# The Benefits, Challenges and Insights of a Dynamic Panel assessment of Life Satisfaction

Piper, Alan T.

Europa-Universität Flensburg

October 2014

# The Benefits, Challenges and Insights of a Dynamic Panel assessment of Life Satisfaction

Alan T. Piper*, **

Europa-Universität Flensburg

Internationales Institut für Management und ökonomische Bildung

Munketoft 3B, 24937 Flensburg, Deutschland

October 2014

Abstract: This study discusses and employs dynamic panel data to investigate life satisfaction. A key result is that approximately 90% of the impact of any commonly measured variable on well-being is contemporaneous. This reflects the finding that lagged life satisfaction has a small, positive and significant effect on current life satisfaction. A related key benefit of dynamic panel models is the ability to determine short and long run values of coefficients. Additionally such models make it possible for researchers to choose which explanatory variables are potentially endogenous or exogenous. The challenges of such an analysis are also detailed, which are linked to its complexity particularly when compared with the more standard fixed effects models.

Keywords: Life Satisfaction, Dynamic Panel Analysis, GMM, Happiness, Subjective Well-Being

JEL codes: C23, I31

---

*Email address for correspondence: alan.piper@uni-flensburg.de

** For useful comments I am grateful to Nick Adnett, Andrew Clark, Clemens Hetschko, Geoff Pugh, Anita Ratcliffe and Christoph Wunder as well as two anonymous reviewers. I also thank, and have benefited from suggestions from, Yannis Georgellis, Gerd Grözinger, and Mehtap Hisarciklilar, Maarten Vendrik and participants at the 2014 European Association of Labour Economists conference (Ljubljana). The British Household Panel data used here were made available by the UK data archive. Neither the original collectors of the data nor the Archive bears any responsibility for the analyses or interpretations presented here.

The Benefits, Challenges and Insights of a Dynamic Panel assessment of Life Satisfaction

## 1. Introduction

Dynamic panel models have become increasingly popular in many areas of economic enquiry, and their use has provided new insights. Some recent examples include investigations into corporate finance (Flannery and Hankins 2013), economic growth (Lee et al. 2012), and foreign aid (Dutta et al. 2013) as well as the relationship between school expenditure and school performance (Pugh et al. 2014). This increase in use is due, in part, to increasingly sophisticated software which has followed a greater theoretical understanding of dynamic panel analysis. As an indication of their popularity, key papers for the development of these models have (at the time of writing) several thousand citations. Despite this popularity in economics generally, the use of such models in well-being research is sparse, and this paper both describes the benefits and challenges of this use in a well-being context, and employs the model to provide new insights for the understanding of well-being.

Dynamics are often modelled by including a lag of the dependent variable on the right-hand side of the regression equation. Such an inclusion changes the interpretation of the right-hand side variables, which now indicate contemporaneous correlations conditional on the history of the model. The history of the model is itself contained within the lagged dependent variable (see appendix 1 for the algebra). Discussed in more detail below, together they enable researchers to find overall (long-run) coefficients for the explanatory variables as well as the contemporaneous (or short-run) ones. Very much connected with this is the possibility to determine the influence of the past, which the lagged dependent variable represents.

The dynamic model used here is System General Method of Moments (GMM) (also discussed more in section 2) and this enables the explanatory variables to be treated as potentially endogenous or exogenous. This is potentially important for well-being, enabling the investigation of variables that may once have been *verboten* (as well as being better suited than standard models for determining coefficients for time invariant variables). As well as being informed by theory, there are diagnostic tests available to help with this choice of endogeneity or exogeneoity. Furthermore, system GMM does not have to be dynamic so this benefit is available to researchers who are not interested in dynamics, though a central argument of this paper is that dynamics are interesting and should be given consideration even if they are not ultimately employed.

The challenges of a dynamic panel analysis provide reasons why researchers may not ultimately undertake such an analysis. This paper provides a frank illustration of some of the challenges of a dynamic panel analysis for well-being. The complexity of the model and its diagnostic tests are the main sources of these challenges, and a reason why some previous well-being studies using this model are not fully successful. A particular weakness with these models is that its estimations are computationally intensive and memory hungry and this can mean large samples cannot be estimated. To 'solve' this problem, this paper has split the sample by gender, estimating the equations separately for males and females. Due to initially poor diagnostic test results in the female case, other female samples were used with differing diagnostic test result outcomes (which themselves reflect different samples and different choices of endogeneity and exogeneity). The results obtained for the various coefficients, however, are robust to different diagnostic test outcomes. A recommendation is made, in section 3 when this is discussed, that, due to the seeming difficulty of passing all the diagnostic tests and the novelty of such estimations in the 'life satisfaction' area, researchers should do the following: test different samples; present all of the results from the different samples; present all of the different diagnostic test results; and make sure that the discussion of the results emphasises the degree of caution necessary (if any) despite the apparent robustness of the results. The results discussion of section 3 is an example of this.

Despite the substantial challenges and complexity, there are rewards from a dynamic panel assessment of life satisfaction.[1] With the more common fixed effects analysis, it would not be possible to learn about the influence of the past, the contemporaneous associations of the explanatory variables and well-being (and the relative importance of contemporaneous effects), and similarly, demonstrate the short- and long-run effect of these right-hand side variables. [2] Furthermore, dynamic panel methods (as a species of random effects models) can incorporate both within and between variation to determine the coefficients in the model, more efficiently estimating coefficients that are near time invariant. Also, when using fixed effects, the results cannot be generalised out of the sample to the wider population (though this is often done implicitly or explicitly). This problem of the impossibility of generalising out of the sample is not shared with the dynamic model employed here. The typical random effects problem, and the reason why fixed effects estimation is usually preferred with well-being panel work, is that the fixed effects appear in the error term and are generally correlated with one or more right hand side variables. However, with General Method of Moments estimation, suitable instruments can easily be found for the right hand side variables which are then tested for exogeneity with respect to the error term (which, as mentioned, includes any individual specific effects). This is what the particular chosen dynamic model, discussed in section 2, does enabling out-of-sample generalisation and other benefits.

The results from dynamic panel estimation are detailed in the results section, section 3, and discussion section, section 4. Dynamic panel analysis deserves wider consideration in the 'economics of happiness' literature. That such models can, in a straightforward fashion, let some variables be endogenous with life satisfaction adds substantially to their usefulness.[3] Furthermore, in the

---

[1] With more work (and subsequently more understanding of the dynamics of life satisfaction) some of these challenges may subside somewhat. This article represents, in part, a step on this journey.

[2] Moreover, tests for serial correlation in the idiosyncratic part of the error term almost always reject the null of no serial correlation: a strong sign that there are dynamics omitted from static panel analyses.

[3] A recent example of an investigation taking advantage of this ability is Piper (2014a) where the fear of the future is assessed with respect to well-being. Such fears are likely to be endogenous, however via GMM analysis this was instrumented for (in a fashion typical for GMM analysis) and the instruments successfully tested as exogenous and hence suitable for estimation.

discussion section it is explained how these dynamic results provide potential explanations for some existing results in the literature regarding the well-being legacy of different events. These sections (3 and 4) follow section 2 which details some of the methodological issues of a dynamic panel assessment of well-being, as well as discussing the diagnostic tests. Section 5 concludes. Finally, for comparison purposes, appendix 2 contains the results from a fixed effects analysis of the same data.

## 2. Data and Methodology

This section starts with a brief description of the dataset and sample used, before moving on to discuss, in two subsections, aspects of methodology with respect to dynamic panel modelling: the first subsection explains why system GMM was chosen; the second discusses the diagnostic testing for such a model.

The data come from the British Household Panel Survey (BHPS), a nationally representative survey, which was established in 1991. Widely used in the literature on the 'economics of happiness', it is a major source of micro-level panel data in the UK with the same representative sample of individuals interviewed repeatedly over a period of years. From 1996, the BHPS contains a direct satisfaction question where the interviewee is asked 'how dissatisfied or satisfied are you with your life overall' with possible responses running from 1 to 7 representing not satisfied at all to completely satisfied. The sample used in this investigation uses everyone in the dataset from the years 1996 to 2007, aged between 15 and 60. This represents over 100,000 person-year observations. As discussed later, the regressions are often estimated separately by gender because the dynamic panel model used is computationally intensive. The instruments created by such models, as well as the in-built diagnostic tests, are a cause of this computational intensity and these are discussed in subsections 2.1 and 2.2.

### 2.1 Choosing a suitable estimator.

Reasons were advanced in the introduction for a consideration of dynamics in an analysis of well-being. And as Bond states, even when the dynamics themselves are not of direct interest "allowing for dynamics in the underlying process may be crucial for recovering consistent estimates of other parameters" (2002, p.1. see also p.20). Adding dynamics to the model is usually undertaken by including a lag of the dependent variable as a right hand side variable. Hence, what is estimated is the following standard equation (with the other explanatory variables excluded for clarity):

$$y_{it} \quad = \beta y_{i,t-1} + (v_i + \epsilon_{it}) \qquad (1)$$

As this is a panel model each observation is indexed over $i$ (= 1…N) cross-section groups (here, individuals) and $t$ (= 1…T) time periods (here, annual observations). Equation 1 is a first-order dynamic panel model, because the explanatory variables on the right-hand side include the first lag of the dependent variable ($y_{i,\ t-1}$). The composed error term in parentheses combines an individual-specific random effect to control for all unobservable effects on the dependent variable that are unique to the individual and do not vary over time ($v_i$), which captures specific ignorance about individual $i$, and an error that varies over both individuals and time ($\epsilon_{it}$), which captures our general ignorance of the determinates of $y_{it}$. However, this cannot be estimated accurately by OLS or by fixed effects estimation. An OLS estimator of $\beta$ in equation 1 is inconsistent, because the explanatory variable $y_{i,t-1}$ is positively correlated with the error term due to the presence of individual effects. A fixed effects estimation does not have this inconsistency because the equation is transformed to remove the individual effect, as in equation 2.

$$y_{it} - y_{i,t-1} = \beta\left(y_{i,t-1} - y_{i,t-2}\right) + (\epsilon_{it} - \epsilon_{i,t-1}) \qquad (2)$$

However, equation (2) exhibits the different problem of correlation between the transformed lagged dependent variable and transformed error term. Here the overall impact of the correlations is negative, and is the well-known Nickell (1981) bias. Bond (2002) states that these biases can be used to provide an informal test for an estimator of the lagged dependent variable: the estimated coefficient should be bounded below by the outcome from OLS (which gives the maximum upwards bias) but above by the fixed effects estimate (which gives the maximum downwards bias).[4]

Due to these problems, the standard approach is to find a suitable instrument that is correlated with the potentially endogenous variable (the more strongly correlated the better), but uncorrelated with $\varepsilon_{it}$. Because, with GMM, instrumentation is not confined to one instrument per parameter to be estimated, the possibility exists of defining more than one moment condition per parameter to be estimated. It is this possibility that is exploited in the GMM estimation of dynamic panel models, first proposed by Holtz-Eakin et al. (1988).[5] The two models popularly implemented are the "difference" GMM estimator (Arellano and Bond, 1991) and the "system" GMM estimator (Arellano and Bover 1995). Greene (2002, p.308) explains that suitable instruments come from within the dataset: the lagged difference ($y_{it-2} - y_{it-3}$) ;and the lagged level $y_{it-2}$. Both of these should satisfy the two conditions for valid instruments, since they are likely to be highly correlated with ( $y_{i,t-1} - y_{i,t-2}$ ) but not with $(\varepsilon_{it} - \varepsilon_{i,t-1})$. It is this easy availability of such "internal" instruments (i.e., from within the dataset) that the GMM estimators exploit. The "difference" GMM estimator follows the Arellano and Bond (1991) data transformation, where *differences are instrumented by levels*. The "system" GMM estimator adds to this one extra layer of instrumentation where the original *levels are instrumented with differences* (Arellano and Bover 1995). Here, for three main reasons, system GMM is used rather than difference GMM. Firstly, system GMM allows for more instruments and can dramatically improve efficiency (compared to difference GMM) (Roodman 2009, p.86). Secondly, any gaps in a panel – and this BHPS sample is unbalanced - are magnified by difference GMM (when compared to

[4]This bias has been misunderstood in some of the well-being work which estimates similar equations. Della Giusta et al. (2010) state that the biases are general, and "therefore, we have reported both of the [whole of] OLS and fixed effects results as a comparison (both of which do not include a lagged dependent variable)" (p.10). This is also wrong because the coefficients for the independent variables from dynamic GMM analysis and those from OLS and fixed effects are not referring to the same things, and should not be directly compared. This is an important point for dynamic panel analysis, and is discussed later to aid the results interpretation (and subsequent discussion).

[5]GMM was developed by Lars Peter Hansen, work that led, in part, to him being selected as one of the three Nobel Prize winners for Economics in 2013. See Hansen (1982) for more information on the General Method of Moments,or Hall (2005) for a detailed textbook treatment.

system GMM. Indeed this was a motivating factor for the creation and development of system GMM) (Roodman 2009, p. 104). And thirdly, unlike difference GMM, system GMM does not expunge the fixed effects (which are important in a well-being context) (Roodman 2009, p.114). These estimators, unlike OLS, FE and RE estimation, do not require distributional assumptions, like normality, and can allow for heteroscedasticity of unknown form (Verbeek, 2000, pp. 143 and 331; Greene, 2002, pp.201, 525 and 523). A more extensive discussion of these methods is beyond the scope of this investigation, but the references provided above and papers by Roodman (e.g. 2006, 2007, and 2009) are very informative.[6]

Before estimating any dynamic panel model there are two important (and linked) considerations. Firstly, which of the regressors are to be treated as potentially endogenous and which strictly exogenous? Secondly, how many instruments to use? With happiness equations some of the regressors are potentially endogenous: does marriage, for example, make someone happy or are happy people more likely to get married (or are both determined by underlying but omitted variables)? There is (as yet) little theoretical guidance to help with this decision, though some evidence that marriage is potentially endogenous (Stutzer and Frey 2006). Arguments could be advanced for income and health being endogenous variables too. Diagnostic tests are available and built in with *xtabond2*, the Stata command employed for the empirical analysis, to help with this choice. The actual choice made is based on theoretical considerations along with the diagnostic test outcomes (which are explained in more detail below). Here, such testing (along with a consideration of the likely relationships between life satisfaction and the right-hand side variables) led, for males, to the treatment of marital status only as potentially endogenous, and everything else treated as endogenous. For females, there are no suitable outcomes (in terms of the diagnostic testing) regarding which variables should be treated as endogenous and which exogenous. Initially, and consistent with the male estimation, marital status only is treated as endogenous for females. In

[6] The Roodman papers are particularly useful for applied researchers because they explain how to use the Stata software programme, xtabond2, that he created to implement the GMM dynamic estimators.

another regression health and income are also treated as endogenous. As explained below, the results are consistent despite differing diagnostic test outcomes.

## 2.2 Diagnostic tests

The choice of which regressors are to be treated as endogenous and exogenous is coupled with the consideration of how many instruments should be used, because that choice, in part, generates the instruments. A high number of regressors treated as endogenous means that a higher number of instruments are employed, ceteris paribus. Researchers can also affect the instrument count by changing the lag length to be used for instrumentation, and good practice is to test results for their robustness to different lag length choices (and hence different instrument counts).[7] Diagnostic tests are available for the appropriateness of the instrumentation collectively, and also for the subsets of instruments created by the regressors that are treated as exogenous or endogenous, as well as those generated by the lagged dependent variable. (Indeed, with *xtabond2* any subset of instruments can be tested, should the researcher want or need to.) These tests are asking whether the instruments are exogenous to the error term, and are returned to below.

Additionally, *xtabond2* contains a built in check on first and second order autocorrelation in first differences, which is an additional check on the appropriateness of the instrumentation.[8] For this investigation, the "system" GMM estimation was undertaken with the sample separated by gender. The reason is wholly pragmatic: such estimations are computationally intensive and it was not possible to perform the estimate for the whole sample.[9] The diagnostics of the chosen models should indicate that first order autocorrelation is present, but second order is not. Specific outcomes

[7]In a life satisfaction context, this choice appears to makes little difference to the subsequently obtained coefficients, but can matter for the autocorrelation diagnostic tests (discussed below). The association of well-being with the various right-hand side variables are robust to different lag length choices.

[8] Recall the explanation presented above utilising Greene (2002), regarding suitable instruments.

[9]Every dynamic regression both shown here, and undertaken as part of the diagnostic testing, employed the twostep robust procedure that utilises the Windmeijer (2005) finite sample correction for the two-step covariance matrix. Without this, standard errors have been demonstrated to be biased downwards (Windmeijer 2005).

from the regressions undertaken in this investigation are presented below along with the outcomes from other diagnostic tests including the Hansen *J* and *C* test.

The Hansen (1982) test *J* statistic[10] has a null hypothesis of exogenous instruments and refers to all of the instruments collectively. Rather than rejecting or (not rejecting) the null hypothesis with the typical value of 0.05, Roodman offers what he calls a 'common sense' value instead. Roodman's recommended minimum threshold is a p-value of at least 0.25 and he (2007, p.10) warns that researchers :

> should not view a value above a conventional significance level of 0.05 of 0.10 with complacency. Even leaving aside the potential weakness of the test, those thresholds are conservative when trying to decide on the significance of a coefficient estimate, but they are liberal when trying to rule out correlation between instruments and the error term. A *p* value as high as, say, 0.25 should be viewed with concern. Taken at face value, it means that if the specification is valid, the odds are less than 1 in 4 that one would observe a *J* statistic so large.

Thus, the J tests, Hansen and Sargan, inspect all of the generated instruments together, with a null hypothesis of exogenous instruments. Low p-values mean that the instruments are not exogenous and thus do not satisfy the orthogonality conditions for their use. Within the well-being area, some of the GMM studies do not test (or at least report) the Hansen J test result, risking what Sargan calls, more generally, a 'pious fraud'. (Godfrey 1991, p.145). Other dynamic well-being studies report a very low p-value and incorrectly assert that this indicates that the instruments are appropriate for estimation.[11] Discussed below, some of the estimates the p-value for the Hansen *J* test are low and thus caution is attached to those results, no matter how plausible they seem.

---

[10]This has the advantage over the Sargan *J* test (also reported by default) because it works in the presence of heteroscedasticity. Indeed, if the errors are believed to be homoscedastic then the Hansen test is the same as the Sargan test.

[11]Bottan and Perez-Truglia (2011), for example, report p-values of $<0.001$ (Table 1A) and incorrectly state that they cannot "reject the null of the Sargan test at the 1% level" (p.230). This value, however, is a strong rejection of the null. In this study, only once is the p-value of the Sargan test above 0.25. However, this may not necessarily invalidate all of the results because, for the reason put forward in footnote 11, the Hansen test

Valuable, but perhaps even more neglected in the well-being GMM literature, are the difference-in-Hansen (or C) tests. These are diagnostic tests that inspect the exogeneity of a particular subset of instruments, and are, by default, reported by *xtabond2*.[12] Thus, this means that researchers can test their choices (and alternative choices) of which regressors should be treated as exogenous and which endogenous. [13] This decision crucial since it can affect the overall *J* test result and, as mentioned above, the choice can alter somewhat the coefficients obtained for the independent variables (although not qualitatively the lagged dependent variable). This *C* test is well explained in Baum et al. (2003, sections 4.2 and 4.4) as well as the Roodman papers referred to above.

The difference-in-Hansen tests also inspect the 'initial conditions' problem, which refers to the relationship between the unobserved fixed effects and the observables at the time of the start of the panel subset employed. For estimation to be valid, it is necessary that changes in the instrumenting variables are uncorrelated with the individual-specific part of the error term. This is tested by the difference-in-Hansen GMM test for levels, reported by *xtabond2.* Roodman (2009, section 4) discusses this, and in the conclusion of the same article offers advice regarding what diagnostic tests should be reported along with the results: "several practices ought to become standard in using difference and system GMM. Researchers should report the number of instruments generated for their regressions. In system GMM, difference-in-Hansen tests for the full set of instruments for the levels equation, as well as the subset based on the dependent variable, should be reported" (Roodman 2009, p.156).

---

(unreported) is the more appropriate *J* test. Powdthavee (2009) reports the Hansen version of the *J* test, but the p-values are often under 0.25. In that article there is a supporting claim that values between 0.1 and 0.25 are within Roodman's (2007) acceptable range: as we can see from the Roodman quote just above this is incorrect.

[12]It does this by re-estimating the Hansen *J* test without the subset of interest, and comparing the result with that for the overall (full instrumentation) Hansen test.

[13]Wunder (2012) does not discuss this decision but treats all the regressors as exogenous. Whether this is appropriate or not, it is impossible to judge from the study. This may be a consequence of the paper's brevity: published in Economic Letters it is just over two pages long. Della Giusta et al. (2010), follows Powdthavee (2009) in treating all of the independent variables as endogenous apart from the age and wave dummies. Their reported *J* test result suggests that, for females, like the outcome here, this is likely to be invalid. Here, as table 1 shows, the first attempt at estimating female life satisfaction also has this problem.

As recommended these are presented in the results table of the next section. Importantly, the next section commences with a discussion regarding how the coefficients need to be interpreted. An understanding of the interpretation of the coefficients, and particularly the coefficient on the lagged dependent variable, is important generally, and for the subsequent discussion in Section 4.

## 3. Results

This section presents and discusses the results from dynamic panel estimation, after an explanation of how the coefficients need to be interpreted, and then proceeds to discuss the diagnostic test results. Regarding interpretation, a footnote above states that coefficients obtained via OLS or FE are substantially different from those obtained by dynamic panel methods and cannot directly be compared and this is now explained. As Greene asserts

> *Adding dynamics to a model … creates a major change in the interpretation of the equation*. Without the lagged variable, the "independent variables" represent the full set of information that produce observed outcome $y_{it}$. With the lagged variable, we now have in the equation the entire history of the right-hand-side variables, *so that any measured influence is conditional on this history*; in this case, any impact of (the independent variables) $\boldsymbol{x}_{it}$ represents the effect of *new* information. (2008, p.468, emphasis added).

Thus, in a dynamic panel model, the 'independent variables' only reflect new or contemporaneous information conditional both on the other controls and the lagged dependent variable, which itself represents the history of the model (i.e. the past). This means that contemporaneous associations of variables with life satisfaction can be usefully assessed via dynamic panel methods, whereas anything historic (perhaps, for example, education) is captured in the 'black box' of the lagged dependent variable itself.[14]In the appendix, the lagged dependent variable is shown algebraically to be the entire history of the model and not just a fixed effect (as sometimes assumed).

[14]Piper (2013) discusses this in more detail along with the implications for modelling.

Table 1 displays the results for four estimations, one of which is for males and three are for females. The explanation of the table starts with males, as this is easier to explain (and perhaps understand), and then proceeds onto the other three columns. For males, the estimation uses default instrumentation, i.e. it uses all available lags as instruments, utilising the full length of the sample. Furthermore, as discussed in the previous section, only marital status is treated as potentially endogenous. The coefficients obtained are robust to other choices of lag length which start at the first available lag and do not employ every additional available lag (unlike default instrumentation).

[TABLE ONE ABOUT HERE]

For males, positive and statistically significant for life satisfaction are real annual income (though the size is negligible with an income increase of £1000 increasing life satisfaction by less than 0.002), marriage, health (both self-reported as good or excellent relative to a dummy variable capturing fair health and worse responses); negative and statistically significant for male life satisfaction are unemployment, being long-term sick or disabled, being a family carer, having a labour force status as other[15] and medium and high levels of education, as assessed by qualifications obtained. The coefficients on the age-range dummy variables are in line with the well-known U shape. The coefficient obtained for the lagged dependent variable is discussed below. These results are robust to the number of instruments used which, for most variables, give qualitatively the same outcome. In the male case, the diagnostic tests are all supportive of the estimation choices made. Second order autocorrelation is ruled out, and the p-values for the *J* and *C* tests are above Roodman's 'common sense' minimum of 0.25 (as discussed in the previous section).

For females, there are three columns of results (reflecting differences in the diagnostic test outcomes, discussed just below). The first column is every female in the sample, only marital status is treated as potentially endogenous, and the diagnostics of this estimation highlight that the

15 This might be caring for someone on maternity leave, on a government training scheme or one of a handful of people in the dataset who fit none of the possible labour force categories.

instruments created are invalid.[16] Second order autocorrelation cannot be ruled out, and the null of instrument validity for the whole set of instruments (the *J* test) can be rejected with a 0.053 chance of error. The *C* test for valid instruments created for the lagged dependent variable can be rejected with a chance of error less than 0.01. Thus for the second column (the first female column) the instruments are endogenous with the error term and therefore invalid. Any discussion of the results from the second column needs a large caveat. The problem regarding the presence of second order autocorrelation can be solved by using longer lag lengths (i.e. starting further back in the dataset) but this is only a technical solution. The AR(2) test would then result in a preferred outcome, but how appropriate is it to instrument for life satisfaction levels (and other explanatory variables) and differences, the differences and levels of at least two years previously? There is a debate in the wider literature about weak and strong instrumentation, and not just valid and invalid instrumentation (Clemens et al 2004; Bazzi and Clemens 2009). However, this concern over weak – as opposed to valid - instruments in (difference and) system GMM estimation, and particularly regarding corresponding solutions, still seems to be at a rather tentative stage, with no agreed approaches. Different samples result in different diagnostic test outcomes. The third column (in table 1) focuses on females aged between 15 and 35 and has similarly valid instrumentation.

When restricting the sample to those females aged 35 and under, the four diagnostic tests support the instruments used for estimation: the various null hypotheses of exogenous instruments are supported (not rejected) in each case. Here, again, only marital status was treated as potentially endogenous. The final column treats health and income as potentially endogenous as well as marital status, and extends the sample's age range upwards to 50. For the final column of results, three of the four diagnostic tests indicate exogenous instruments, and one test – the *C* test for the lagged dependent variable - indicates that some caution is necessary. This last column is a good example of the need to not stop diagnostic testing with AR(2) and the *J* test (which is, in the main, as far as the

[16] The diagnostic problems for GMM estimation regarding females in the BHPS are also found by Della Giusta et al (2010). In that paper, the null hypothesis of having exogenous instruments overall (i.e. Hansen J test) is comfortably rejected.

most conscientious dynamic panel work goes in the well-being area). Subsets of instruments should also be investigated. Despite the differences in the diagnostic test results in the three female columns, the age ranges examined, and the differing choice of what is potentially endogenous the coefficients obtained are very similar and, while not directly comparable, similar to those obtained by fixed effects (discussed below).[17]

In table 1, for females (based on the consistency of results from all three estimates), positive and statistically significant for life satisfaction are the following: being married, reporting health as good or excellent, and having a labour force status as other. This latter effect appears to reflect maternity leave, which may be the reason for the different sign when compared to males.[18] Negative and statistically significant for female well-being (again in all three estimates) are the following: unemployment, being long-term sick or disabled and being a family carer. Once again, the age coefficients are in line with the U-shape finding. For females in the younger age range only, education has a positive effect on life satisfaction, perhaps reflecting the possibility that any well-being effects of education, on average, fade as individuals age. None of these results – for females and males – are surprising, and the results from dynamic panel analysis support, reasonably well, results from most fixed effects analyses in the well-being area.

Not yet discussed is the lagged dependent variable, and its obtained coefficient. This provides the central insight of this investigation enabling the determination of the influence of the past history of the model and is necessary for the calculation of and the ratio between the contemporaneous effect and historical effect of the explanatory variables. This is discussed in detail in the next section, along with greater consideration of these contemporaneous coefficients.

[17] This similarity perhaps indicates that researchers should report the results and add a caveat regarding the diagnostic results rather than just dismissing the obtained results (or not reporting all of the diagnostics). A second best solution is to demonstrate robust results to differing diagnostic outcomes, rather than a first best outcome of perfect diagnostic results, which is perhaps not possible with valuable panel data surveys.
[18] See D'Addio et al. (2013) for more information regarding the well-being effect of maternity leave and other birth-related policies.

4. DISCUSSION

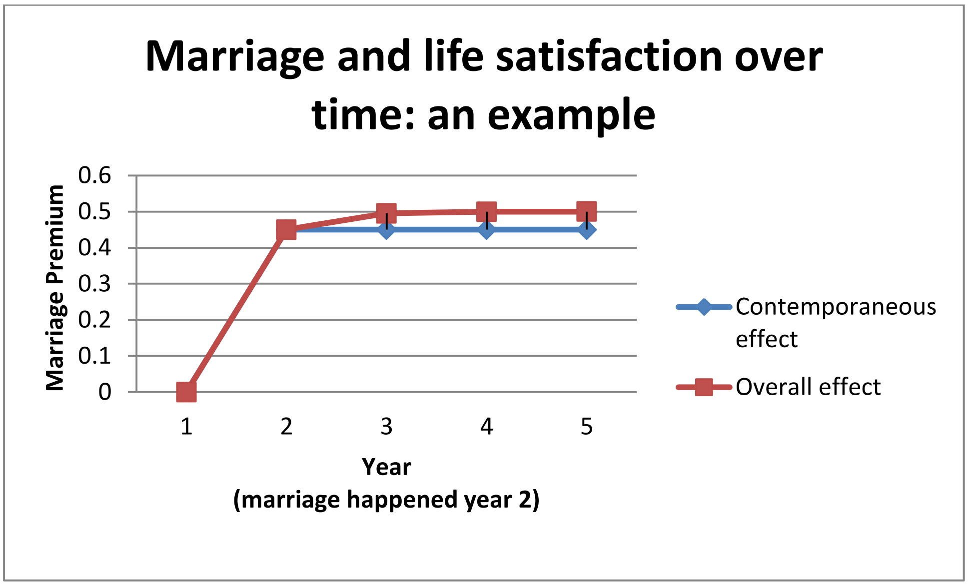
A key finding from the results of table 1 is the coefficient obtained for lagged dependent variable. In all four columns it is small (around 0.1), positive and significant. This, as the quote from William Greene (in section 3) and the algebra (in the appendix) shows, represents the entire history of the model. Thus, the entire history of the model has only a small influence (0.1) on current life satisfaction, indicative of life satisfaction being largely contemporaneous. Much of what contributes to life satisfaction are current circumstances and events, with this small influence from the past. The result that determines this, the 0.1 value, is robust being found in other studies. To a greater or lesser degree, every study mentioned previously that uses GMM for dynamic estimation finds a small, positive coefficient for the lagged dependent variable (Powdthavee 2009; Della Giusta et al 2010; Bottan and Perez-Truglia 2011; Piper 2012; Wunder 2012)[19].More recently, the GSOEP has been used to investigate the impact of how individuals perceive the future in general, taking advantage of GMM's ability to create exogenous instruments for potentially endogenous explanatory variables, and also finds a small, yet significant positive influence of lagged life satisfaction on current life satisfaction (Piper 2014). These similar results from the different studies for the lagged dependent variable are obtained despite many differences including: the equation estimated; the datasets employed; alternate choices of exogneneity and endogeneity; diagnostic test results (and their differing appropriateness); and the use of lags for other independent variables.

[19]Powdthavee (2009) does not consistently find a significant effect of lagged life satisfaction, however as mentioned previously the estimations do not exhibit good diagnostic test results. In the estimations that are closest to those of this investigation, (columns 7 and 8 of Table 2) he finds a small, positive significant effect of past life satisfaction of current life satisfaction. The empirical results of Bottan and Perez-Truglia (2011) for the (Arellano-Bond) autoregressive happiness estimates (Tables 1A-1D), based on panel data from four countries (Britain, Germany, Japan and Switzerland) overwhelmingly find a small positive and statistically significant coefficient. Piper (2012) has also found a very similar coefficient for lagged life satisfaction for the twenties age range, fifties age range, and when using the Caseness and Likert General Health Questionnaire composites as a proxy for life satisfaction. With German data Wunder obtains almost exactly the same coefficient as those reported in section 4 in regressions that do not employ the additional lags of the dependent variable. This is not reported in Wunder (2012) because it is not diagnostically appropriate, there is AR(2) serial correlation in the such estimates with the GSOEP (email correspondence). Running similar estimates, I have also found figures around 0.1 to 0.12 for various estimations using the GSOEP too, but like Wunder's work the diagnostics do not sufficiently support the estimation.

The introduction mentioned the property of dynamic panel analysis in splitting up contemporaneous (or short-term) effects and overall (or long-term effects), and this splitting up highlights the finding that life satisfaction responses largely reflect contemporaneous concerns. Overall effects can be found via a quick post-estimation calculation: the contemporaneous coefficient divided by 1 minus the lagged dependent variable. Taking marriage as an example, the contemporaneous effect of marriage for males (column 1 of table 1) compared to being single is approximately 0.45, indicating that married people are, on average, nearly half a point more satisfied with life (ceteris paribus) on the BHPS's 1-7 scale. The overall value (or long-run value) is of 0.5 (which comes from 0.45/1-0.1), approximately 90% of which reflects the contemporaneous effect of being married. Being married in the past contributes only a small amount to life satisfaction.[20] The following graph illustrates the life satisfaction premium for marriage over time (assuming the marriage takes place in year 2). The analysis here supports Qari (2014) and indicates that marriage is not adapted to, or to restate in a different (preferred way) being married contributes, on average, to well-being sometime after getting married.

[20] To help illustrate what this means, in practical terms, I paraphrase a colleague (who will remain anonymous, and is not necessarily listed in the acknowledgements): "I think that's right. Most of the well-being effect of marriage for me is being married currently. I feel a residual satisfaction that I have found someone who has put up with me for nearly thirty years – a small sense of satisfaction in that – but most of the happiness effect, for me, is in being married now."

Figure 1: illustration of life satisfaction effect of marriage




The size of the coefficient for the lagged dependent variable influences the ratio between the contemporaneous influence and the 'historical influence'. A coefficient size of 0.05 would further decrease the influence of the past, increasing the contemporaneous effect of marriage; a coefficient of 0.4 would have the opposite effect: the past would matter considerably, nearly as much as the contemporaneous effect. As it is, at 0.1, much of any well-being effect of being married (unemployed, health, or any other statistically significant variable) is contemporaneous. Unemployment, as another example, is very similar: the contemporaneous coefficient for males is -0.43, so the long run coefficient is approximately -0.48. Much of the negative impact of being unemployed is contemporaneous, and like marriage the impact is cumulative over time. As an individual remains unemployed, his life satisfaction decreases a little more (until a new equilibrium is reached) though most of the impact is contemporaneous. As a further illiustration, consider the long-term unemployed: for them, most of the overall life satisfaction penalty is due to being currently unemployed, with prior years of unemployment adding a little to this penalty. The contemporaneous experience is what is important.

Figure 2: illustration of life satisfaction effect of unemployment


Both of these results help illustrate the highly contemporaneous nature of life satisfaction. That most of the impact of well-being is contemporaneous may explain some previously found results in the well-being literature. Steiner et al. (2013) investigate the individual life satisfaction or well-being impact of a city being the European Capital of Culture. They find, on average, a significant negative impact in the year a city is the European Capital Culture, but no impact in the years before or afterwards.[21] The results here regarding the dynamics of happiness suggest that an event like this is unlikely to have a substantial effect (if any) on the day to day lives of individuals in any other year than the year of the associated celebrations, life satisfaction being a largely contemporaneous phenomenon. Similarly, Kavetsos and Szymanski (2010) find that hosting the FIFA World Cup or the Olympics increases life satisfaction only in the year of the event and has no long term effects on life satisfaction. In the language of time-series econometrics, such events are crash dummy variables occurring once, in comparison with shift dummies which reflect a more permanent change (for example, being married). Figure 1 illustrates this, with an assumption that the events (Olympics, marriage) happens in year four.

[21] The authors suggest that this negative effect may reflect dissatisfaction with associated high levels of public expenditure, transport disruptions, general overcrowding or an increase in housing prices.

Figure 3: illustration of crash and shift dummy variables

| | Year | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 |
| Crash dummy variable | 0 | 0 | 0 | 1 | 0 | 0 |
| Shift dummy variable | 0 | 0 | 0 | 1 | 1 | 1 |

Given the (largely) contemporaneous nature of life satisfaction, any event or situation that has the property of a crash dummy is highly unlikely to have a well-being effect in the years before or subsequent to the year of the event (i.e. in the years when the dummy variables takes the value 0). For shift dummies, there may well be a contemporaneous effect in the years after event takes place (when the dummy variable takes the value of 1). Marriage, for example, may contribute to well-being in the years after marriage and this is what the adaptation literature tries to determine.

Such a result – life satisfaction scores largely reflecting contemporaneous events with minor influence from the past - offers a reframing of the adaptation question. Thinking about adaptation as getting used to an event from the past (e.g. marriage) can obscure what seems to be occurring with well-being. Well-being appears to (largely) reflect what is going on now rather than what happened in the past: being married mattering more than the act of marriage; being unemployed mattering more than entering unemployment. The question researchers should perhaps ask instead is: does this event, or situation, have a contemporaneous effect on life satisfaction? In other words, is an individual's happiness affected by this situation or status *now*? For an event to have a legacy or long term impact on an individual's life satisfaction it seems likely that it must have a profound effect on the individual's day to day life sometime after the event is entered into. Dynamic panel analyses, with their contemporaneous coefficients, can discover this.

## 5. Conclusion

Subjective well-being, as assessed by life satisfaction data, is heavily influenced by contemporaneous circumstances and events. Any direct influence of the measured past is somewhat minimal.[22] This could not be found via standard fixed effects (or other non-dynamic) methods. This investigation into the concept of life satisfaction and its dynamics, which has taken advantage of theoretical advances coupled with increases in our collective understanding of using General Method of Moments procedures to estimate dynamic panel models. This, along with the subsequent technical and computational advances, makes running such models possible and somewhat straightforward.

Roodman (2009) warns that such apparent simplicity can mean that such models are estimated incorrectly and without full diagnostic testing. Indeed, as this paper has shown, studies in the well-being area sometimes misunderstand the diagnostics or fail to report them (or discuss them) sufficiently. Future research using these models needs to remedy this, especially because the choices that a researcher makes regarding instrumentation can have a substantial impact on the subsequent results, as well as on the subsequent diagnostic test outcomes, and these need to be explained. Here the diagnostics did not always fully support the estimations, though the coefficients obtained appear very robust which offers some confidence regarding the estimations. Future work may well encounter similar concerns regarding the diagnostic test results, and these results should be shown and a note of caution attached to them. Testing the robustness of the obtained coefficients is important and there are many ways to do this.

The analysis and results of this study both support and extend recent research. The central finding of a small, positive coefficient on the lag of life satisfaction (which represents the history of the model) means that most of what makes up current life satisfaction scores reflects contemporaneous concerns and situations. For example, being married contributes, on average, to well-being

[22] This does not, however, rule out indirect influences where individuals make contemporaneous decisions which may partly reflect their past.

sometime after getting married; most of the contribution to life satisfaction comes from being currently married (as opposed to the event of marriage or previous years of marriage). This is likely to be in contrast to events that are one-offs. The analysis here suggests reasons for the previous findings that any feel good factor from events like the Olympics do not have a legacy in terms of individual well-being.

The consistent, positive, yet small influence of the past on current life satisfaction could not have been found using the 'workhorse' static models. An initial reason for a dynamic panel analysis was the possibility that many static models are misspecified. They may well suffer from serial correlation, indicating missing dynamics. One way of taking advantage of this finding is to employ a dynamic panel model. Indeed such a model may be important to obtain more accurate associations between the right-hand side variables and well-being. Given that life satisfaction appears to be a largely contemporaneous phenomenon, models (like system GMM) that can provide contemporaneous coefficients are very useful. A further advantage of such models is the ability to straightforwardly instrument potentially endogenous variables, better accounting for the relationship between well-being and some commonly used explanatory variables and enabling the estimation of coefficients for variables previously deemed impossible to assess.

Studies in the well-being area have started to employ dynamic panel methods, but often do not adequately consider the diagnostics nor appear to fully understand how such models need to be interpreted. Such methods are more complex than the standard fixed effects and this additional complexity needs to be better understood. It is not enough just to include a lagged dependent variable in standard well-being estimations without considering the additional complexity involved. Dynamic analyses of well-being are at a nascent stage but have many advantages (and challenges) and offer an interesting path for future well-being research.

References


Arellano, M. and Bond. S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, *The Review of Economic Studies* 58, pp. 277 – 297.

Arellano, M, and Bover, O. (1995) Another look at the instrumental variable estimation of error -components models, *Journal of Econometrics* 68, pp. 29–51.

Baum, C. F., Schaffer, M. E. and Stillman, S. (2003) Instrumental variables and GMM: Estimation and testing. *Stata Journal* 3: 1–31.

Bazzi, S. and Clemens, M. (2009) Blunt Instruments: On Establishing the Causes of Economic Growth, *Working Papers* 171 *Center for Global Development*.

Bond, S. R. (2002) Dynamic Panel Models: A Guide to Micro Data Methods and Practice. *Institute for Fiscal Studies / Department of Economics, UCL, CEMMAP (Centre for Microdata Methods and Practice)Working Paper* No.CWPO9/02.

Bottan, N. L. and Perez-Truglia, R.N. (2011) Deconstructing the Hedonic Treadmill: Is Happiness Autoregressive? *Journal of Socio-Economics* 40, 3: pp. 224-236.

Clemens, M., Radelet, S., and Bhavnani, R. (2004) Counting Chickens When They Hatch: The Short-term Effect of Aid on Growth, *Working Papers* 44*, Center for Global Development.*

D'Addio, A. C., Chapple, S., Hoherz, A., & Van Landeghem, B. (2013). Using a quasi-natural experiment to identify the effects of birth-related leave policies on subjective well-being in Europe. *OECD Journal: Economic Studies*, 1, pp.1-34.

Della Giusta, M.,Jewell, S., and Kambhampati, U. (2010)Anything to Keep You Happy?, *Economics & Management Discussion Papers* 01, Henley Business School, Reading University.

Dutta, N., Leeson, P. T. and Williamson, C. R. (2013) The Amplification Effect: Foreign Aid's Impact on Political Institutions, *Kyklos* 66: 2, pp. 208-228.

Drukker, D. M. (2003) Testing for serial correlation in linear panel-data models. *Stata Journal* 3, pp. 168-177.

Flannery, M. J., & Hankins, K. W. (2013) Estimating dynamic panel models in corporate finance *Journal of Corporate Finance*, 19, pp.1-19.

Godfrey,L. G., 1991. Misspecification Tests in Econometrics. Cambridge Books, Cambridge University Press.

Greene, W. (2002) *Econometric Analysis,* Fifth Edition. Upper Saddle River, New Jersey: Prentice Hall.

Greene, W. (2008) *Econometrics Analysis,* Sixth Edition. Upper Saddle River, New Jersey: Prentice Hall.

Hall, A. R. (2004) *Generalized Method of Moments*. Oxford: Oxford University Press.

Hansen, L. P. (1982) Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society* pp.1029-1054.

Holtz-Eakin, D., Newey, W., & Rosen, H. S. (1988) Estimating vector autoregressions with panel data. *Econometrica: Journal of the Econometric Society*, pp. 1371-1395.

Kavetsos, G. and Szymanski, S. (2010) National well-being and international sports events. *Journal of Economic Psychology*, vol. 31:2, pp.158-171.

Lee, S. H., Levendis, J., & Gutierrez, L. (2012) Telecommunications and economic growth: an empirical analysis of sub-Saharan Africa. *Applied Economics*, 44:4, pp.461-469.

Nickel, S. J. (1981) Biases in Dynamic Models with Fixed Effects, *Econometrica*, 49 , pp. 1417-1426

Piper, A. T. (2012) Dynamic Analysis and the Economics of Happiness: Rationale, Results and Rules, *MPRA Paper* 43248, University Library of Munich, Germany.

Piper, A. T. (2013) A Note on Modelling Dynamics in Happiness Estimations, *MPRA Paper* 49364, University Library of Munich, Germany.

Piper, A. T. (2014) Zukunfstsangst! Fear of (and hope for) the future and its impact on life satisfaction, *MPRA Paper* 59557, University Library of Munich, Germany.

Powdthavee, N. (2009) I can't smile without you: Spousal correlation in life satisfaction, *Journal of Economic Psychology*, vol. 30: 4, pp. 675-689.

Pugh, G., Mangan, J., Blackburn, V., & Radicic, D. (2014) School expenditure and school performance: evidence from New South Wales schools using a dynamic panel analysis. *British Educational Research Journal*. doi: 10.1002/berj.3146

Qari, S. (2014). Marriage, adaptation and happiness: Are there long-lasting gains to marriage?. *Journal of Behavioral and Experimental Economics*, 50, pp.29-39.

Roodman, D. (2006) How to do xtabond2: An Introduction to “Difference” and “System” GMM in STATA, *Center for Global Development, Working Paper* No. 103.

Roodman, D. (2007) A Note on the Theme of Too Many Instruments, *Center for Global Development, Working Paper* No.125.

Roodman, D. (2009) How to do xtabond2: An Introduction to Difference and System GMM in Stata, *The Stata Journal*, 9(1) pp.86-136.

Steiner, L., Frey, B.S. and Hotz, S. (2013) European capitals of culture and life satisfaction, *ECON - Working Papers* 117, Department of Economics - University of Zurich.

Stutzer, A. and Frey, B.S. (2006). Does Marriage Make People Happy, Or Do Happy People Get Married? *Journal of Socio-Economics*, vol. 35, pp. 326-347.

Verbeek, M. (2000) *A Guide to Modern Econometrics.* Chichester: Wiley.

Windmeijer, F. (2005) A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics* 126: 25-51.

Wooldridge, J. M. (2002) *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MA: MIT Press.

Wunder, C. (2012) Does subjective well-being dynamically adjust to circumstances?, *Economics Letters*, vol. 117: 3, pp. 750-752.

Table 1 life satisfaction of British people, assessed via GMM dynamic panel analysis.

| VARIABLES | Males<br>All | Females<br>All | Females≠<br>Age 15-35 | FemalesΨ<br>Age 15-50 |
|---|---|---|---|---|
| Lagged Life Satisfaction | 0.09*** | 0.09*** | 0.10*** | 0.09*** |
| | (0.014) | (0.012) | (0.019) | (0.013) |
| Log wage | 0.00*** | -0.00 | -0.00 | -0.01* |
| | (0.000) | (0.000) | (0.001) | (0.003) |
| Self-employed | 0.04* | 0.04 | 0.02 | 0.05 |
| | (0.023) | (0.031) | (0.058) | (0.036) |
| Unemployed | -0.43*** | -0.30*** | -0.33*** | -0.34*** |
| | (0.039) | (0.043) | (0.061) | (0.050) |
| Retired | 0.01 | 0.12** | | -0.31 |
| | (0.058) | (0.047) | | (0.204) |
| LT Sick or Disabled | -0.75*** | -0.57*** | -0.56*** | -0.55*** |
| | (0.063) | (0.052) | (0.108) | (0.087) |
| FT Student | 0.01 | 0.06* | 0.06* | 0.02 |
| | (0.036) | (0.033) | (0.034) | (0.035) |
| Family Carer | -0.38*** | -0.15*** | -0.20*** | -0.19*** |
| | (0.097) | (0.025) | (0.036) | (0.032) |
| Other Labour Force Status | -0.31*** | 0.11*** | 0.14*** | 0.12*** |
| | (0.091) | (0.039) | (0.045) | (0.039) |
| Married | 0.45*** | 0.47*** | 0.43*** | 0.47*** |
| | (0.096) | (0.100) | (0.081) | (0.095) |
| Separated | -0.10 | -0.17 | -0.27 | -0.08 |
| | (0.200) | (0.176) | (0.283) | (0.175) |
| Divorced | 0.19 | -0.06 | -0.08 | -0.04 |
| | (0.161) | (0.145) | (0.157) | (0.138) |
| Widowed | 0.17 | -0.24 | -0.13 | 0.19 |
| | (0.328) | (0.252) | (0.573) | (0.237) |
| Education: High | -0.12*** | 0.01 | 0.11** | 0.06* |
| | (0.028) | (0.028) | (0.045) | (0.035) |
| Education: Medium | -0.10*** | -0.02 | 0.08* | 0.03 |
| | (0.029) | (0.028) | (0.045) | (0.033) |
| Health: Excellent | 0.62*** | 0.71*** | 0.70*** | 0.90*** |
| | (0.022) | (0.020) | (0.030) | (0.141) |
| Health: Good | 0.41*** | 0.45*** | 0.43*** | 0.58*** |
| | (0.019) | (0.017) | (0.026) | (0.131) |
| Age: 21 – 30 years old | -0.29*** | -0.12*** | -0.09** | -0.09** |
| | (0.037) | (0.041) | (0.037) | (0.041) |
| Age: 31 – 40 years old | -0.53*** | -0.29*** | -0.20*** | -0.26*** |
| | (0.071) | (0.078) | (0.059) | (0.076) |
| Age: 41 – 50 years old | -0.61*** | -0.39*** | | -0.36*** |
| | (0.085) | (0.092) | | (0.089) |
| Age: 51 – 60 years old | -0.44*** | -0.23** | | |
| | (0.090) | (0.096) | | |
| Wave Dummies | Yes | Yes | Yes | Yes |
| Region Dummies | Yes | Yes | Yes | Yes |
| Constant | 4.53*** | 4.30*** | 4.22*** | 4.17*** |
| | (0.086) | (0.077) | (0.112) | (0.115) |
| Number of observations | 34801 | 41644 | 17064 | 32858 |
| Number of instruments | 274 | 278 | 255 | 418 |
| Number of Individuals | 7820 | 8963 | 4765 | 7547 |

| | | | | |
|---|---|---|---|---|
| AR (2) | 0.147 | 0.016 | 0.842 | 0.364 |
| Hansen's *J* test | 0.935 | 0.053 | 0.551 | 0.447 |
| Diff-in-Hansen for Levels | 0.552 | 0.456 | 0.917 | 0.770 |
| Diff-in-Hansen (lag depvar) | 0.382 | 0.005 | 0.288 | 0.134 |

Note: data from individuals in the BHPS, 1996-2007, aged 15 to 60. Standard errors in parentheses*** p<0.01, ** p<0.05, * p<0.1. Missing categories: employed, single, low education, fair to very poor health, 16 – 20 years old. Key ≠: the 10 females aged 35 or lower in the data set are included in the other labour force status category; Ψ here, health and real income are treated as endogenous as well as marital status.

## APPENDIX 1

The coefficient for lagged life satisfaction in these dynamic estimations is itself interesting and, as Greene informs us (see the quote that introduces the results section), this coefficient represents the 'entire history of the model' i.e. the history of the process that generates current levels of happiness. A little algebra expanding the lagged dependent variable demonstrates this. In equation (1) $LS_{it}$ is the life satisfaction of individual i in year t, $\hat{\beta}x_{it}$ is an independent variable and $\epsilon_{it}$ is the usual error term. Starting with our simplified specification in equation (1), we repeatedly substitute for the lagged dependent variable.

$$LS_{it} = \hat{\alpha}LS_{it-1} + \hat{\beta}x_{it} + \epsilon_{it} \tag{1}$$

Substitute for $LS_{it-1}$ in (1):

$$LS_{it-1} = \hat{\alpha}LS_{it-2} + \hat{\beta}x_{it-1} + \epsilon_{it-1} \tag{2}$$

Substitute (2) into (1)

$$LS_{it} = \hat{\alpha}(\hat{\alpha}LS_{it-2} + \hat{\beta}x_{it-1} + \epsilon_{it-1}) + \hat{\beta}x_{it} + \epsilon_{it} \tag{3}$$

Substitute for $LS_{it-2}$ in (3):

$$LS_{it-2} = \hat{\alpha}LS_{it-3} + \hat{\beta}x_{it-2} + \epsilon_{it-2} \tag{4}$$

Substitute (4) into (3)

$$LS_{it} = \hat{\alpha}(\hat{\alpha}\left[\hat{\alpha}LS_{it-3} + \hat{\beta}x_{it-2} + \epsilon_{it-2}\right] + \hat{\beta}x_{it-1} + \epsilon_{it-1}) + \hat{\beta}x_{it} + \epsilon_{it} \tag{5}$$

Gather terms

$$\mathrm{LS_{it}} = \hat{\alpha}^3\mathrm{LS_{it-3}} + \hat{\alpha}^2\hat{\beta}\mathrm{x_{it-2}} + \hat{\alpha}\hat{\beta}\mathrm{x_{it-1}} + \hat{\beta}\mathrm{x_{it}} + \hat{\alpha}^2\epsilon_{\mathrm{it-2}} + \hat{\alpha}\epsilon_{\mathrm{it-1}} + \epsilon_{\mathrm{it}} \tag{5'}$$

Going back further than four lags introduces more past values and more idiosyncratic error terms too. By repeated substitution, it can be demonstrated that through the lagged dependent variable dynamic specifications contain the entire history of the independent variable(s). Clearly this is not just a fixed effect (as sometimes assumed).

APPENDIX 2

Fixed effects life satisfaction regressions for British individuals aged 15-60

| VARIABLES | Both genders Life Satisfaction | Males Life Satisfaction | Females Life Satisfaction |
|---|---|---|---|
| Real Annual Income ('000s) | 0.00* | 0.00** | -0.00 |
| | (0.000) | (0.000) | (0.000) |
| Self-employed | 0.00 | -0.01 | 0.00 |
| | (0.019) | (0.023) | (0.031) |
| Unemployed | -0.33*** | -0.41*** | -0.26*** |
| | (0.018) | (0.025) | (0.027) |
| Retired | 0.02 | -0.01 | 0.04 |
| | (0.028) | (0.044) | (0.036) |
| LT Sick or Disabled | -0.52*** | -0.70*** | -0.41*** |
| | (0.025) | (0.038) | (0.032) |
| FT Student | 0.03 | -0.01 | 0.05** |
| | (0.019) | (0.029) | (0.026) |
| Family Carer | -0.12*** | -0.20*** | -0.10*** |
| | (0.017) | (0.069) | (0.019) |
| Other Labour Force Status | 0.08*** | -0.12** | 0.14*** |
| | (0.027) | (0.055) | (0.032) |
| Married | 0.08*** | 0.07*** | 0.07*** |
| | (0.019) | (0.027) | (0.027) |
| Separated | -0.10*** | -0.14*** | -0.08** |
| | (0.031) | (0.047) | (0.042) |
| Divorced | 0.06** | 0.06 | 0.06 |
| | (0.028) | (0.041) | (0.038) |
| Widowed | -0.17*** | -0.13 | -0.19*** |
| | (0.060) | (0.114) | (0.073) |
| Education: High | 0.05* | 0.06 | 0.03 |
| | (0.026) | (0.038) | (0.037) |
| Education: Medium | 0.04* | 0.06 | 0.03 |
| | (0.027) | (0.039) | (0.037) |
| Health: Excellent | 0.44*** | 0.43*** | 0.46*** |
| | (0.012) | (0.017) | (0.016) |
| Health: Good | 0.30*** | 0.30*** | 0.30*** |
| | (0.009) | (0.013) | (0.012) |
| Age: 21-30 | -0.10*** | -0.18*** | -0.04 |
| | (0.019) | (0.028) | (0.027) |
| Age: 31-40 | -0.12*** | -0.20*** | -0.05 |
| | (0.027) | (0.039) | (0.038) |
| Age: 41-50 | -0.16*** | -0.23*** | -0.10** |
| | (0.034) | (0.049) | (0.048) |
| Age: 51-60 | -0.11*** | -0.15** | -0.08 |
| | (0.042) | (0.060) | (0.058) |
| Wave Dummies | Yes | Yes | Yes |
| Region Dummies | Yes | Yes | Yes |
| Constant | 4.96*** | 4.94*** | 4.98*** |
| | (0.058) | (0.081) | (0.083) |
| Observations | 107,858 | 49,534 | 58,324 |
| R-squared | 0.033 | 0.040 | 0.030 |
| Number of Individuals | 21,004 | 9,905 | 11,099 |

Note: data from individuals in the BHPS, 1996-2007; standard errors in parentheses; significance levels: *** p<0.01; ** p<0.05; * p<0.1; baseline categories: employed, never married, low education, health self-reported as fair or worse, age range 16-20.