of individual heterogeneity. It will be interesting to determine how much these available measures contribute to the explanation of heterogeneity when compared to other likely sources of heterogeneity, including an individual's health and background.

The panel constructed is unbalanced. All individuals are observed in 2005 and at least one other time point. The period under analysis is 6 years. This time-frame is considerably shorter than the available data in GSOEP and means that the data set is likely to have a lower within-person variation than a longer panel. Estimation by FEVD will therefore be falsely superior according to the conditions set out by Plumper and Troeger (2007). However, a short data set is needed to ensure that personality measures from 2005 are adequate proxies across the entire period under analysis and can be adequately used to attempt an understanding of the fixed effect. To counter the potential low-within person variation problem the descriptive statistics across a 12 year panel from GSOEP are also shown alongside the 6 year panel in Table 1. Across both panels most characteristics have between variations that exceed the within variation. Many important characteristics are also observed to have within-to-between variations exceeding 2 and this suggests that a great deal of observable information will be discarded when using the FE model. Depending on the correlation between the unobservable heterogeneity and any characteristics of interests, as will be empirically approximated later, estimation using FEVD might be the preferable estimation strategy.

Concentrating on the 6 year panel used in the main analysis there are 93016 individual-year observations coming from 17210 unique individuals. For ease of interpretation in the later analyses all the variables with intrinsically non-meaningful scales,

life satisfaction, self-rated health and all of the personality measures, are standardised with a mean of zero and standard deviation of one.

5. Results

Analysis begins by showing the importance of controlling for individual heterogeneity. In Table 2 standardised life satisfaction equations are estimated using pooled OLS, FE and REMT models. All models offer some interesting insights into what makes an individual satisfied with their life. However, there are important differences between the models. Both the FE and REMT appropriately control for individual heterogeneity and the FE model has often been interpreted as producing causal effects. The pooled OLS represents a mere association. For example, concentrating on the effect size on the log of household income, the FE model has a coefficient of 0.17, whereas the pooled OLS estimator has a coefficient that is nearly twice as large at 0.33. The coefficient on the FE model suggests that if the individual's household income were doubled then their life satisfaction could be expected to increase by 0.17 standard deviations. The pooled OLS model coefficient would suggest that, *ceteris paribus*, if individual x were observed to have a household income twice the size of an individual y, then individual x would be on average more satisfied with their life by 0.33 standard deviations. The prediction from the pooled model is after having controlled for all other observable characteristics and presents the main drawback of the pooled OLS model – there are important factors correlated with both income and life satisfaction that are unobservable. Not controlling for these unobservable factors results in biased coefficients. At the sacrifice of efficiency it is sensible to opt for an unbiased estimator such as the FE model.

There are, however, other important differences across the pooled OLS and FE models. Some variables do not have enough within-person variation to enable reliable estimation. In Table 2 variables with zero within-person variation, such as gender, cannot be included in the fixed effects model. Controlling for age can also be problematic. Age changes within all individuals in the same way and when included in the FE model is only

interpretable as a linear time trend. As a result when time dummies are included there is little reason to include age as an explanatory variable in the FE model. Contrastingly, both age and gender can be included in the pooled OLS and the REMT models. These models suggest that women are more satisfied with life and that there is a u-shape relationship between age and life satisfaction (life satisfaction minimises at around 39 in the pooled OLS model). More importantly a closer observation across all the models highlights the difficulty of obtaining reliable estimates on variables that could be termed as slow changing. Table 1 showed that characteristics like education and marital status had very low within-person variations. In both the REMT and FE models in Table 2 the coefficient on education is indeterminable. The coefficients on marital status, on the other hand, vary considerably and in conflicting directions from the estimates given in the pooled OLS model. This leaves some concern over the reliability of the coefficients using both FE and REMT estimations on variables with low within-person variations.

The FE model in Table 2 is used to estimate a fixed effect residual for each individual. Observable characteristics can then be used to decompose the fixed effect residual in Table 3. Column 1 begins by including only observable demographic characteristics; these variables explain 7% of the variation. Column 2 extends the model by further including what could be considered as very slow moving variables; marital status, education and whether the individual is retired. Adding these variables increases the explanation of the fixed effect residual to 10%. Column 3 indicates how much personality contributes to an explanation of the fixed effect residual. When grouped into their 9 meaningful constructs the personality measures collectively explain 18% of the fixed effect residual, with most of this explanation coming from an individual's level of pessimism, locus of control and neuroticism. The explanation of the fixed effect residual rises to 20% when personality is included as 31 separate scores. Maximising the explanation of the fixed effect residual is most important here so all 31 measures are used in the subsequent analysis. In column 4 these personality variables are appended to the explanatory variables considered to be slow moving. The overall

explanation of the fixed effect residual rises to 29%. Finally column 5 includes the between-person information about the household's income, whether there are children in the house and an individual's health and disabilities. The explanation rises to 33%.

Table 3 illustrates that personality provides the greatest explanation of individual heterogeneity when compared to other observable characteristics. Pessimism, locus of control, agreeableness and neuroticism are particularly important components of individual heterogeneity. The individual's health is also observed to be an important component. This suggests that individual heterogeneity is mostly, although not exclusively, personality. Since a large proportion of individual heterogeneity is in fact observable it may be sensible to consider alternative estimation strategies. As has been discussed, many of the favoured models are based on an assumed correlation between unobservable heterogeneity and the observable characteristics.

The fixed effect residual can also be used to empirically approximate the likely correlation between individual heterogeneity and other observable characteristics. The first column in Table 4 shows that there is only low to moderate correlation between observable characteristics and the fixed effect residual. All variables are below 0.2. The second column controls for demographic characteristics with the correlation rising across many characteristics. However, as personality and other characteristics with very low within-person variation are included in Columns 3 and 4 respectively the correlations substantial reduce. As Plumper and Troeger (2007) show when an observable characteristic is correlated with the unobservable component of individual heterogeneity by just 0.3 the between-to-within person ratio only needs to be 1.7 to make estimation by FEVD superior to the FE model. Focusing specifically on income, the correlation is just 0.09 and has a between-to-within ratio of 2.17, suggesting that estimation may be preferable using FEVD.

The estimations carried out in Table 3 represent the second stage of the FEVD technique. Table 5 therefore displays the third stage FEVD results by including the error terms from Table 3 as explanatory variables in a pooled OLS estimation. In column 1 the

error term from the second column in Table 3, where only demographic and slow moving variables were used to predict the fixed effect residual in the second stage, is used in estimation. Although efficiency of estimation has been increased it is important to note that the coefficients on variables not used in stage two are the same as those using the standard FE model seen in Table 2. On the other characteristics used in stage 2, however, some interesting changes occur. The estimation in the first column of Table 5 suggests that education has a positive effect on life satisfaction, contrasting the negative and insignificant coefficient seen in the FE and REMT models in Table 2.

Here, it is useful to comment on the coefficients on marital status in the third stage of the FEVD when compared with the FE model in Table 2. The change in the coefficients across the models is not consistent. Married has a coefficient that is nearly three times larger in the FEVD model. In contrast, the coefficient on widowhood is at least three times smaller using FEVD, and the variable divorced reverts from a positive to a negative coefficient. These changes across marital status are inconsistent and this highlights an important issue with the FE model. When comparing the FE and FEVD models it is important to consider the difference between a change in circumstance and maintaining a permanent situation. In marital status, for example, only the changes to the individual's marital status are important in the FE model. Thus, if there is no change to the individual's marital status then the individual will yield no useful information for the FE estimation. However, that same individual may have greater life satisfaction simply from sustaining a particular marital situation for a considerable period of time. For example, it is likely that getting married will increase the individual's life satisfaction but additionally there will also be well-being benefits from sustaining a healthy marriage for an extended period of time. Similarly, getting divorced may bring immediate life satisfaction benefits as seems to be supported in work by Gardner and Oswald (2006). However, if the individual were to remain in a divorced state for a sustained period then it would seem plausible that there could eventually be adverse consequences for their well-being. The coefficients on widowed invite a similar argument –

becoming a widow initially has a large negative impact on the individual's life. Over time, however, the impact reduces as the individual adapts to their loss, supporting conclusions in both Clark et al (2008) and Gardner and Oswald (2006). The issue associated with marital status generalises to other variables and highlights a practical concern with the FE model.

Column 2 of Table 5 shows a further estimation using FEVD using the error term from column 5 of Table 3. Since the use of personality measures substantially reduce the correlation between unobservable heterogeneity and important characteristics it enables many of the observable variables to be classified as slow moving enough to be preferably estimated using FEVD. The coefficients are different to the pooled OLS, FE and REMT models seen in Table 2 and, given the core model assumptions, possibly reflect more accurate estimates. Another way of using personality variables, given that the correlation with unobservable heterogeneity and observable characteristics is substantially reduced, could be by using a standard RE model. Columns 3 and 4 show the results from an RE estimation both without and with personality variables respectively. The coefficients are fairly similar to the FEVD coefficients with the personality variables attenuating many of the coefficients in comparison to the pooled OLS model. The FEVD, however, is the preferred model since it satisfies the specific conditions set out by Plumper and Troeger (2007).

The pooled OLS model, by not appropriately dealing with unobservable correlated factors, has a tendency to produce biased coefficients. The FE model, on the other hand, discards all between-person information and without a true understanding of individual heterogeneity underestimates the effect on life satisfaction of various individual characteristics. The FEVD by offering an alternative way to deal with individual heterogeneity combines elements of both techniques to enable efficient yet unbiased estimates. The results from column 2 in Table 5 suggest that doubling the individual's household income will actually increase their life satisfaction by 0.29 standard deviations. The discrepancy with the FE model arises due to the well-being benefits that come about from having a permanently high level of household income as well as increases to household

income. Additionally the FEVD provides more reliable coefficients on the effect of age, education, marital status, disabilities and having children.

6. Conclusion

This paper attempts to understand individual heterogeneity, which has been shown to substantially attenuate estimates of effects when moving from a pooled OLS to an FE model in subjective well-being studies. Here, personality measures are used to increase the understanding of individual heterogeneity and help confirm that personality is one of the main components of individual heterogeneity. Health and other demographic characteristics also provide some explanation of heterogeneity. A greater understanding of individual heterogeneity reduces the correlation between the remaining unobservable heterogeneity. Reducing this correlation is key, since it allows alternative techniques to be explored, and enables more reliable estimates on variables that have low within-person variations, such as income, education, marital status, disabilities and having children.

The use of a FEVD model with personality variables has a tendency to produce estimates that lie somewhat between estimates on the pooled OLS and FE estimations. For example, using the FEVD the individual's household income is estimated to be more than 1.6 times more beneficial for the individual's well-being than is suggested by the FE and REMT models, but around 0.85 that of a pooled OLS model. One potential explanation is that there are still other important unobservable components not controlled for using the FEVD. However, another reason could be that the FE and REMT models, by making strong assumptions about the correlation with the unobservable heterogeneity, are simply too restrictive. Specifically, the FE model leaves no room for uncovering improvements to the individual's subjective well-being that may simply arise, for example, from having a permanently high income or being in a permanently healthy relationship. Only focusing on changes detracts from the benefits to well-being that may accrue from sustaining a high level or state.

The observable between-person information is shown to predict 13% of the fixed effect residual. This decomposition of the fixed effect residual suggests that the fixed effect residual should not be completely termed as unobservable individual heterogeneity and simply disregarded. It contains valuable observable information. The fixed effect residual is potentially an untapped source in providing answers as to why some individuals have higher subjective well-being than others. The personality measures used here alone explain 20% of the fixed effect residual. Compared to the explanation given by other characteristics, such as health and other demographic characteristics, this contribution is large. However, there still remains a substantial unexplained component. There are three possible explanations. The unexplained component could be due individual heterogeneity that is still largely unknown, for example an individual's ability. Alternatively, the measures used, particularly the one-item life satisfaction scale, are imperfect and are likely to be measured with some error. Lastly, the FE model, focusing on only changes as discussed earlier, may underestimate the importance of permanent state effects. It is likely that there is some combination of the three but simply terming the individual heterogeneity simply as personality traits and discarding the information seems inappropriate.

An important consideration for the future is the availability of personality measures in the large data sets commonly used by economists. Currently many representative national surveys like the GSOEP do not include questions on an individual's personality. Such unavailability may prove problematic for the approach outlined in this paper. Economists are relatively unfamiliar with the idea that personality can be measured and this has no doubt influenced the demand for inclusion of such measures in their data sets – this needs to change. Personality has already been shown to be an important determinant of wages (Bowles et al., 2001a) but further work is needed around this area. More generally Borghans et al. (2008) have convincingly argued that economic research has much to gain from using reliable and valid personality measures that are used extensively by psychologists. Personality appears to be one of the biggest and most consistent predictors of well-being (Diener & Lucas, 1999)

and as shown here is important for understanding both fixed and variable components of well-being. This paper adds to the support for the increased use of personality measures in subjective well-being research (Anand et al., 2009, Boyce & Wood, 2009). A wider inclusion of personality measures in data sets like GSOEP would be of enormous benefit to both personality and economic research.

This paper has gone some way in understanding the important fixed effect and shown the importance of exploiting between person information. The use of personality in this context is novel and may allow researchers to relax the statistical technique used to estimate subjective well-being equations. The use of personality measures combined with the FEVD technique may therefore provide an important methodological advance for subjective well-being research.

Appendix

Personality variables in GSOEP

In the questionnaire section entitled “What kind of personality do you have?” individuals are asked 30 questions. 15 of these relate to the “Big five” personality inventory, whilst a further 15 cover aspects of the individual’s reciprocity and control in life. A further question, on pessimism, comes from the “Attitudes and opinions” section. Using factor analysis the measures were found to load onto 9 different personality constructs: openness to experience, conscientiousness, extroversion, agreeableness, neuroticism, positive and negative reciprocity, locus of control and pessimism.

Big five personality inventory

Individuals are asked whether they see themselves as someone who...

1. *...does a thorough job*
2. *...is communicative, talkative*
3. *...is sometimes somewhat rude to others*
4. *...is original, comes up with new ideas*
5. *...worries a lot*
6. *...has a forgiving nature*
7. *...tends to be lazy*
8. *...is outgoing, sociable*
9. *...values artistic experiences*
10. *...gets nervous easily*
11. *...does things effectively and efficiently*
12. *...is reserved*
13. *...is considerate and kind to others*
14. *...has an active imagination*
15. *...is relaxed, handles stress well*

Individuals are asked whether the statement applies to them on a 1 to 7 scale, with 1 meaning the statement does not apply to them at all and 7 that it applies perfectly. These 15 variables load onto five personality dimensions: Openness to experience, conscientiousness, extroversion, agreeableness and neuroticism. Questions 4, 9 and 14 relate to an individual's openness to experience; questions 1, 7 & 11 relate to conscientiousness, questions 2, 8 & 12 relate to extroversion; questions 3, 6 & 13 relate to agreeableness; questions 5, 10 & 15 relate to neuroticism. These groups of questions can be reverse coded (as appropriate) and combined to give an underlying score of the personality dimension.

Positive and negative reciprocity

Individuals are asked to what extent the following apply to them

16. *If someone does me a favour, I am prepared to return it*
17. *If I suffer a serious wrong, I will take revenge as soon as possible, no matter what the cost*
18. *If somebody puts me in a difficult position, I will do the same to him/her*
19. *I go out of my way to help somebody who has been kind to me before*
20. *If somebody offends me, I will offend him/her back*
21. *I am ready to undergo personal costs to help somebody who helped me before*

Individuals are asked whether the statement applies to them on a 1 to 7 scale, with 1 meaning the statement does not apply to them at all and 7 that it applies perfectly. Questions 16, 19 & 21 load onto a construct termed positive reciprocity and questions 17, 18 & 20 load onto the individual's negative reciprocity.

Locus of control

Individuals are asked their attitudes towards their life and future.

22. *How my life goes depends on me*
23. *What a person achieves in life is above all a question of fate or luck*
24. *If a person is socially or politically active, he/she can have an effect on social conditions*
25. *I frequently have the experience that other people have a controlling influence over my life*
26. *One has to work hard in order to succeed*
27. *If I run up against difficulties in life, I often doubt my own abilities*
28. *The opportunities that I have in life are determined by the social conditions*
29. *Inborn abilities are more important than any efforts one can make*
30. *I have little control over the things that happen in my life*

Individuals are asked whether they agree with the statements on a 1 to 7 scale, with 1 representing complete disagreement and 7 that they completely agree. These measures reflect an individual's locus of control and factor analysis show that questions 22, 25, 27 & 30 can be grouped together to give an indication of this trait.

Pessimism

This is a one item scale. Individuals are asked

31. *When you think about the future, are you...optimistic, more optimistic than pessimistic, more pessimistic than optimistic, pessimistic?*

This variable is treated as cardinal.

References

- Allport, G.W. (1937) *Personality: A psychological interpretation* (New York, Henry Holt).
- Anand, P., Hunter, G., Carter, I., Dowding, K., Guala, F. & Van Hees, M. (2009) The development of capability indicators, *Journal of Human Development and Capabilities*, 10(1), pp. 125-152.
- Argyle, M. (1999) Causes and correlates of happiness, in: D. Kahneman, E. Diener & N. Schwarz (Eds) *Well-being: The foundations of hedonic psychology* (New York, Russell Sage Foundation).
- Beck, N. & Katz, J.N. (2001) Throwing out the baby with the bath water: A comment on Green, Kim, and Yoon, *International Organization*, 55(2), pp. 487- 495.
- Benet-Martinez, V. & John, O.P. (1998) Los Cinco Grandes across cultures and ethnic groups: Multitrait multimethod analyses of the Big Five in Spanish and English, *Journal of Personality and Social Psychology*, 75(3), pp. 729-750.
- Blanchflower, D.G. & Oswald, A.J. (2004) Well-being over time in Britain and the USA, *Journal of Public Economics*, 88(7-8), pp. 1359-1386.
- Booth, A.L. & van Ours, J.C. (2008) Job satisfaction and family happiness: The part-time work puzzle, *Economic Journal*, 118(526), pp. F77-F99.
- Borghans, L., Duckworth, A.L., Heckman, J.J. & ter Weel, B. (2008) The economics and psychology of personality traits, *Journal of Human Resources*, 43(4), pp. 972-1059.
- Borkenau, P. & Ostendorf, F. (1990) Comparing exploratory and confirmatory factor-analysis - a study on the 5-factor model of personality, *Personality and Individual Differences*, 11(5), pp. 515-524.
- Bowles, S., Gintis, H. & Osborne, M.O. (2001a) The determinants of earnings: A behavioral approach, *Journal of Economic Literature*, 39(4), pp. 1137-1176.
- Bowles, S., Gintis, H. & Osborne, M.O. (2001b) Incentive-enhancing preferences: Personality, behavior, and earnings, *American Economic Review*, 91(2), pp. 155-158.
- Boyce, C.J. & Wood, A.M. (2009) Which personality types have the highest marginal utilities of income? University of Warwick.
- Cattell, C.R. (1946) *Description and measurement of personality* (Yonkers-on-Hudson, NY: World Book).
- Clark, A.E., Diener, E., Georgellis, Y. & Lucas, R.E. (2008) Lags and leads in life satisfaction: A test of the baseline hypothesis, *Economic Journal*, 118(529), pp. F222-F243.
- Clark, A.E., Etilé, F., Postel-Vinay, F., Senik, C. & Van der Straeten, K. (2005) Heterogeneity in reported well-being: Evidence from twelve European countries, *Economic Journal*, 115(502), pp. C118-C132.
- Clark, A.E. & Oswald, A.J. (1994) Unhappiness and unemployment, *Economic Journal*, 104(424), pp. 648-659.

- Clark, A.E. & Oswald, A.J. (1996) Satisfaction and comparison income, *Journal of Public Economics*, 61(3), pp. 359-381.
- Costa, P.T. & McCrae, R.R. (1980) Influence of extraversion and neuroticism on subjective well-being - happy and unhappy people, *Journal of Personality and Social Psychology*, 38(4), pp. 668-678.
- Costa, P.T. & McCrae, R.R. (1988) Personality in adulthood - a 6-year longitudinal-study of self-reports and spouse ratings on the Neo Personality-Inventory, *Journal of Personality and Social Psychology*, 54(5), pp. 853-863.
- Di Tella, R. & MacCulloch, R. (2006) Some uses of happiness data in economics, *Journal of Economic Perspectives*, 20(1), pp. 25-46.
- Di Tella, R., MacCulloch, R.J. & Oswald, A.J. (2001) Preferences over inflation and unemployment: Evidence from surveys of happiness, *American Economic Review*, 91(1), pp. 335-341.
- Diener, E. & Lucas, R.E. (1999) Personality and subjective well-being, in: D. Kahneman, E. Diener & N. Schwarz (Eds) *Well-being: The foundations of hedonic psychology* (New York, Russell Sage Foundation).
- Dohmen, T., Falk, A., Huffman, D. & Sunde, U. (2008) Representative trust and reciprocity: Prevalence and determinants, *Economic Inquiry*, 46(1), pp. 84-90.
- Epstein, S. (1979) Stability of behavior .1. Predicting most of the people much of the time, *Journal of Personality and Social Psychology*, 37(7), pp. 1097-1126.
- Ferrer-i-Carbonell, A. (2005) Income and well-being: an empirical analysis of the comparison income effect, *Journal of Public Economics*, 89(5-6), pp. 997-1019.
- Ferrer-i-Carbonell, A. & Frijters, P. (2004) How important is methodology for the estimates of the determinants of happiness?, *Economic Journal*, 114(497), pp. 641-659.
- Fliessbach, K., Weber, B., Trautner, P., Dohmen, T., Sunde, U., Elger, C.E. & Falk, A. (2007) Social comparison affects reward-related brain activity in the human ventral striatum, *Science*, 318(5854), pp. 1305-1308.
- Frijters, P., Haisken-DeNew, J.P. & Shields, M.A. (2004) Money does matter! Evidence from increasing real income and life satisfaction in East Germany following reunification, *American Economic Review*, 94(3), pp. 730-740.
- Gardner, J. & Oswald, A.J. (2006) Do divorcing couples become happier by breaking up?, *Journal of the Royal Statistical Society Series a-Statistics in Society*, 169, pp. 319-336.
- Gerlitz, J. & Schupp, J. (2005) Zur erhebung der Big-Fivebasierten persönlichkeitsmerkmale im SOEP. DIW Research Notes 2004–5.(Berlin, DIW Berlin).
- Groves, M.O. (2005) How important is your personality? Labor market returns to personality for women in the US and UK, *Journal of Economic Psychology*, 26(6), pp. 827-841.
- Gutierrez, J.L.G., Jimenez, B.M., Hernandez, E.G. & Puente, C.P. (2005) Personality and subjective well-being: Big Five correlates and demographic variables, *Personality and Individual Differences*, 38(7), pp. 1561-1569.

- Hogan, R. (2005) In defense of personality measurement: New wine for old whiners, *Human Performance*, 18(4), pp. 331-341.
- John, O.P. & Srivastava, S. (1999) The Big Five trait taxonomy: History, measurement, and theoretical perspectives, in: L.A. Pervin & O.P. John (Eds) *Handbook of Personality: Theory and Research* (New York: London, Guildford Press).
- Luttmer, E.F.P. (2005) Neighbors as negatives: Relative earnings and well-being, *Quarterly Journal of Economics*, 120(3), pp. 963-1002.
- Matzler, K., Renzl, B., Muller, J., Herting, S. & Mooradian, T.A. (2008) Personality traits and knowledge sharing, *Journal of Economic Psychology*, 29(3), pp. 301-313.
- McCrae, R.R. & Costa, P.T. (1987) Validation of the 5-factor model of personality across instruments and observers, *Journal of Personality and Social Psychology*, 52(1), pp. 81-90.
- McCrae, R.R. & Costa, P.T. (1990) *Personality in adulthood* (New York, The Guildford Press).
- Mueller, G. & Plug, E.J.S. (2006) Estimating the effect of personality on male and female earnings, *Industrial & Labor Relations Review*, 60(1), pp. 3-22.
- Mundlak, Y. (1978) Pooling of time-series and cross-section data, *Econometrica*, 46(1), pp. 69-85.
- Nyhus, E.K. & Pons, E. (2005) The effects of personality on earnings, *Journal of Economic Psychology*, 26(3), pp. 363-384.
- O'Cleirigh, C., Ironson, G., Weiss, A. & Costa, P.T. (2007) Conscientiousness predicts disease progression (CD4 number and viral load) in people living with HIV, *Health Psychology*, 26(4), pp. 473-480.
- Plumper, T. & Troeger, V.E. (2007) Efficient estimation of time-invariant and rarely changing variables in finite sample panel analyses with unit fixed effects, *Political Analysis*, 15(2), pp. 124-139.
- Rotter, J.B. (1966) Generalized expectancies for internal versus external control of reinforcement, *Psychological Monographs*, 80(1), pp. 1-28.
- Ryff, C.D., Love, G.D., Urry, H.L., Muller, D., Rosenkranz, M.A., Friedman, E.M., Davidson, R.J. & Singer, B. (2006) Psychological well-being and ill-being: Do they have distinct or mirrored biological correlates?, *Psychotherapy and Psychosomatics*, 75(2), pp. 85-95.
- Senik, C. (2004) When information dominates comparison - Learning from Russian subjective panel data, *Journal of Public Economics*, 88(9-10), pp. 2099-2123.
- Vendrik, M.C.M. & Woltjer, G.B. (2007) Happiness and loss aversion: Is utility concave or convex in relative income?, *Journal of Public Economics*, 91(7-8), pp. 1423-1448.
- Winkelmann, L. & Winkelmann, R. (1998) Why are the unemployed so unhappy? Evidence from panel data, *Economica*, 65(257), pp. 1-15.

Winkelmann, L. & Winkelmann, R. (2008) Personality, work, and satisfaction: Evidence from the German Socio-Economic Panel, *The Journal of Positive Psychology*, 3(4), pp. 266-275.

Winter, D.G. & Barenbaum, N.B. (1999) History of modern personality theory and research, in: L.A. Pervin & O.P. John (Eds) *Handbook of personality* (New York, Guildford Press).

Note to Tables

Variable	Description
Life Satisfaction	A self reported measure of how satisfied the individual is with their life, all things considered, where 0=completely dissatisfied and 10=completely satisfied
Monthly Household Income (Euros)	The household's income in which the individual resides
Age	Individual's age
Female	Individual is female (excluded dummy: male)
Education (years)	Number of years of education
Household Size	The number of members in the individual's household
Married, Separated, Divorced, Widowed	Individual is married, separated, divorced or widowed (excluded dummy: single)
Self-Rated Health	Individuals are asked to give a self rating of their current health where 1=Very Good, 2=Good, 3=Satisfactory, 4=Poor and 5=Bad
Unemployed	Individual is unemployed (excluded dummies: any other responses to occupation position except retired)
Retired	Individual is retired (excluded dummies: any other responses to occupation position except unemployed)
Disabled	Disability status of the individual
Child dummy	Whether there is at least one child in the household (excluded dummy: no children in the household)
Personality variables	31 personality variables that measured 8 underlying constructs; openness-to-experience, conscientiousness, extroversion, agreeableness, neuroticism, individual autonomy, social responsibility and pessimism. See appendix for a full description of the personality questions and how questions grouped into constructs.

Table 1: Summary statistics across the 6 year panel used in analysis and a longer 12 year panel (N = 93016/135486) – non-standardised

Variable:		Mean	6 Year Panel		Mean	12 Year Panel	
			Standard Deviation	Between-to-within		Standard Deviation	Between-to-within
Life Satisfaction	Overall	7.00	1.74	1.29	6.98	1.73	1.16
	Between		1.38			1.32	
	Within		1.07			1.14	
Monthly Household Income (Euros)	Overall	2,662	1827.7		2,511	1644.6	
	Between		1779.9	2.17		1749.5	2.18
	Within		821.4			801.7	
Age	Overall	47.45	16.09	9.84	46.06	15.87	5.63
	Between		16.24			16.15	
	Within		1.65			2.87	
Female	Overall	0.52	0.50	n/a	0.52	0.50	n/a
	Between		0.50			0.50	
	Within		n/a			n/a	
Education (years)	Overall	12.07	2.64	8.87	11.91	2.61	3.84
	Between		2.66			2.61	
	Within		0.30			0.68	
Household Size	Overall	2.76	1.28	2.95	2.83	1.30	2.11
	Between		1.21			1.18	
	Within		0.41			0.56	
Married	Overall	0.65	0.48	3.29	0.66	0.48	2.5
	Between		0.46			0.45	
	Within		0.14			0.18	
Separated	Overall	0.02	0.13	1.11	0.02	0.13	0.90
	Between		0.10			0.09	
	Within		0.09			0.10	
Divorced	Overall	0.07	0.26	2.67	0.07	0.26	2.09
	Between		0.24			0.23	
	Within		0.09			0.11	
Widowed	Overall	0.06	0.23	3.67	0.05	0.22	3
	Between		0.22			0.21	
	Within		0.06			0.07	
Self-Rated Health	Overall	2.59	0.93	1.45	2.58	0.92	1.30
	Between		0.77			0.74	
	Within		0.53			0.57	
Unemployed	Overall	0.06	0.24	1.00	0.07	0.25	0.80
	Between		0.17			0.16	
	Within		0.17			0.20	
Retired	Overall	0.22	0.42	2.71	0.20	0.40	2.18
	Between		0.38			0.36	
	Within		0.14			0.17	
Disabled	Overall	0.10	0.30	2.08	0.10	0.30	1.79
	Between		0.27			0.25	
	Within		0.13			0.14	
Child dummy	Overall	0.32	0.47	2.39	0.34	0.47	1.75
	Between		0.43			0.42	
	Within		0.18			0.24	

Table 2: Fixed effect, REMT and pooled OLS life satisfaction regressions

Dependent Variable: Estimation type Independent Variables:	(1)	(2)	(3)
	Pooled OLS	Fixed Effect	REMT
Year Dummies	Yes	Yes	Yes
Regional Dummies	Yes	Yes	Yes
Log of Monthly Household Income (Euros)	0.331 (47.40)**	0.170 (15.94)**	0.165 (15.91)**
Age	-0.021 (16.00)**		-0.012 (6.31)**
Age squared/1000	0.272 (20.90)**		0.207 (10.83)**
Female	0.055 (9.62)**		0.066 (6.97)**
Education (years)	0.007 (5.51)**	-0.003 (0.44)	-0.001 (0.27)
Log of Household Size	-0.197 (19.96)**	-0.081 (4.81)**	-0.083 (5.54)**
Married	0.135 (12.47)**	0.041 (1.65)	0.065 (2.75)**
Separated	-0.255 (10.78)**	-0.123 (3.46)**	-0.137 (4.01)**
Divorced	-0.046 (3.29)**	0.101 (2.96)**	0.077 (2.36)*
Widowed	-0.031 (1.83)	-0.356 (7.86)**	-0.395 (9.15)**
Self-Rated Health (Standardised)	-0.414	-0.233	-0.240
Unemployed	(129.18)** -0.423 (34.72)**	(61.62)** -0.297 (23.13)**	(63.95)** -0.291 (23.03)**
Retired	0.061 (5.29)**	0.016 (1.00)	0.057 (4.38)**
Disabled	-0.087 (8.52)**	-0.087 (5.06)**	-0.037 (2.88)**
Child dummy	0.043 (4.90)**	0.057 (4.32)**	0.052 (5.03)**
Mean(Log of Monthly Household Income)			0.192 (12.44)**
Mean(Log of Household Size)			-0.051 (7.71)**
Mean(Married)			0.034 (1.17)
Mean(Separated)			-0.206 (3.11)**
Mean(Divorced)			-0.138 (3.39)**
Mean(Widowed)			0.423 (8.01)**
Mean(Self-Rated Health)			-0.289 (36.62)**
Mean(Unemployed)			-0.324 (9.72)**
Constant	-2.006 (32.46)**	-1.054 (6.17)**	-2.304 (22.15)**
Observations	93016	93016	93016
Number of Never Changing Person ID		17210	17210
R-Squared (within)		0.08	0.08
R-Squared (between)		0.27	0.41
R-Squared (overall)	0.27	0.19	0.29

Absolute value of t-statistics in parenthesis * significant at 5%; ** significant at 1%

Table 3: Predicting the fixed effects residual (from column 2 of table 2) using the mean levels of various objective characteristics and personality variables

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Fixed Effect Residual (from column 2 of table 1)				
Independent Variables (mean levels):					
Year Dummies	No	No	No	No	No
Regional dummies	Yes	Yes	No	Yes	Yes
Log of Monthly Household Income (Euros)					0.144 (12.18)**
Age	-0.023 (13.21)**	-0.027 (11.16)**		-0.016 (7.42)**	-0.010 (4.79)**
Age squared/1000	0.309 (16.91)**	0.326 (12.68)**		0.234 (10.12)**	0.183 (8.05)**
Female	0.041 (4.03)**	0.043 (4.28)**		0.065 (6.60)**	0.072 (7.51)**
Education (years)		0.035 (17.40)**		0.016 (8.29)**	0.001 (0.30)
Log of Household Size		-0.015 (2.82)**		-0.006 (1.38)	-0.022 (3.73)**
Married		0.072 (3.71)**		0.060 (3.45)**	0.069 (3.87)**
Separated		-0.291 (4.85)**		-0.263 (4.92)**	-0.199 (3.78)**
Divorced		-0.213 (7.91)**		-0.229 (9.49)**	-0.178 (7.35)**
Widowed		0.338 (10.63)**		0.307 (10.80)**	0.360 (12.53)**
Self-Rated Health (Standardised)					-0.195 (27.08)**
Unemployed					
Retired		-0.029 (1.09)		0.015 (0.63)	0.101 (4.23)**
Disabled					0.106 (5.36)**
Child dummy					-0.051 (3.09)**
Standardised Personality Variables: Constructs of the “Big Five”					
Openness			0.019 (3.59)**		
Original				-0.003 (0.57)	-0.003 (0.49)
Values artistic experiences				0.014 (2.80)**	0.012 (2.33)*
Active imagination				0.002 (0.38)	0.004 (0.75)
Conscientiousness			-0.002 (0.45)		
Thorough worker				0.005 (0.99)	0.002 (0.39)
Lazy				0.008 (1.49)	0.011 (2.27)*
Effective and efficient				0.021 (3.69)**	0.016 (2.94)**
Extrovert			0.021 (3.75)**		
Communicative				0.011 (1.84)	0.008 (1.36)
Sociable				0.026 (4.40)**	0.021 (3.72)**
Reserved				0.024 (4.75)**	0.019 (3.89)**
Agreeable			0.045 (8.20)**		
Rude to others				0.002 (0.38)	0.002 (0.34)
Forgiving nature				0.028 (5.62)**	0.027 (5.50)**
Considerate				0.015 (2.74)**	0.019 (3.52)**
Neuroticism			-0.053 (10.24)**		
Worries a lot				-0.078 (14.99)**	-0.050 (9.65)**
Nervous				0.012 (2.31)*	0.012 (2.26)*
Deals well with stress				0.034 (6.31)**	0.017 (3.28)**

Table 3 continues on the next page

Table 3 continued

Positive Reciprocity			0.029 (5.66)**		
Returns favours			0.022 (4.26)**	0.021 (4.15)**	
Helps others even at a cost			0.011 (2.09)*	0.002 (0.45)	
Reciprocates help			-0.001 (0.22)	0.003 (0.58)	
Negative Reciprocity			0.020 (3.77)**		
Takes revenge			-0.001 (0.16)	-0.003 (0.48)	
Offends others			0.002 (0.32)	0.002 (0.43)	
Causes problems for others			-0.003 (0.49)	-0.005 (0.73)	
Locus of control			0.140 (25.25)**		
Control over life			0.077 (15.22)**	0.069 (14.00)**	
Belief in luck			0.007 (1.42)	0.011 (2.25)*	
Influencing social conditions			0.028 (6.09)**	0.023 (5.20)**	
Others control their life			-0.054 (10.37)**	-0.048 (9.56)**	
Success comes from hard work			-0.020 (4.05)**	-0.019 (3.93)**	
Doubts own abilities			-0.017 (3.18)**	-0.018 (3.52)**	
Opportunities depend on social conditions			-0.016 (3.27)**	-0.012 (2.52)*	
Ability is more important than effort			0.006 (1.34)	0.008 (1.68)	
Little control in life			-0.044 (8.20)**	-0.036 (6.77)**	
Pessimism			-0.149 (28.74)**	-0.150 (29.82)**	-0.127 (25.62)**
Constant	0.444 (6.03)**	0.144 (1.76)	0.004 (0.79)	-0.071 (0.96)	-1.136 (11.13)**
Observations	17210	17210	17210	17210	17210
R-squared	0.07	0.10	0.18	0.29	0.33
Absolute value of t-statistics in parenthesis * significant at 5%; ** significant at 1%					

Table 4: Correlations between observable characteristics and the unobservable component of the fixed effect residual errors

Fixed Effect Residual controlling for:	(1) No controls	(2) Demographics	(3) Demographics and personality	(4) Demographics, personality and characteristics with low within-person variation
Error from predicting the fixed effect residual in:		Table 3 Column 1		Table 3 Column 2
Log of Monthly Household Income (Euros)	0.13**	0.16**	0.10**	0.09**
Age	0.14**	-0.01*	-0.01**	-0.01
Female	0.03**	-0.00	0.00	0.00
Education (years)	0.08**	0.12**	0.05**	0.00
Log of Household Size	-0.07**	0.01**	0.02**	0.03**
Married	0.04**	0.04**	0.05**	0.03**
Separated	-0.05**	-0.05**	-0.05**	-0.03**
Divorced	-0.11**	-0.10**	-0.12**	-0.02**
Widowed	0.15**	0.07**	0.07**	-0.01
Self-Rated Health (Standardised)	0.20**	0.26**	0.10**	0.09**
Unemployed	-0.13**	-0.10**	-0.06**	-0.05**
Retired	0.13**	-0.01**	-0.00	-0.00
Disabled	-0.03**	-0.08**	-0.04**	-0.04**
Child dummy	-0.08**	-0.01**	-0.01	-0.01

* significant at 5%; ** significant at 1%

Table 5: Introducing personality into life satisfaction regressions using the fixed effect vector decomposition technique (3rd stage) and the random effects model

	(1)	(2)	(3)	(4)
Dependent variable:	Life Satisfaction (Standardised)			
Independent variables:				
Year dummies	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes
31 Personality Variables	No	Yes	No	Yes
Log of Monthly Household Income (Euros)	0.178	0.288	0.276	0.238
	(36.53)**	(57.76)**	(33.41)**	(29.45)**
Age	-0.025	-0.015	-0.021	-0.015
	(28.17)**	(16.06)**	(11.41)**	(8.33)**
Age squared/1000	0.305	0.221	0.274	0.221
	(33.78)**	(24.04)**	(14.59)**	(12.42)**
Female	0.044	0.067	0.051	0.066
	(11.09)**	(15.31)**	(5.40)**	(6.89)**
Education (years)	0.030	0.001	0.016	0.006
	(34.98)**	(1.65)	(8.66)**	(3.14)**
Log of Household Size	-0.109	-0.129	-0.148	-0.109
	(15.94)**	(18.61)**	(12.13)**	(9.24)**
Married	0.100	0.101	0.098	0.094
	(13.30)**	(13.24)**	(6.55)**	(6.65)**
Separated	-0.270	-0.223	-0.198	-0.203
	(16.46)**	(13.43)**	(7.63)**	(8.06)**
Divorced	-0.089	-0.064	-0.029	-0.045
	(9.15)**	(6.49)**	(1.49)	(2.39)*
Widowed	-0.069	-0.060	-0.110	-0.118
	(5.95)**	(5.07)**	(4.50)**	(5.10)**
Self-Rated Health (Standardised)	-0.235	-0.332	-0.308	-0.279
	(102.62)**	(141.91)**	(92.25)**	(83.99)**
Unemployed	-0.296	-0.295	-0.345	-0.326
	(35.00)**	(34.42)**	(29.12)**	(27.99)**
Retired	-0.001	0.059	0.026	0.031
	(0.19)	(7.38)**	(1.99)*	(2.44)*
Disabled	-0.091	-0.070	-0.134	-0.102
	(12.87)**	(9.69)**	(10.51)**	(8.29)**
Child dummy	0.058	0.019	0.043	0.032
	(9.62)**	(3.13)**	(4.13)**	(3.20)**
Error from predicting the fixed effect residual in:	Table 3 Column 2	Table 3 Column 5		
	0.999	1.001		
	(316.84)**	(283.39)**		
Constant	-0.968	-2.220	-1.681	-1.852
	(22.53)**	(42.41)**	(20.03)**	(19.40)**
Observations	93016	93016	93016	93016
R-Squared (within)			0.08	0.08
R-Squared (between)			0.40	0.48
R-Squared (overall)	0.65	0.64	0.26	0.33

Absolute value of t-statistics in parenthesis * significant at 5%; ** significant at 1%