Making inference of British household's happiness efficiency: a Bayesian latent model.

Emmanuel C. Mamatzakis[a] and Mike G. Tsionas[b]

January 2021

In this paper, we propose a novel approach whereby happiness for British households is identified within a latent model frontier analysis using longitudinal data. By doing so we overcome issues related to the measurement of happiness. To estimate happiness frontier and thereby happiness efficiency, we employ a Bayesian inference procedure organized around Sequential Monte Carlo (SMC) particle filtering techniques. In addition, we propose to consider individual-specific characteristics by estimating happiness efficiency models with individual-specific thresholds to happiness. This is the first study that treats happiness as a latent variable and departs from restrictions that happiness efficiency would be time invariant. Our results show that happiness efficiency is related to the welfare loss associated with potentially misusing the resources that British individuals have at their disposal. Key to happiness is to have certain personality traits, such as being agreeable and extravert as they assist efforts to enhance happiness efficiency. On the other hand, being neurotic impairs happiness efficiency.

Keywords: Behavioural OR, happiness, latent modelling, Bayesian inference.

JEL Codes: C11, C13, C35, I3, D1.

Acknowledgements: The authors wish to thank David Peel for useful discussions on a previous version. We would like to thank the Editor, Professor Robert Graham Dyson, and three Reviewers for all their constructive comments. All remaining errors are ours.

[a]Department of Management, Birkbeck College, University of London, London WC1E 7HX, UK. E-mail: e.mamatzakis@bbk.ac.uk. [b] Lancaster University Management School, Bailrigg, Lancaster, LA1 4YX, UK. E-mail: m.tsionas@lancaster.ac.uk.

1. Introduction

To trace the origins of what constitutes happiness from its content point of view one should go back in time thousands of years. Unravelling Ariadne's thread of what constitutes happiness would lead to Aristotelian virtue ethics where '*eudaimonia*' is its centre of gravity. The Aristotle's concept of 'eudemonia' provides the foundations of modern theorisations of happiness. According to Aristotle virtue, health, wealth, and beauty all contribute to eudemonia and thereby happiness. As often the case, Aristotle's philosophy did raise criticism at the time as the Stoics philosophers draw attention only to virtue as the determinant of eudaimonia and thus happiness. The debate of defining happiness and searching for the factors that would contribute to happiness remains vivid to this date. In this study we follow from Aristotle's eudemonia to derive happiness and thereby quantifying happiness efficiency. In line with Aristotle's eudaimonia, an individuals' endowment of socio-economic resources, but also psychological resources, would determine her/his potential to achieve an optimal, steady-state level of happiness.

Coming to our era, happiness studies have gained considerable attention since early in 2000 by both academics (see Kahneman et al. 2004, Stiglitz et al. 2009; Binder and Broekel, 2012; Cordero et al. 2017) and policy makers alike (see Debnath et al. 2017; Frey and Stutzer, 2012; Conzo et al. 2017; United Nations World Happiness Report while the OECD also publishes its own happiness report) as standard macro-economic and micro-economic measures and models have been criticised for failing to reveal the true individual's well-being. Alas, defining the content of happiness is not by any means an easy task mainly due to the fact that is subjective and to a large extent unobservable. The precise definition of happiness has been widely debated since the days of Aristotle's eudemonia. One of the seminal contributions in the literature Kahneman (2003) and Kahneman and Krueger (2006) highlight the challenges of defining happiness. Kahneman and Krueger (2006) reason that understanding happiness one should differentiate between experiencing happiness, which is related to how we feel while we live, and satisfaction with our life, which is how we feel when we think about our life. The authors suggest that happiness could be identified by two factors that derive from the above differentiation: one factor relates to the current experience of the feeling of an emotion such as pleasure or joy, which is not easily measurable and it might be spontaneous; and the second factor relates to the appraisal of life satisfaction, such as of the quality of life over a period of time. Kahneman (2003) and Kahneman and Krueger (2006) argue that the appraisal of life satisfaction is more important to people.

In this paper, we build on the content of happiness from an economic point of view as provided by Kahneman and Krueger (2006). To this end, we employ data from the British Household Panel Survey that provides information about an individual's current experience of her life (such as emotions, moods, and feelings that would classify under the first factor of Kahneman and Krueger) and her appraisal of life satisfaction (that classifies as the second factor of Kahneman and Krueger). Modelling happiness does not, also, come without considerable challenges. From a content point of view Rayo and Becker (2007) provide the theory of an individual's happiness function. They reason that the individual is mainly concerned not with her absolute level of happiness, but rather with the difference between her happiness and a benchmark of happiness. This reference on the existence of a benchmark is key as Rayo and Becker (2007) show that the happiness function provides a decision-making device that allows an individual to rank alternative actions, regarding, for example, what inputs and how much of them the individual will select in her happiness function. This process of ranking individual's actions within the framework of happiness function is of importance as it is at the core of pursuing happiness efficiency first theorized by Rayo and Becker (2007).[1] In some detail, Rayo and Becker (2007) show that an individual would always aim at maximizing her happiness given her characteristics and resource endowments, and to do so it is necessary to reach for higher values of happiness efficiency that are strictly preferred. The maximization of happiness underpins most of the empirical models in the literature (Cordero et al. 2017; Binder and Broekel, 2012; Graham and Oswald 2010). However, Rayo and Becker (2007) predict that not all individuals are equally efficient in utilizing their resources and have certain characteristics that would lead to the maximum happiness. Rayo and Becker (2007) suggest, therefore, that there could be a degree of happiness inefficiency which would explain not achieving the maximum happiness. The existing literature has paid little attention on estimating the extent to which happiness inefficiency is prevalent at an individual level (Cordero et al. 2017), while it has also been largely silent about explaining the causes of happiness inefficiency.

We build on the content of happiness efficiency of Rayo and Becker (2007) and reason that it is of importance to provide empirical evidence and thereby reveal individual's happiness and happiness efficiency. Rayo and Becker (2007) also discuss the underlying *'inputs'* of the happiness function, that is the individual's characteristics and resource endowments. On the determinants of the happiness function, there are also challenges, some go back to Aristotle's notion of '*eudemonia*'. Lucas and

[1] Rayo and Becker (2007) show that happiness and the happiness efficiency should be seen in the light of the principal agent problem whereby the principal designs the happiness function so as the agent, which is the individual, maximises her happiness. Rayo and Becker (2007) provide proof that maximizing the happiness efficiency would lead to the maximum happiness. This theory is key to our modelling as it provides its theoretical content whereby higher values of happiness efficiency are strictly preferred by the principal and the agent.

Diener (2015) provides a comprehensive review of the complexities involved regarding the plethora of potential determinants of happiness. In previous studies happiness is a function of social-economic variables such as: education, income, employment, health, marital status traits (Clark et al., 2008; Diener 2009; Diener, et al. 2018; Lucas and Diener 2015; Graham and Oswald 2010; Binder and Broekel, 2012); but also variables such as geographic location and personality (Anand, et al. 2011; Cordero et al. 2017).

In this study, we consider these challenges and opt to employ both time invariant and time variant determinants of happiness. To this end, happiness would be a function of time invariant characteristics of individuals such as gender (Dolan and Kahnemann, 2008; Kahneman, 2003), but also on time variant characteristics such as income (Clark et al., 2008), health (Dolan and Kahnemann, 2008), employment (Winkelmann, 1998). We shall also consider important life events such as the marital status in line with the adaptation doctrine (Lucas, 2007), but also personality traits given their importance (Lucas and Diener 2015; Gosling, et al. 2003; Diener, et al. 1999). Lucas and Diener (2015) argue that, typically, the correlation between happiness and, for example, income is in most empirical research half of the one between happiness and personality traits like extraversion and neuroticism.

When it comes to the estimation of the happiness function there are also challenges: some studies employ parametric methods (Diener 2009; Diener, et al. 2018; Rayo and Becker 2007; Graham and Oswald 2010; and Binder and Broekel, 2012; Anand, et al. 2011), while other studies opt for non-parametric methods (see Despotis, 2005; Cordero et al. 2017; Barberio et al. 2015; Mizobuchi 2017; Debnath and Shankar 2014; Tsurumi and Manage 2017).[2] In addition, there is another thread in the literature that predicts persistence in the happiness function. Di Tella et al., (2010) and Bottan and Perez-Truglia, (2010) argue that happiness is subject to such persistence and it should be modelled taking into account underlying dynamics. Di Tella et al., (2010) show that persistence in happiness could be explained but what they called '*reference dependent preferences*' impended in the serial correlation of the idiosyncratic error term.[3] In this paper, we propose a novel Bayesian stochastic

[2] It is worth noting that studies of Despotis (2005) and Barberio-Mariano (2015) focus on the concept of quality of life which does not involve subjective aspects.

[3] Extending on Rayo and Becker's model, Di Tella et al., (2010) provide a theory of the importance of reference-dependent preferences. Their theory helps to resolve the Easterlin paradox. Easterlin was the first researcher to study happiness data and observed that happiness is positively correlated with income across individuals and across countries at a given period of time, but this is not true over time. The reference-dependent preference theory reasons that the current happiness may depend directly on past happiness (see also Bottan and Perez-Truglia, 2010). In this paper, we consider the reference dependent preference.

frontier happiness function that addresses concerns regarding the estimation of happiness function in previous literature.

White et al. (2016), White (2018) and White (2016) argue in favor of the significance of social operational research modelling. In line with White et al. (2016) and White (2016) and given the content of Rayo and Becker (2007) happiness efficiency theory, we propose that a way to advance social operational research modeling is to view the measurement of happiness as one that is related to a latent variable modelling. To this end, we propose that accurately measuring happiness should be in the first instance about modeling it as an unobserved variable. By treating happiness as unobservable we tackle the challenges that relate to the definition of happiness. This is a novel contribution to the existed literature. Our modelling provides a Bayesian generalized linear latent model while considering unobserved heterogeneity across individuals through introducing latent individual effects (see Anand, et al. 2011 for linear latent and mixed models).[4] We show that our latent modelling nests both parametric or non-parametric estimation methods and thereby addresses criticism raised about the estimation method. This modelling is also addressing endogeneity and multicollinearity concerns. We proceed by taking the latent modelling some steps further by incorporating a stochastic happiness efficiency frontier so as to provide evidence of Rayo and Becker (2007) prediction about happiness efficiency. The concept of happiness efficiency frontier is key as any deviations from the frontier would imply that individuals deviate from their maximum happiness that could achieve should they optimally use their available resources. In addition, and following Di Tella et al., (2010) we model both persistence and threshold effects in happiness and happiness efficiency. Furthermore, as heterogeneity across individuals is key our model provides happiness efficiency across individuals and over time.

In terms of the estimation of happiness function and happiness efficiency, we employ Bayesian inference procedures organized around Sequential Monte Carlo (SMC) particle-filtering techniques so as to explore the posterior distribution of happiness efficiency, whilst we also allow for key individual's specific characteristics, whether time varying like income, or time invariant like gender, to impact upon happiness efficiency. Our modeling permits to do so in a single stage that enhances statistical significance and statistical efficiency. Compared to previous methods of estimation in the

[4] This analysis follows earlier research by Krishnakumar, (2007), Anand et al., (2009), Durante, et al. (2017) and Sewell and Chen (2015).

literature the Bayes approach developed here is better suited for finite survey samples and data which embodies substantial individual heterogeneity. In our empirical application, we use all available data from the British Household Panel Survey. The survey is a longitudinal one of British households and since the survey commenced there have been 18 waves concerning 5050 British households. We test various models, including static and dynamic ones. We also propose a model that captures heterogeneity across individuals, that departs from standard random effects models, given the cross-sectional heterogeneity of our sample. Results show that the typical British individual could further improve her happiness efficiency. An interesting finding of our new results is that a large part of the observed happiness inefficiency can be explained by personality traits. Individuals with certain personality traits might be more efficient than others in employing their resources so as to move towards their happiness frontier. Personality traits could also impact upon the resilience of individuals when they face an adversity while these traits could also impact upon reaching their happiness frontier.

The remainder of the paper proceeds as follows. The next section presents our methodology, the latent variable modeling. Section 3 presents the data set and the results. Section 4 provides some conclusions.

2. Methodology

2.1 The latent happiness model

We build on the content of Rayo and Becker (2007) who theorize that the individual's objective is to maximize her happiness. Suppose y_{it} denotes the ordinal response of individual i at time t ($i = 1, \cdots, n; t = 1, \cdots, T$), where specifically $y_{it} = c$ for $c = 1, \cdots, C$. Moreover, in terms of our sample, y_{it} comes from the individuals' responses to the question of '*How dissatisfied or satisfied are you with your life overall?*' in the British Household Panel Survey data.[5] If y_{it} takes the value of 1 then it means that the underlying individual is completely unsatisfied, where 7 would indicate the maximum level of life satisfaction. Given the challenges of defining happiness (Dolan and Kahnemann, 2008; Kahneman, 2003), we employ $y_{it} = c$ to model happiness, H_{it}, as a latent variable for the first time, which is unobservable though it exists:

[3]Most papers in the literature (see Binder and Broekel, 2012) opt for the same question to define happiness. This is a direct way of measuring happiness as it comes from the responses in the BHPS data survey. Similar survey questions are also used for other countries, such as in Marcenaro-Gutierrez et al. (2010) that study the level of satisfaction of workers in Spain.

$$y_{it} = c \iff H_{it} \in (\gamma_{c-1}, \gamma_c), \forall\ t = 1, \cdots, T, \forall\ i = 0, 1, \cdots, n, \quad (1)$$

where γ_c denotes a cutoff point. We have: $-\infty = \gamma_0 < \gamma_1 = 0 < \gamma_2 < \cdots < \gamma_C = \infty$, where the condition $\gamma_1 = 0$ is needed for identification (Albert and Chib, 1993, p. 673) and Johnson and Albert (1999, p.131).

It is worth noting that since our data is ordinal, we need to have an appropriate statistical model to take account of this fact. So, we assume that when, for example, $y_{it} = 1$ then the latent happiness indicator is in the interval $(\gamma_0 = -\infty, \gamma_1 = 0)$, when $y_{it} = 2$ then the latent happiness indicator is in the interval $(\gamma_1 = 0, \gamma_2)$, etc., where γ_2, γ_3 are parameters to be estimated. Therefore, the "*cutoff points*" determine sub-intervals of latent happiness which correspond one-to-one to observed values of the response.

We treat happiness as a latent variable that follows the process:

$$H_{it} = \rho_i H_{i,t-1} + \beta_i' x_{it} + \delta_i' w_i + u_{it}, \forall\ t = 1, \cdots, T, i = 1, \cdots, n, \quad (2)$$

and

$$H_{i0} = \beta_{0i}' x_{i0} + \delta_{0i}' w_i + u_{i0}, \forall\ i = 1, \cdots, n, \quad (3)$$

where x_{it} is a $k \times 1$ vector of covariates, w_i is an $m \times 1$ vector of time-invariant characteristics and β_i, δ are conformable parameter vectors. Happiness efficiency is derived from equation (3) decomposition of u_{it} to random error and efficiency term following Cornwell, et al. (1990) estimator. Note that ρ is the coefficient of the lagged happiness, capturing possibly persistence of happiness over time thus taking on board Di Tella et al., (2010) reference-dependent preferences. Note that testing for persistence is key in our modelling and has been rather overlooked in the literature. However, if persistence is ignored this would lead to incorrect estimates and inferences of happiness.

Moreover, x_{it} is a $k \times 1$ vector of primary inputs (Rayo and Becker 2007 refer to these as resource endowments) into the happiness function in line with Binder and Broekel (2012) and Clark et al. (2008) while w_i is a $m \times 1$ vector of an individual's specific characteristics (see also Cheng et al. 2015; Oswald and Powdthavee, 2008; Tiefenbach and Kohlbacher, 2015). For example, x_{it} includes variables such as the annual household income, health and education (see Clark et al. 2008; Cheng et al. 2015), whereas w_i include variables such as gender, age (see Oswald and Powdthavee, 2008; Tiefenbach and Kohlbacher, 2015). Given that we would focus on British households and individuals we consider where the geographic location of a household would impact upon happiness by including regional dummy variables.[6]

Conditionally on the regressors, x_{it} and w_i, parameters, we assume:

$$u_{it} \mid x_{it}, w_i, \beta, \delta \sim N(0,1), \forall t = 0,1,\cdots,T, i = 1,\cdots,n. \quad (4)$$

Relative to previous work, we allow, in the first instance, for random coefficients to capture heterogeneity:

$$\lambda_i = [\rho_i, \beta'_i, \delta_i]' \sim N(\bar{\lambda}, \Sigma_\lambda), \forall i = 1,\cdots,n; \quad (5)$$

$$\lambda_0 = [\rho_0, \beta'_0, \delta_0]' \sim N(\bar{\lambda}_0, \Sigma_{\lambda 0}), \forall i = 1,\cdots,n. \quad (6)$$

Markov Chain Monte Carlo (MCMC thereafter) techniques for simulating the posterior distribution have been proposed by Albert and Chib (1993, 2001), and Chen and Dey (2000), see also Jeliazkov et al. (2009). The main problem is to draw the parameters $\gamma = [\gamma_2,\cdots,\gamma_{C-1}]'$. Albert and Chib (2001) propose a reparameterization of the form $z_c = \log(\gamma_c - \gamma_{c-1}), \forall c = 2,\cdots,C-1$ so the vector $z = [z_2,\cdots,z_{C-1}]'$ is unrestricted (see also Hasegawa, 2009).

2.2 Individual-specific characteristics

The assumption that the underlying individual's characteristics are the same across all is not realistic as there is heterogeneity. We treat for this heterogeneity across individuals and complement (2) with the following specification:

$$y_{it} = c \Leftrightarrow H_{it} \in (\gamma_{c-1,it}, \gamma_{c,it}), \forall t = 1,\cdots,T, \forall i = 0,1,\cdots,n, \quad (7)$$

where $\gamma_{it} = [\gamma_{2,it},\cdots,\gamma_{C-1,it}]'$ are individual- and time-varying parameters.

We use the Albert and Chib (2001) parametrization,

$$z_{c,it} = \log(\gamma_{c,it} - \gamma_{c-1,it}), \forall c = 2,\cdots,C-1 \quad (8)$$

along with the following assumptions:

$$\zeta_{c,it} = \eta_i + \xi_i \zeta_{c,i,t-1} + \varepsilon_{c,it}, \varepsilon_{c,it} \sim iidN(0, \omega_{ci}^2), \quad (9)$$

where

$$\mu_i = [\eta_i, \xi_i]' \sim N(\bar{\mu}, \Sigma_\mu). \quad (10)$$

[6] Geographic location in previous studies appear to affect happiness; for example, Cordero, et al. (2017) using nonparametric frontier analysis show that happiness varies from region to region.

In this specification, the transformed cutoff points follow an AR(1) process with individual-specific coefficients in μ_i. The variance parameters, ω_{ci}^2, are individual-specific. This modeling perspective allows for considerable flexibility and provides the means of estimating happiness efficiency at the individual level without assuming that the transformation of resources (explanatory variables) to happiness is exactly the same for all individuals.

2.3 Monte Carlo methods

Conditionally on γ, Markov Chain Monte Carlo (MCMC thereafter) for this class of models is quite standard. In our model, however, the cutoff points are both category-, individual- and time-specific. This extension creates problems with standard application of MCMC methods like the Gibbs sampler. We can write (2), the latent happiness equation, in the form:

$$H_{it} = \tilde{\mathbf{x}}'_{it}\lambda_i + u_{it} = \tilde{\mathbf{x}}'_{it}\bar{\lambda} + U_{it}, \qquad (11)$$

where $\tilde{\mathbf{x}}_{it} := [H_{i,t-1}, x'_{it}, w_i]'$, and $U_{it} \sim N(0, 1+\tilde{\mathbf{x}}'_{it}\Sigma_\lambda \tilde{\mathbf{x}}_{it})$. Therefore,

$$\begin{aligned} p(H_{it} \mid H_{i,t-1}, x_{it}, w_i, \theta) &:= p(H_{it} \mid \tilde{\mathbf{x}}_{it}, \theta) \propto \\ &(1+\tilde{\mathbf{x}}'_{it}\Sigma_\lambda \tilde{\mathbf{x}}_{it})^{-1/2} \exp\left\{ -\tfrac{1}{2} \frac{\left(H_{it} - \tilde{\mathbf{x}}'_{it}\bar{\lambda}\right)^2}{\left(1+\tilde{\mathbf{x}}'_{it}\Sigma_\lambda \tilde{\mathbf{x}}_{it}\right)} \right\}. \end{aligned} \qquad (12)$$

In this conditional distribution, the parameters are $\bar{\lambda}$ and the different elements of Σ_λ. Moreover, $\theta \in \Theta \subseteq \Re^K$ denotes the entire vector of parameters in the model. If H_{it} was observed, the product of the above would provide directly the likelihood function along with the different distributional assumptions about $H_{i0}, \forall i = 1, \cdots, n$. Therefore, for individual i, we would have:

$$p(H_i \mid \{\tilde{\mathbf{x}}_{it}\}, \theta) \propto \prod_{t=1}^{T} \left[(1+\tilde{\mathbf{x}}'_{it}\Sigma_\lambda \tilde{\mathbf{x}}_{it})^{-1/2} \exp\left\{ -\tfrac{1}{2} \frac{\left(H_{it} - \tilde{\mathbf{x}}'_{it}\bar{\lambda}\right)^2}{\left(1+\tilde{\mathbf{x}}'_{it}\Sigma_\lambda \tilde{\mathbf{x}}_{it}\right)} \right\} \right] \forall i = 1, \cdots, n, \quad (13)$$

where $H_i = [H_{it}, t = 1, \cdots, T]'$.

For (8) and (9) we can make a similar reduction to get:

$$\zeta_{c,it} = \bar{\mu}_1 + \bar{\mu}_2 \zeta_{c,i,t-1} + \varphi_{c,it}, \varphi_{c,it} \sim N(0, \omega_{ci}^2 + \mathbf{q}'_{it}\Sigma_\mu \mathbf{q}_{it}), \quad (14)$$

where $\mathbf{q}_{it} = [1, \zeta_{c,i,t-1}]'$, for all $c = 2, \cdots, C-1$. Collecting for all classes, we have:

$$\zeta_{it} = \bar{\mu}_1 1_{C-2} + \zeta_{i,t-1}\bar{\mu}_2 + \varphi_{it}, \forall i = 1, \cdots, n, t = 1, \cdots, T, \qquad (15)$$

where $\varphi_{it} = [\varphi_{c,it}, c = 2, \cdots, C-1]'$ and 1_{C-2} denotes a vector of ones. If $\Omega_{it} := \text{diag}\left[\omega_{ci}^2, c = 2, \cdots, C-1\right] + \mathbf{q}'_{it}\Sigma_\mu \mathbf{q}_{it}$, we have:

$$\varphi_{it} \sim N_{C-2}(0, \Omega_{it}) \qquad (16)$$

As there are only three different elements in Σ_{μ} and two different elements in $\bar{\mu}$, marginalizing with respect to the random coefficients results in a vast reduction to the number of parameters. Finally, we can write (15) for all individuals as:

$$z_t = \bar{\mu}_1 1_{n(C-2)} + z_{t-1}\bar{\mu}_2 + \xi_t, \forall\ t = 1,\cdots,T \qquad (17)$$

where $z_t = [z_{it}, i = 1,\cdots,n]'$. As individuals are independent, it follows that $\operatorname{cov}(\xi_t) := \Omega_t = \operatorname{diag}[\Omega_{it}, i = 1,\cdots,n]$. Although (17) is high-dimensional, it is very tightly parametrized.

Our Markov Chain Monte Carlo (MCMC) procedure can be described as follows (for further details see Appendix A.1):

(i) Drawing H_{it} from (13)

(ii) Drawing ζ_t from (17) and (13) and transforming to $\gamma_{c,it}$ via (7).

(iii)Drawing the *'deep'* parameters in θ, viz. $\bar{\lambda} := [\bar{\rho}, \bar{\beta}', \bar{\delta}']', \bar{\mu}, \Sigma_{\lambda}, \Sigma_{\mu}, \{\omega_{ci}^2\}$.

The bulk of computation lies in steps (i) and (ii). Regarding (i), we draw $H_i = \{H_{it}, t = 1,\cdots,T\}$ as a block to avoid introducing heavy autocorrelation in MCMC. We do the same in step (ii) where we draw ζ_t block wise for all individuals and time periods at once. For H_{it} we have the observation restrictions in (1). A draw for H_i can be obtained using a Langevin Metropolis algorithm due to Girolami and Calderhead (2011) as described in the technical Appendix A.2. This is repeated for each $i = 1,\cdots,n$. Even when the number of individuals is large the procedure is efficient as it uses gradient and Hessian information from (13).

For step (ii) to draw ζ_t as a block for all individuals and time periods we have to combine (16) and (12) -to form the posterior conditional kernel density. This is facilitated by a vector Particle Filtering (PF) approach, which evaluates the posterior for each θ. The version of PF we adopt is described in the technical Appendix A.3 and is based on Nemeth, Sherlock and Fearnhead (2014). We use $2^{16} = 65{,}536$ particles per component and iteration. In total, we use 250,000 MCMC iterations the first 50,000 of which are discarded to mitigate the impact of start-up effects. Convergence is assessed using Geweke's diagnostic, whilst effective sample size along with numerical standard errors and relative numerical efficiency have been monitored (Geweke, 1992). Our final results were robust to 2^8 and 2^{14} particles per component.

Part (iii), drawing the deep parameters θ, we follow Liu and West (2001) as explained in the technical Appendix A.3. Of course, covariance matrices like Σ_{λ} and Σ_{μ} are parametrized in terms of the elements of their Cholesky factorizations. For ω_{ci}^2 we parametrize in terms of p_{ci}, where $\omega_{ci} = \exp(p_{ci}), \forall\ i = 1,\cdots,n, c = 2,\cdots,C-1$.

2.4 Priors

Regarding priors, they are informative. Starting from the happiness equation in (2) and (4) we adopt:

$$\bar{\rho} \sim N(0,1) \qquad (18)$$

$$\bar{\beta} \sim N(0,10^2\mathbf{I}) \qquad (19)$$

$$\delta \sim N(0,10^2\mathbf{I}) \qquad (20)$$

For (8) and (9) in the Albert-Chib parametrization (7) we have:

$$\bar{\eta} \sim N(0,1), \bar{\xi} \sim N(0,1) \quad (21)$$

For Σ_λ and Σ_μ denoted generically by Σ we assume a Wishart prior:

$$p(\Sigma) \propto |\Sigma|^{-(d+\bar{a}+1)/2} \exp(-\mathrm{tr}\bar{A}\Sigma^{-1}), \quad (22)$$

where d is the dimensionality, $\bar{a} = 0.1$ and $\bar{A} = 0.1\mathbf{I}$. Since the covariance matrices are parametrized in terms of the elements of their Cholesky factorizations the priors are not conditionally conjugate, so they are not particularly easy to work with in terms of MCMC computation. Fortunately, our Langevin MCMC approach does not rely on conditional conjugate priors. Finally, for elements ω_{ci}^2 we assume:

$$\frac{\bar{b}}{\omega_{ci}^2} \sim \chi^2(\bar{v}), \quad (23)$$

where $\bar{b} = 0.01$ and $\bar{v} = 1$.

3. The data set and empirical results

3.1 The British Household Panel Survey (BHPS)

We employ the dataset of British Household Panel Survey (BHPS) available from the UK Data Archive. We include in our empirical analysis all available data since 1994. This survey is a longitudinal one that includes information regarding individuals of interviewed British households.[7] Since the survey commenced there have been 18 waves concerning 5050 British households. The British Household Panel Survey interviews every adult member of sample household and thereby it

[7] In the event that individuals change address or new individuals are added in the same household, the survey continues to include them and thereby interview them. The survey is asking the life satisfaction question over time. So, the sample of the present paper includes variability over time that enhances the information content. Individuals over time might change their views about how perceive their underlying life satisfaction. This information is present in our sample

follows the same representative sample of individuals over a period of years.

The survey is quite comprehensive and refers to marital status, age, household composition, housing conditions, residential mobility, education and training, health and the usage of health services, labour market, socio-economic values, and income from employment, benefits and pensions.[8] In our study, there are 166,381 observations made up of 26,027 individuals. The main variable of interest is overall life satisfaction, which ranges from a scale of 1 to 7, with 1 indicating not satisfaction at all and 7 completely satisfaction. The question for respondents is *'How dissatisfied or satisfied are you with your life overall?'*. In line with Binder and Broekel (2012), we consider individual's resources such as annual income, health, and education as variables that would affect life satisfaction and thus happiness.[9] We also account for a number of individual-specific characteristics (Oswald and Powdthavee, 2008; Tiefenbach and Kohlbacher, 2015). They are age, age squared, number of child, dummies for marital status (such as divorced, married, never married, separated, widowed), employment status (self-employed, in paid employment, unemployed, retired from paid work altogether, on maternity leave, looking after family or home, full time student/at school, long term sick or disabled, on a government training scheme, and something else), gender (male/female), and region (19 areas across Great Britain).

3.2. Estimation of Happiness Efficiency

We opt for Markov Chain Monte Carlo (MCMC) procedure to estimate the happiness efficiency. Moreover, we employ a vector Particle Filtering (PF) approach which evaluates the posterior as discussed in the methodological section. This is rather cumbersome as we use $2^{16} = 65{,}536$ particles per component and iteration. Overall, we use 250,000 MCMC iterations, the first 50,000 of which are discarded to mitigate the impact of start-up effects. Convergence is assessed using Geweke's diagnostic, whilst effective sample size along with numerical standard errors and relative numerical efficiency have been monitored (Geweke, 1992). Our final results were robust to 2^{8} and 2^{14} particles per component.

Table 1 reports the Bayesian estimation of happiness function for three models: Model A is a simple static model that does not include lagged happiness, Model B includes dynamics, namely lagged

[8] Liberini et al. (2017) in a recent paper explore yet another dimension of this sample, that is the significance of being happy for the outcomes of retrospective voting.

[9] Note in equation (3) we include is a vector of covariates that includes time-invariant characteristics. To this end, includes variables such as the annual household income, health and education (see Clark et al. 2008; Cheng et al. 2015), whereas include variables such as gender, age (see Oswald and Powdthavee, 2008; Tiefenbach and Kohlbacher, 2015).

happiness. We estimate this model because we expect (see Rayo and Becker 2007; Di Tella et al., 2010) that happiness is persistent although this is clearly a testable statistical hypothesis. Most empirical studies have ignored this possibility which, to anticipate some of our empirical results, could be present. At the very least, this suggests that ignoring persistence leads to incorrect estimates and inferences for all parameters in the model. Last, Model C is a dynamic model with random coefficients in all parameters so as to capture, also, heterogeneity across individuals. Intuitively, given the cross-sectional heterogeneity of our sample, Model C could better fit our data set. The rationale is that a model of the type "*one size fits all*" in terms of assuming common coefficients across individuals is hard to defend. More specifically, happiness levels and happiness efficiency are sensitive on assumptions made about the nature of parameters. If they are fixed and common for all individuals it means that explanatory variables affect happiness in the same way no matter what the individual's characteristics are. Again, on prior grounds, this assumption is highly questionable and, at the very least we would like to: i) test it empirically, and ii) examine whether it makes any difference in terms of estimated happiness efficiency. In what follows, we would test for whether Model C is a better fit rather than imposing it.

Also, note that post z is the posterior mean divided by the posterior standard deviation. We employ post z to evaluate the statistical significance of our estimation. We have statistical significance in the frequentist or sampling-theory context when post z exceeds 1.96 in absolute value (at the 5% level of statistical significance).

The results, presented in Table 1, indicate that most right-hand side variables assert the expected impact on happiness. Household income has a positive impact on happiness, albeit its magnitude is small. Blanchflower and Oswald (2004) also show that income positively affect happiness that has been also verified by Binder and Coad (2011) and Clark et al., (2008). However, Easterlin et al. (2010) argue that happiness would increase as income increases, but there is a threshold of a satiation point beyond which any further increase in income might not impact on life satisfaction. Conventional wisdom might dictate that money does not always buy happiness. Compared to women, men are generally less happy. There is a U-shape relationship between age and happiness, with happiness declining until middle age and improving thereafter. Marital status also affects happiness. Single, separated, divorced, and widowed individuals are all less happy with their lives than married individuals. Fernández-Ballesteros et al., (2001) show that married people are happier than those who are separated, divorced, widowed, or were never married. Married people could profit from physical and emotional provisions (Stack and Eshleman, 1998). In addition, Blanchflower and Oswald (2004) reasons that the frequency of sexual activity is positively related to happiness while the former is significantly higher for married people compared to single, divorced, widowed, or separated.

The number of children in the household has a negative impact on happiness. Having and raising children necessitates substantial resources that could come at the cost of happiness, see Binder and Coad (2011). Indeed, most studies show that having children does not make people happier (Clark *et al.* 2008). However, Mastekaasa, (1994) argues that growing up children could increase happiness as they enlarge their parents' social network and engaging in social activities improve happiness.

The retired, students and those in family care also enjoy a positive boost in their happiness. Consistent with existing evidence (Clark et al. 2008; Cheng et al. 2015) and a widely held belief, employees are less satisfied with their lives than the self-employed (reference category).

As expected, unemployment has one of the strongest negative effects on happiness which is in line with previous studies (Lucas et al., 2008; Stutzer, 2004; Di Tella et al., 2010). It is worth noting that Lucas et al. (2008) shows that unemployment could persistently negatively affect happiness as it acts as a shock that is difficult to recover from. Interestingly, self-employed seem to be happier than employees as in Binder and Coad (2013).

Equally unsurprisingly, there is a positive correlation between health and happiness. As health deteriorates from excellent (reference category) to good, to fairly or very poor, happiness decreases. This positive effect of health is in line with the literature (Diener and Seligman, 2002; Sabatini, 2014). Higher educational attainment impacts negatively on happiness. While this result appears to be puzzling, it is not out of line with previous research, which has produced mixed results about the link between education and happiness (Powdthavee et al. 2015). In a recent paper, Brennan et al. (2014) show that the significance of education varies as they demonstrate that the productivity in education is affected by non-discretionary inputs such as socio-economic variables like income and parental education. However there is some variability in reported results as Binder and Coad (2011) report no significant impact of education on life satisfaction for UK, whereas in another Binder and Coad (2013) they show a significant and negative relationship between education and life satisfaction.

Moreover, in line with US study of Blanchflower and Oswald (2004) gender and age are negatively related to happiness while there is an underlying U-shape type of age-happiness relationship (see also

Frijters and Beatton, 2012). Happiness increases in older age could be connected to old-age social activities.

In model C the coefficient on lagged happiness is positive and statistically significant and implies a slow adaptation to changed circumstances. This slow adaptation contrasts with other reported estimates of adaptation (see e.g., Botan and Truglia 2011; Chrostek 2013; Botan and Truglia 2011; Botan and Truglia 2011).

Table 1: Bayesian parameter estimates of the happiness function.

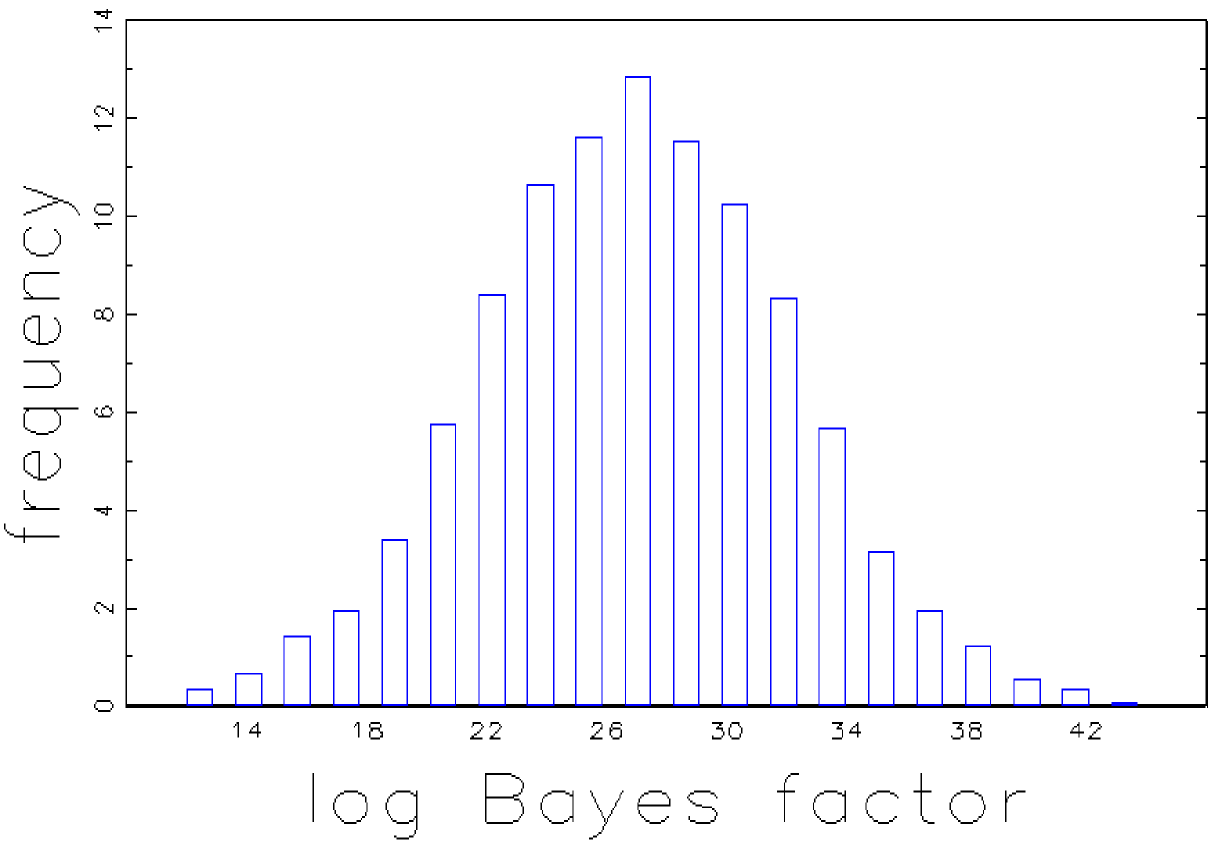
	Model A Static model		**Model B Dynamic model -no random coefficients**		**Model C Dynamic model - with random coefficients**	
	post. mean	post. z	post. mean	post. z	post. mean	post. z
const	5.92**	115.72	4.92**	115.72	1.69**	115.72
Lag happiness			0.72**	50.41	0.85**	34.79
age	-0.02**	-10.76	-0.01**	-10.76	-0.01**	-10.76
age2	0.002**	17.87	0.001**	17.87	0.001**	17.87
hhldinc	0.001**	5.51	0.002**	5.51	0.001**	5.51
nchild	-0.03**	-7.10	-0.03**	-7.10	-0.02**	-7.10
Marst1	-0.26**	-14.23	-0.29**	-14.23	-0.17**	-14.23
Marst2	-0.14**	-9.59	-0.16**	-9.59	-0.11**	-9.59
Marst3	-0.23**	-13.47	-0.18**	-13.47	-0.14**	-13.47
Marst4	-0.12**	-9.58	-0.10**	-9.58	-0.06**	-9.58
Empstat1	0.01	0.58	0.01**	-2.35	0.02**	3.23
Empstat2	-0.29**	-17.78	-0.31**	-17.78	-0.23**	-17.78
Empstat3	0.02	1.52	0.01	1.52	0.01	1.52
Empstat4	0.27**	7.95	0.25**	7.95	0.09**	7.95
Empstat5	-0.07**	-3.78	-0.06**	-3.78	-0.02**	-3.78
Empstat6	0.04*	2.71	0.03*	2.71	0.01*	2.71
Empstat7	-0.48**	-22.55	-0.49**	-22.55	-0.24**	-22.55
Empstat8	0.00	-0.02	0.00	-0.02	0.00	-0.02
Empstat9	-0.07*	-2.20	-0.04*	-2.20	-0.03*	-2.20
sex1	-0.06**	-6.10	-0.06**	-6.10	-0.03**	-6.10
Hlth1	0.19**	30.52	0.13**	30.52	0.07**	30.52
Hlth2	-0.51**	-68.28	-0.35**	-68.28	-0.35**	-68.28
Educ1	0.11**	11.13	0.12**	11.13	0.08**	11.13
Educ2	-0.09**	-6.94	-0.10**	-6.94	-0.03**	-6.94
Region	yes		yes		yes	
Years	yes		yes		yes	
Bayes factors	1.000		4.717		12.282	
R^2	0.744		0.797		0.935	

Note: ** notes significance at 1%, * at 5%. age captures the age of the individual, age2 is the squared term, hhdinc is the household income /1000, nchild number of children, marst captures marital status (marst1 separated, marst2 divorced, marst3 widowed, marst4 never married, empstat the employment status (empstat1 in paid employment, empstat2 unemployed, empstat3 retired, empstat4 family care, empstat5 family care, empstat6 full time student, empstat7 long term sick, empstat8 maternity leave, empstat9 government training, sex counts for male (sex1)/female, hlth health (hlth1 is good, hlth2 very poor), educ education (educ1 high, educ2 medium). Region captures dummies for the different regions of UK (19 areas across Great Britain). Years captures time effects from year 1 to year 12. The Bayesian estimations is based on a sample of 166,381 person-year observations made up of 26,027 individuals, through 250,000 MCMC iterations (with the first 50,000 of which are discarded to mitigate the impact of start-up effects). A vector Particle Filtering (PF) approach evaluates the posterior with 2^{16}=65536 particles per component and iteration. Geweke's diagnostic confirms convergence to 2^8 and 2^{14} particles per component.

Provided that we estimate two dynamic models (Model B which includes dynamic terms, namely lagged happiness and Model C which is a dynamic model with random coefficients in all parameters to capture heterogeneity across individuals) we report next the frequency of the log of Bayes factor of model C compared to model B so as to test their performance.

Results show that Model C preforms 10^{26} better than model B (see Bayes factors in Table 1, but also Figure 1). In addition, the goodness of fit as captured by R^2 also shows that Model C is superior. To this end, we shall focus on happiness efficiency as derived from Model C.

Figure 1. The density of the log Bayes factor of model C against model B.



Note: Authors' estimations. The log Bayes factor is based on a sample of 166,381 person-year observations made up of 26,027 individuals, through 250,000 MCMC iterations (with the first 50,000 of which are discarded to mitigate the impact of start-up effects). A vector Particle Filtering (PF) approach evaluates the posterior with 2^{16}=65536 particles per component and iteration. Geweke's diagnostic confirms convergence to 2^8 and 2^{14} particles per component.

Table 2 reports descriptive statistics of happiness efficiency scores for the three models. The happiness efficiency score is always positive. There is some variability in the efficiency scores across models that range from 0.55 in Model A to 0.79 in Model C, the preferred model. The mean efficiency score over the three models is around 70 per cent, which is indicative of a substantial happiness efficiency deficit. This implies that British individuals are not reaching their happiness frontier and that there is some considerable room for improvement in happiness efficiency by around 30 per cent.[10] Note that the reported results show that there is some happiness inefficiency present in our sample. In terms of intuition, our modelling reveals that British individuals face a 'happiness loss' in terms of happiness inefficiency of around 30 per cent. This 'happiness loss' is associated with using their resources that they have at their disposal at lower level of efficiency.

Table 2: The Bayesian happiness efficiency.

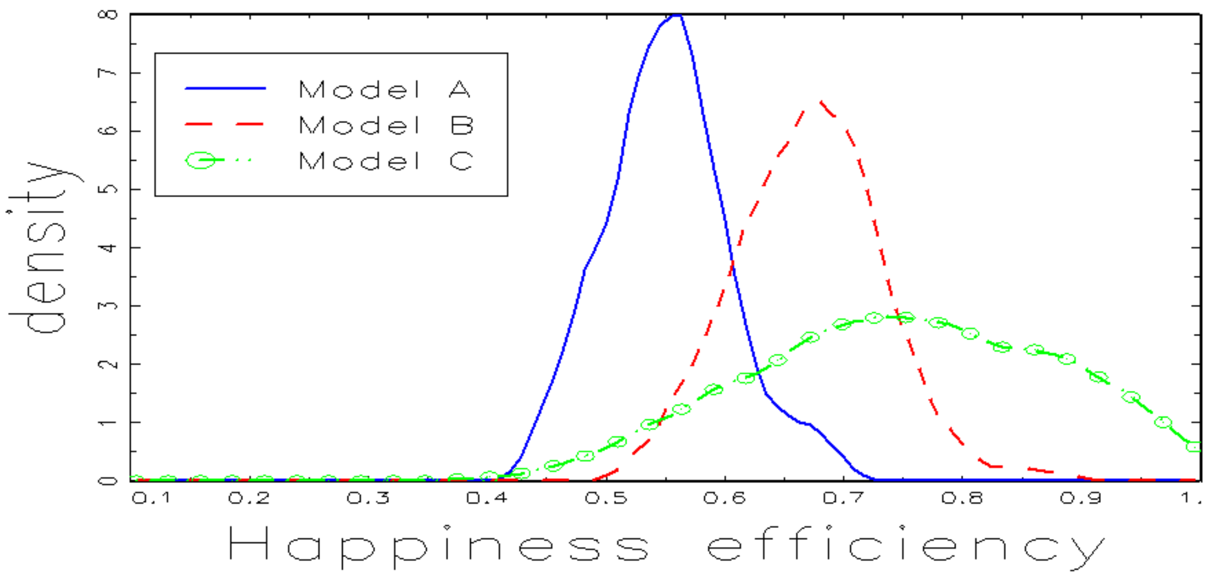
	Obs	Mean	Std. Dev.	Min	Max
Model A	166,381	0.552321	0.103045	1.43E-06	0.672202
Model B	166,381	0.715468	0.147443	2.18E-06	0.948433
Model C	166,381	0.795471	0.071959	1.60E-06	0.985544

Note: Model A is the static one, Model B is the dynamic without random coefficients and Model C is the dynamic with random coefficients. Authors' estimations. The happiness efficiency is based on a sample of 166,381 person-year observations made up of 26,027 individuals, through 250,000 MCMC iterations (with the first 50,000 of which are discarded to mitigate the impact of start-up effects). A vector Particle Filtering (PF) approach evaluates the posterior with 2^{16}=65536 particles per component and iteration. Geweke's diagnostic confirms convergence to 2^8 and 2^{14} particles per component.

The advantage of our modelling is that estimates the entire distribution of the happiness efficiency scores over the three models, see Figure 2. Model C corrects for the leptokurtic characteristic of Model B and Model A, though it reports higher average happiness efficiency. The results show that British households have, on average, achieved 70 percent of their happiness efficiency. This figure of happiness efficiency is not small, though it shows that British households could further increase their happiness efficiency by 30% by using more efficiently existing resources. This result implies that the typical British household could improve its happiness efficiency without relying on increasing its resources, for example its income.

Figure 2. The sample distributions of happiness efficiency.

[10] This result complements Marcenaro-Gutierrez et al. (2010) that were able to identify the level of satisfaction of Spanish workers. However, the present approach has the advantage of modelling based on latent variables that allows revealing the whole distribution of happiness efficiency across all individuals that take part in the sample. In a sense, this approach is permitting a ranking of individuals based on the happiness efficiency.



Note: Model A is the static one, Model B is the dynamic without random coefficients and Model C is the dynamic with random coefficients. Authors' estimations. The happiness efficiency is based on a sample of 166,381 person-year observations made up of 26,027 individuals, through 250,000 MCMC iterations (with the first 50,000 of which are discarded to mitigate the impact of start-up effects). A vector Particle Filtering (PF) approach evaluates the posterior with 2^{16}=65536 particles per component and iteration. Geweke's diagnostic confirms convergence to 2^8 and 2^{14} particles per component.

Overall, by treating happiness as a latent variable we reveal that, on average, the British happiness efficiency has significant room to further improvement given the existed individuals' characteristics and resources. The results show that given a higher happiness efficiency is strictly preferred as the individual seek to maximize her happiness in line with Rayo and Becker (2007) theory, a typical British individual could improve her happiness efficiency by a further 30% so as to reach the maximum feasible happiness. This result is of interest as we confirm Rayo and Becker (2007) prediction that there is happiness efficiency.

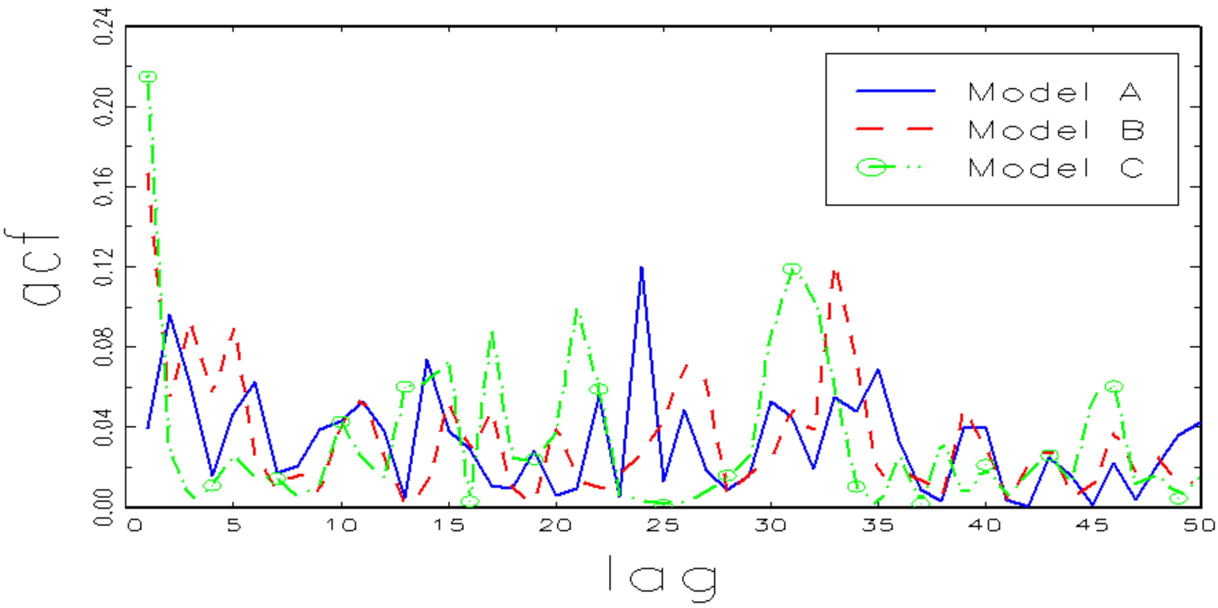
Fleurbaey and Schwandt (2015) conduct an on-line survey of a plethora of previous empirical studies on happiness to show that 90% of the respondents report that they seek to maximize their happiness. What the online survey did not show is whether respondents reach their maximum happiness. In this paper, we provide evidence that British household could seek to maximize their happiness, though there is considerable ground to be gained in terms of higher values of happiness efficiency.

Operational research modelling has been providing valuable information in various subject areas. Following from White (2018) who argue of the importance of providing also social operational research modeling and given the content of Rayo and Becker (2007) theory, the results above on British individual's happiness efficiency advances our knowledge and thereby information on a subject, that of measuring happiness, that has attracted much attention in recent years by academics and policy makers alike.

3.3 Testing for the fitness of happiness efficiency: autocorrelation function (acf) and relative numerical efficiency (RNE).

In this section we provide evidence of the statistical validity of our happiness and happiness efficiency estimations. In some detail, we use a Girolami and Calderhead (2012) algorithm (GC thereafter) as the first stage GMM estimator and then we employ MCMC until convergence. Depending on the model and the subsample this takes 1,000 to 3,000 iterations. For safety we run 10,000 iterations. Then we run another 50,000 MCMC iterations to obtain final results for posterior moments and densities of parameters of happiness function. Figure 3 reports the autocorrelation function that shows that the GC algorithm for Model C performs better than Model A or Model B.

Figure 3. Autocorrelation function (acf) of model A, B and C.

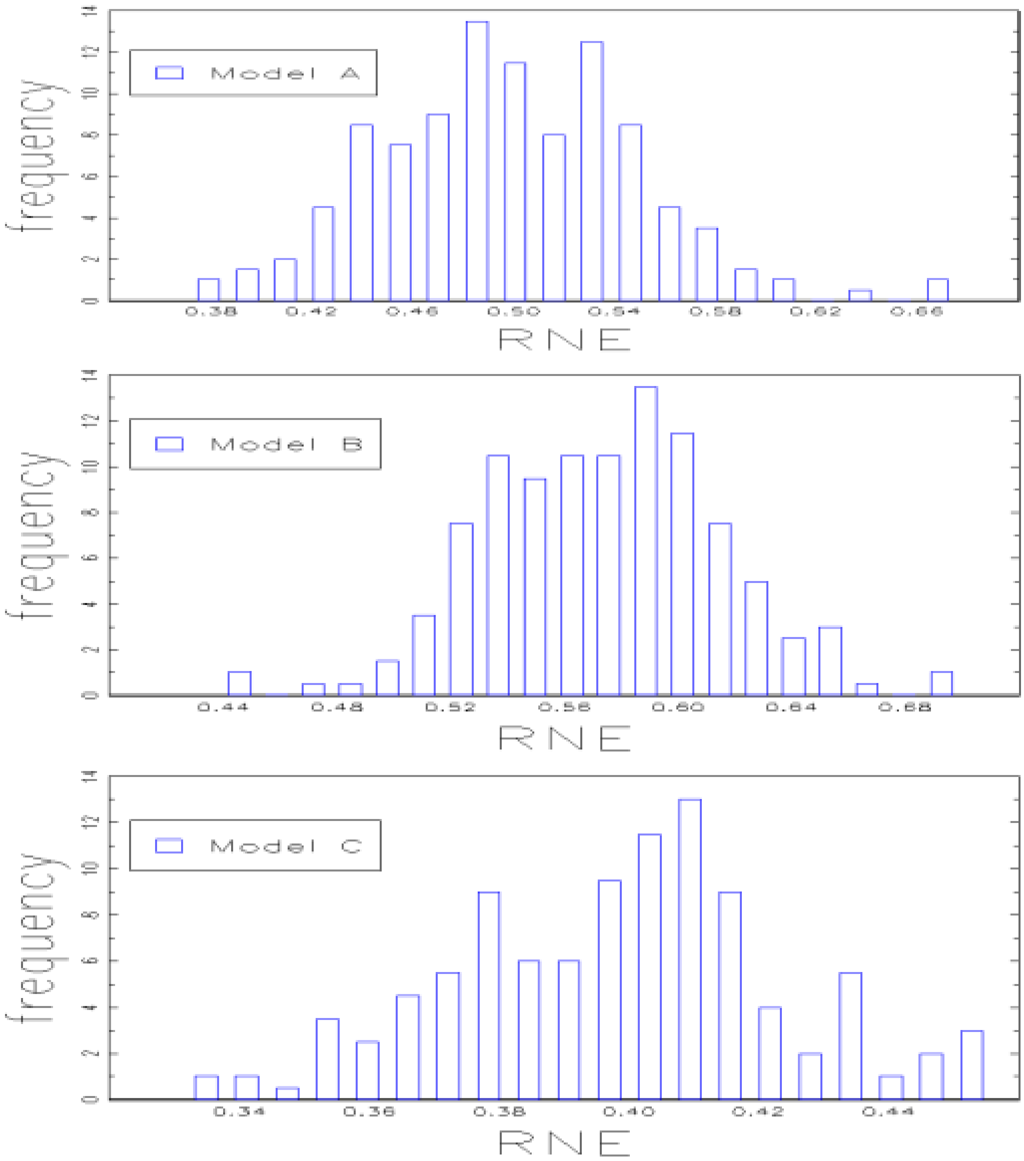


Note: Authors' estimations. 250,000 MCMC iterations (with the first 50,000 of which are discarded to mitigate the impact of start-up effects). A vector Particle Filtering (PF) approach evaluates the posterior with 2^{16}=65536 particles per component and iteration. Geweke's diagnostic confirms convergence to 2^8 and 2^{14} particles per component.

This statistical evidence is of some importance as we prove that a dynamic model with random coefficients in all parameters would capture the heterogeneity across individuals. A common assumption in the literature is that happiness function parameters are fixed and common for all individuals which would imply that explanatory variables would affect their happiness in the same manner (Fleurbaey and Schwandt 2015; Cheng et al. 2015; Clark et al. 2008;). That is individuals are treated as highly homogenous. This is certainly not a plausible assumption. Here we provide evidence that underlying heterogeneity is of importance.

To assess convergence of our Bayesian MCMC we employ, in addition, Geweke's diagnostic that allows monitoring relative numerical efficiency. We confirm that our estimations were robust to 2^8 and 2^{14} particles per component. Figure 4, in addition, reports the relative numerical efficiency (RNE) of Model A, B and C. The frequency of RNEs shows the superiority of Model C over Model A and Model B.

Figure 4. Relative numerical efficiency (RNE) of Model A, B and C.



Note: Authors' estimations. 250,000 MCMC iterations (with the first 50,000 of which are discarded to mitigate the impact of start-up effects). A vector Particle Filtering (PF) approach evaluates the posterior with 2^{16}=65536 particles per component and iteration. Relative numerical efficiency had been monitored by Geweke's diagnostic that converges to 2^8 and 2^{14} particles per component.

3.4 The Role of Personality Traits on Happiness Efficiency

Previous research highlights that the big five personality traits would affect happiness (see Lucas and Diener 2015; Gosling, et al. 2003). Following from this literature whether personality traits would also impact on happiness efficiency. Table 3 reports results for the importance of personality traits for a variety of models for consistency and robustness. The main variables of interest are: openness to experience (OE), agreeableness (AG), consciousness (CON), extraversion (EX), neuroticism (NR). Various models are estimated, including lagged values of happiness, differences in household income and different configurations of the explanatory variables.

In addition, we propose to consider individual-specific thresholds to happiness efficiency. Relative to standard random effects (see Model C of the previous section) by incorporating individual-specific thresholds to happiness we consider heterogeneity due to unobserved time-invariant characteristics across individuals. It is worth noting that it is not customary in similar studies (see Binder and Broekel, 2012) to consider heterogeneity across individuals. This has been a serious limitation in previous studies. To treat happiness efficiency as time-invariant is not plausible given that any random or fixed effects across individuals in the sample might not be part of inefficiency. This study departs from similar strong assumption that happiness efficiency, and thereby inefficiency, would be time invariant.

Table 3 presents considers individual's characteristics, including personality traits, and also persistence in happiness. It is evident from these estimates that individuals who score high, in particular, in agreeableness and extraversion have also high happiness efficiency, i.e., they are closer to their happiness frontier (see Model E, where we also include interaction terms between lagged happiness and personality traits). In contrast, those who score highly in neuroticism would face lower levels of happiness efficiency and thus would move further away from the happiness frontier. These results are broadly in line with the theoretical hypotheses of Gosling and Rentfrow (2003). The same is true for individuals who score highly in openness to experience and consciousness, when the latter interact with lagged happiness.

There is plethora of studies that show that greater extraversion, lower neuroticism, more agreeableness and conscientiousness are positively associated with higher happiness (McCrae and Costa, 1991; Brebner *et al.*, 1995; Diener and Seligman, 2002; Steel and Ones, 2002; Cheng and Furnham, 2003; Hayes and Joseph, 2003). Individuals who are open to experience are more efficient at achieving their potential happiness in line with Furnham and Cheng (1997) and Steel et al. (2008). However, Diener and Seligman (2002) report that openness is not a personality trait that would define what they call ‘very happy’ individuals while Chamorro-Premuzic et al. (2007) show no statistically significant relationship between openness and happiness.

Table 3. The role of personality traits.

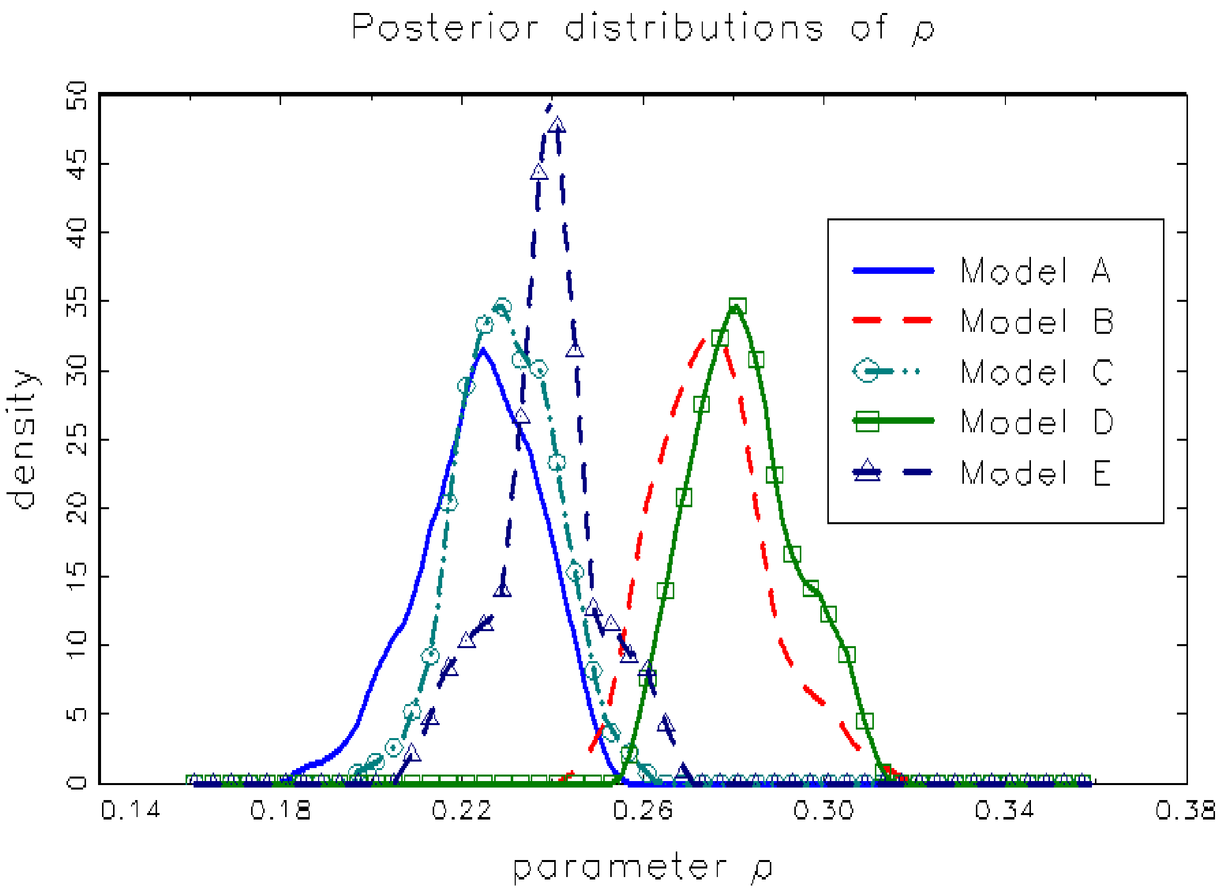
	Model A	Model B	Model C	Model D		Model E
			Differences	Lagged variables	Differences	
H_{t-1}	0.225	0.277	0.231	0.282		0.240
	(0.012)	(0.011)	(0.010)	(0.010)		(0.013)
age	-0.012	-0.013	-0.021	0.0012	-0.017	-0.015
	(0.005)	(0.008)	(0.004)	(0.0041)	(0.003)	(0.004)
age^2	-0.001	-0.002	-0.003	-0.003	-0.003	-0.002
	(0.001)	(0.002)	(0.001)	(0.005)	(0.001)	(0.0003)
hhldinc	0.0022		0.005	-0.002	0.0032	0.001
	(0.0012)		(0.002)	(0.005)	(0.0002)	(0.0002)
Δhhldinc	0.013	0.0092				
	(0.007)	(0.003)				
|Δhhldinc|	0.0025	0.0027	0.0012		0.0027	
	(0.0011)	(0.001)	(0.0004)		(0.0012)	
$Hhldinc_{t-1}$		0.0032				
		(0.044)				
nchild	-0.0022	-0.0012	-0.011	0.011	-0.012	-0.012
	(0.0011)	(0.0013)	(0.003)	(0.035)	(0.002)	(0.002)
Marst1	0.025	0.017	0.012	0.020	0.013	0.012
	(0.003)	(0.002)	(0.001)	(0.011)	(0.004)	(0.003)
Marst2	0.019	0.013	0.014	0.0052	0.021	0.019
	(0.002)	(0.003)	(0.005)	(0.018)	(0.002)	(0.003)
Marst3	0.027	0.015	0.007	-0.005	0.012	0.018
	(0.004)	(0.003)	(0.001)	(0.013)	(0.001)	(0.002)
Marst4	0.013	0.007	0.0011	0.013	0.017	0.012
	(0.004)	(0.002)	(0.002)	(0.012)	(0.003)	(0.002)
Empstat1	-0.121	-0.015	-0.112	-0.033	-0.044	-0.033
	(0.003)	(0.004)	(0.005)	(0.044)	(0.002)	(0.002)
Empstat2	0.007	0.005	0.013	0.021	0.012	0.013
	(0.002)	(0.001)	(0.002)	(0.012)	(0.004)	(0.002)
Empstat3	0.001	0.003	0.001	0.002	0.0027	0.0015
	(0.001)	(0.002)	(0.001)	(0.014)	(0.001)	(0.0003)
Empstat4	0.0025	0.0041	0.001	0.002	0.001	0.001
	(0.002)	(0.001)	(0.0002)	(0.007)	(0.0004)	(0.002)
Empstat5	0.005	0.004	0.0044	0.008	0.0014	0.0025
	(0.002)	(0.001)	(0.001)	(0.007)	(0.001)	(0.001)
Empstat6	-0.004	-0.002	-0.0032	-0.002	-0.002	-0.006
	(0.001)	(0.001)	(0.0001)	(0.003)	(0.0002)	(0.002)
Empstat7	0.008	0.009	0.0024	0.013	0.0162	0.012
	(0.003)	(0.002)	(0.002)	(0.012)	(0.002)	(0.002)
Empstat8	-0.015	-0.009	-0.0015	0.003	-0.012	-0.013
	(0.022)	(0.008)	(0.0002)	(0.004)	(0.003)	(0.008)
Empstat9	-0.044	-0.021	-0.0011	0.005	-0.007	-0.022
	(0.003)	(0.002)	(0.0002)	(0.004)	(0.0014)	(0.003)
Sex1	-0.021	-0.051	-0.012*	-0.013	-0.024	-0.020
	(0.007)	(0.009)	(0.002)	(0.014)	(0.003)	(0.002)
Hlth1	-0.111		-0.032	0.0021	-0.030	-0.113
	(0.005)		(0.002)	(0.013)	(0.003)	(0.002)
Hlth2	-0.102		0.044	0.015	0.021	0.018
	(0.002)		(0.001)	(0.014)	(0.002)	(0.003)
Educ1	0.005		0.014	0.014	0.011	0.012
	(0.002)		(0.012)	(0.077)	(0.003)	(0.003)
Educ2	0.009		0.002	0.015	0.012	0.013
	(0.001)		(0.001)	(0.012)	(0.003)	(0.001)
TRAITS						
OE		0.015		0.014		0.017
		(0.003)		(0.001)		(0.001)
AG		0.122		0.012		0.144

		(0.004)		(0.004)		(0.012)
CON		0.015		0.015		0.015
		(0.022)		(0.004)		(0.002)
EX		0.135		0.026		0.041
		(0.013)		(0.011)		(0.0042)
NR		-0.172		-0.044		-0.074
		(0.011)		(0.005)		(0.003)
H_{t-1} *OE						-0.011
						(0.001)
H_{t-1} *AG						0.012
						(0.001)
H_{t-1} *CON						-0.013
						(0.002)
H_{t-1} *EX						0.014
						(0.002)
H_{t-1} *NR						-0.132
						(0.004)
region	yes	yes	yes	yes	yes	yes
years	yes	yes	yes	yes	yes	yes
BAYES FACTOR	6.267	7.335	13.544	14.351	16.233	15.217

Notes: Coefficients are posterior means and posterior standard deviations are in parentheses. Age captures the age of the individual, age2 is the squared term, hhdinc is the household income /1000, nchild number of children, marst captures marital status (marst1 separated, marst2 divorced, marst3 widowed, marst4 never married, empstat the employment status (empstat1 in paid employment, empstat2 unemployed, empstat3 retired, empstat4 family care, empstat5 family care, empstat6 full time student, empstat7 long term sick, empstat8 maternity leave, empstat9 government training, sex counts for male (sex1)/female, hlth health (hlth1 is good, hlth2 very poor), educ education (educ1 high, educ2 medium). Region captures dummies for the different regions of UK (19 areas across Great Britain). Years captures time effects from year 1 to year 12. Bayesian inference is based on a sample of 166,381 person-year observations made up of 26,027 individuals, through 250,000 MCMC iterations (with the first 50,000 of which are discarded to mitigate the impact of start-up effects). A vector Particle Filtering (PF) approach evaluates the posterior with 2^{16}=65536 particles per component and iteration. Geweke's diagnostic confirms convergence to 2^8 and 2^{14} particles per component. TRAITS are OE=openness to experience, AG=agreeableness, CON=consciousness, EX=extraversion, NR=neuroticism. Age and age^2 are included in levels and never differenced. Model VI includes both lagged values and differences. For sex, it includes only the dummy variable. Regional and time dummy variables are included in all models. Bayes factors are ratios of marginal (or integrated) likelihoods of any model relative to model I. H_{t-1} denotes lagged happiness. post z reported in parenthesis, the posterior mean divided by the posterior standard deviation.

In addition, we report in Figure 5 the marginal posteriors of persistence in happiness efficiency from Table 3. The Figure 5 reports the persistence parameter ρ (note that this parameter comes from equation 2 of happiness function) which indicates the persistence of happiness efficiency due to intrinsic individual heterogeneity as captured by the models in Table 3. The results show that the persistence parameter ρ shows some variability from model to model, but yet it is quite of substance in terms of magnitude from 0.22 to 0.28 on average. As persistence in happiness efficiency should be viewed in the context of outside the control of the specific individual, any improvement in happiness efficiency should also warrant action to those factors outside the control of the individual.

Figure 5. Marginal Posteriors of Persistence from Table 3.



Note: Authors' estimations. Relative numerical efficiency had been monitored by Geweke's diagnostic that converges to 2^8 and 2^{14} particles per component.

To advance our understanding of how personality traits affect happiness efficiency, we also estimate models treating personality traits as latent variables (see Table 4). By doing so, we are embedding in the models a latent trait indicator which is stochastic, dynamic and depends on the five standard personality traits. Compared again with other models whose properties are known only asymptotically and are known to perform erratically in finite samples, the Bayes approach developed here is better suited for finite samples and data with a lot of individual heterogeneity. Unlike standard regression that focuses on the mean of happiness efficiency distribution, treating personality traits in the context of latent variables allows us to explore further its impact in happiness. In this instance, we model as latent variable the personality trait of being agreeable. It appears that certain personality traits such as being agreeable improves efficiency, suggesting that it would be worth in terms of happiness to raise agreeableness. Conscientiousness is also consistently and significantly positive while neuroticism is consistently and significantly negative (results are available upon request).

Table 4. Modeling personality traits as latent variables.

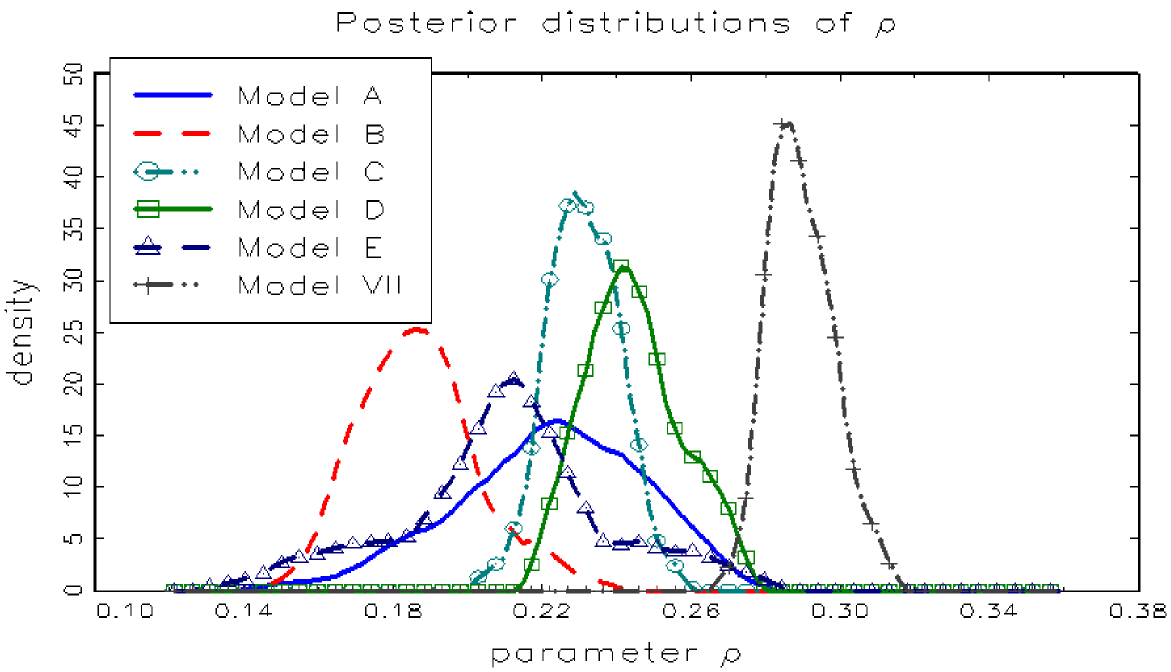
	Model A	Model B	Model C	Model D	Model E	Model F

		Lagged Happiness	Differences	Differences	Lagged Happiness	Differences	
H_{t-1}	0.224	0.189	0.232	0.244	0.213		0.289
	(0.023)	(0.014)	(0.009)	(0.011)	(0.032)		(0.014)
age	-0.004	-0.013	-0.012	-0.014	0.0012	-0.015	-0.019
	(0.001)	(0.012)	(0.011)	(0.002)	(0.023)	(0.013)	(0.014)
age^2	-0.001	0.004	0.002	0.002	0.007	0.011	0.003
	(0.0002)	(0.003)	(0.011)	(0.012)	(0.044)	(0.012)	(0.012)
hhldinc		0.0032	0.0012	0.005	-0.0021	0.0023	0.002
		(0.002)	(0.0001)	(0.0002)	(0.001)	(0.002)	(0.0001)
Δhhldinc	0.0312						0.0171
	(0.0012)						(0.0032)
$\|\Delta$hhldinc$\|$	0.0122	0.0117	0.0013	0.0022		0.0011	0.0087
	(0.0017)	(0.0021)	(0.0005)	(0.00043)		(0.0002)	(0.0022)
nchild	-0.013	-0.013	-0.012	-0.011	0.012	-0.015	-0.022
	(0.009)	(0.012)	(0.003)	(0.002)	(0.032)	(0.002)	(0.023)
Marst1	-0.12	-0.115	0.015	0.012	0.017	0.011	0.017
	(0.071)	(0.055)	(0.003)	(0.001)	(0.012)	(0.003)	(0.004)
Marst2	-0.13	-0.105	0.014	0.014	0.003	0.003	0.012
	(0.072)	(0.022)	(0.003)	(0.002)	(0.014)	(0.001)	(0.003)
Marst3	-0.082	-0.115	0.013	0.017	-0.001	0.004	0.010
	(0.076)	(0.052)	(0.001)	(0.003)	(0.023)	(0.001)	(0.002)
Marst4	-0.073	-0.133	0.031	0.022	0.013	0.013	0.014
	(0.084)	(0.117)	(0.002)	(0.002)	(0.011)	(0.003)	(0.002)
Empstat1	0.015	0.017	-0.013	-0.044	-0.015	-0.012	-0.122
	(0.012)	(0.012)	(0.004)	(0.002)	(0.023)	(0.003)	(0.014)
Empstat2	-0.044	-0.128	0.032	0.016	0.013	0.012	0.010
	(0.071)	(0.013)	(0.004)	(0.002)	(0.011)	(0.002)	(0.002)
Empstat3	0.022	0.021	0.003	0.002	0.001	0.005	0.012
	(0.021)	(0.017)	(0.002)	(0.002)	(0.003)	(0.001)	(0.002)
Empstat4	0.071	0.089	0.004	0.005	0.012	0.006	0.003
	(0.044)	(0.044)	(0.001)	(0.002)	(0.012)	(0.002)	(0.0001)
Empstat5	0.021	0.017	0.003	0.005	0.004	0.0044	0.003
	(0.015)	(0.011)	(0.002)	(0.001)	(0.005)	(0.0011)	(0.0002)
Empstat6	0.013	0.019	-0.002	-0.004	-0.005	-0.007	-0.003
	(0.032)	(0.021)	(0.002)	(0.0002)	(0.005)	(0.0001)	(0.001)
Empstat7	-0.027	-0.019	0.012	0.015	0.012	0.0071	0.0091
	(0.019)	(0.015)	(0.004)	(0.0019)	(0.011)	(0.005)	(0.003)
Empstat8	0.004	0.0017	-0.032	-0.012	0.015	-0.014	-0.0071
	(0.001)	(0.001)	(0.004)	(0.002)	(0.013)	(0.004)	(0.003)
Empstat9	-0.023	-0.012	-0.017	-0.025	0.002	-0.016	-0.013
	(0.017)	(0.012)	(0.002)	(0.003)	(0.031)	(0.0011)	(0.002)
Sex1	-0.022	-0.022	-0.012*	-0.014*	-0.011	-0.010	-0.017
	(0.014)	(0.012)	(0.002)	(0.002)	(0.012)	(0.002)	(0.002)
Hlth1	-0.022	-0.011	-0.111	-0.103	0.022	-0.014	-0.013
	(0.013)	(0.006)	(0.002)	(0.001)	(0.011)	(0.002)	(0.001)
Hlth2	-0.071	-0.065	0.055	0.112	0.077	0.063	0.102
	(0.032)	(0.041)	(0.003)	(0.003)	(0.013)	(0.002)	(0.005)
Educ1	0.122	0.044	0.013	0.012	0.012	0.011	0.005
	(0.074)	(0.082)	(0.003)	(0.003)	(0.044)	(0.001)	(0.002)
Educ2	-0.015	-0.022	0.143	0.103	0.025	0.021	0.0072
	(0.013)	(0.021)	(0.003)	(0.015)	(0.022)	(0.001)	(0.003)
traits, TR^*_{it}	0.115	0.104	0.064	0.072	0.089	0.062	0.017
	(0.022)	(0.044)	(0.032)	(0.0014)	(0.061)	(0.052)	(0.002)
$H_{t-1}\ TR^*_{i,t-1}$							0.017
							(0.0013)
region	yes	yes	yes	yes	yes	yes	yes
years	yes	yes	yes	yes	yes	yes	yes
BAYES FACTOR	11.772	13.44	18.535	17.224		13.255	37.281

Notes: Coefficients are posterior means and posterior standard deviations are in parentheses. Age captures the age of the individual, age^2 is the squared term, hhdinc is the household income /1000, nchild number of children, marst captures marital status (marst1 separated, marst2 divorced, marst3 widowed, marst4 never married, empstat the employment status (empstat1 in paid employment, empstat2 unemployed, empstat3 retired, empstat4 family care, empstat5 family care, empstat6 full time student, empstat7 long term sick, empstat8 maternity leave, empstat9 government training, sex counts for male (sex1)/female, hlth health (hlth1 is good, hlth2 very poor), educ education (educ1 high, educ2 medium). Region captures dummies for the different regions of UK (19 areas across Great Britain). Years captures time effects from year 1 to year 12. Bayesian inference is based on a sample of 166,381 person-year observations made up of 26,027 individuals, through 250,000 MCMC iterations (with the first 50,000 of which are discarded to mitigate the impact of start-up effects). A vector Particle Filtering (PF) approach evaluates the posterior with 2^{16}=65536 particles per component and iteration. Geweke's diagnostic confirms convergence to 2^8 and 2^{14} particles per component. TR^*_{it} denotes latent traits, see Table 3. Model F includes both lagged values and differences. For sex, it includes only the dummy variable. Regional and time dummy variables are included in all models. Bayes factors are ratios of marginal (or integrated) likelihoods of any model relative to model I in Table 1. H_{t-1} denotes lagged happiness. post z reported in parenthesis, the posterior mean divided by the posterior standard deviation.

In Figure 6 we report marginal posterior densities of the persistence parameter ρ. Results of persistence in happiness inefficiency is in line with the one above. Again, the persistence parameter ρ is of some magnitude, in particular for Model F. It is imperative, therefore, to view achieving happiness efficiency in light of personality traits as dynamic process that crucially depends on intrinsic individual characteristics.

Figure 6. Marginal Posteriors of Persistence from Table 4.



Note: Authors' estimations. Relative numerical efficiency has been monitored by Geweke's diagnostic that converges to 2^8 and 2^{14} particles per component.

4. Conclusion

This paper builds on Rayo and Becker (2007) who provide a theory of happiness and happiness efficiency and relaxes strong assumptions about the true definition of happiness. We model happiness as an unobservable variable opting for a Bayesian latent stochastic frontier model that is a function of a plethora of individual characteristics. Our model nests both parametric and non-parametric estimations and accommodates underlying heterogeneity across individuals, persistence and thresholds effects. Our results show that the Bayesian dynamic latent model provides good fit and report the whole underlying density function of happiness efficiency per individual. We confirm the doctrine of Rayo and Becker (2007) that happiness efficiency would vary across individuals. We reveal that the average happiness efficiency is close to 70% and as such a typical British individual would need to cover some ground to get to the optimal happiness efficiency frontier so as to maximise her happiness. Results show that marital status (married vs single), employment and income would enhance happiness, though we demonstrate that money does not always buy happiness. Key is also to have certain personality traits, as being agreeable and extravert would increase happiness.

This study shows that there are gains in happiness efficiency to be made. Specifically, we quantify that those gains could be up to 30%. Proposed measures of happiness indexes (Tsurumi and Manage 2017; Diener, 2000; Cordero et al., 2017, but also World Bank and OECD happiness indexes) are mainly focusing on the importance of availability of social and economic resources. Our study postulates a novel model to quantify happiness and provides evidence of happiness efficiency. The results show that individuals with existing resources could reach a higher frontier of happiness by improving the efficiency with which they use those resources.

However, an interesting finding of our study refers to the persistence of happiness efficiency. Such persistence should be seen in light of threshold effects in happiness efficiency. As thresholds are outside the control of individuals, achieving maximum happiness and driving to higher levels of happiness efficiency might not be feasible. Previous research (Rayo and Becker, 2007; Dolan and Kahneman, 2008; Kahneman and Krueger, 2006; Kahneman, 2003) argue that, when it comes to happiness, individual's decision making within the framework of a happiness function could be also affected by factors such as luck. Such factors as unobservable are rather difficult to measure. Our proposed Bayesian latent modelling of happiness and happiness efficiency provides a social operational research framework to treat unobserved factors and provide social quantitative information that is useful to academics and policy makers alike.

REFERENCES

Albert, J. H. and S. Chib. 1993. Bayesian Analysis of Binary and Polychotomous Response Data. Journal of the American Statistical Association, 88:669-79.
Albert, J. H. and S. Chib (2001). Sequential ordinal modeling with applications to survival data. Biometrics, 57, 829–836.
Andrieu, C. and Roberts, G. O. (2009). The pseudo-marginal approach for efficient Monte Carlo computations. The Annals of Statistics, 37(2):697–9725.
Anand, P., G. Hunter, I. Carter, K. Dowding, F. Guala, M. Van Hees (2009). The development of capability indicators, Journal of Human Development and Capabilities, 10 (1), pp. 125–152.
Anand, Paul, Jaya Krishnakumar, Ngoc Bich Tran (2011). Measuring welfare: Latent variable models for happiness and capabilities in the presence of unobservable heterogeneity. Journal of Public Economics, Volume 95, Issues 3–4, April, Pages 205-215.
Barberio Enzo Mariano, Vinicius Amorim Sobreiro, Daisy Aparecida do Nascimento Rebelatto, (2015). Human development and data envelopment analysis: A structured literature review, Omega, Volume 54, 2015, Pages 33-49.
Binder, M., and Broekel, T. 2012. Happiness no matter the cost? An examination on how efficiently individuals reach their happiness levels. Journal of Happiness Studies, 13(4), 621-645.
Brennan, Shae, Carla Haelermans, John Ruggiero (2014). Nonparametric estimation of education productivity incorporating nondiscretionary inputs with an application to Dutch schools, European Journal of Operational Research, Volume 234, Issue 3, 1 May 2014, Pages 809-818.
Cappe, O.; Godsill, S.; Moulines, E.; (2007). An overview of existing methods and recent advances in sequential Monte Carlo. Proceedings of IEEE, 95 (5): 899.
Cheng, T. C., Powdthavee, N., and Oswald, A. J. (2015). Longitudinal Evidence for a Midlife Nadir in Human Well-being: Results from Four Data Sets. The Economic Journal.
Chen, M.-H., and Dey, D.K. (2000). Bayesian Analysis for Correlated Ordinal Data Models, in Generalized Linear Models: A Bayesian Perspective, eds. D.K. Dey, S. Ghosh 20and B.K. Mallick, pp. 135-162, New York: Marcel Dekker.
Clark, A. E., Diener, E., Georgellis, Y., and Lucas, R. E. (2008). Lags and leads in life satisfaction: a test of the baseline hypothesis. The Economic Journal, 118(529).
Cordero, José Manuel, Javier Salinas-Jiménez, M Mar Salinas-Jiménez, (2017). Exploring factors affecting the level of happiness across countries: A conditional robust nonparametric frontier analysis, European Journal of Operational Research, Volume 256, Issue 2, 16 January, Pages 663-672.
Cornwell, C., P. Schmidt, and R. C. Sickles. 1990. Production frontiers with cross-sectional and time-series variation in efficiency levels. Journal of Econometrics46:185–200.
Conzo, P., Aassve, A., Fuochi, G., & Mencarini, L.(2017). The cultural foundations of happiness, Journal of Economic Psychology, 62, 268-283.
Debnath, R.M., & Shankar, R. (2014). Does good governance enhance happiness: A cross nation study. Social Indicators Research, 116(1), 235-253.
Despotis, D., (2005). Measuring human development via data envelopment analysis: the case of Asia and the Pacific, Omega, 33, October, pp. 385-390.
Doucet, A., Godsill, S., Andrieu, C. (2000). On sequential Monte Carlo sampling methods for Bayesian filtering. Statistics and Computing, 10 (3): 197–208.
Doucet, A., de Freitas, J., and Gordon, N. (eds.) (2001). Sequential Monte Carlo Methods in Practice. Springer, New York.
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 (pp. 213–229). New York: Russell Sage Foundation.
Diener, E. (2009). Subjective well-being, The science of well-being, pages 11-58, The Collected Works of Ed Diener, Springer.

Diener, E., S. Oishi, and L. Tay (2018). Advances in subjective well-being research, Nature Human Behaviour, 2, pages 253–260.
Di Tella, R., MacCulloch, R. J., and Oswald, A. J. (2001). Preferences over inflation and unemployment: Evidence from surveys of happiness. American Economic Review, 335-341.
Di Tella, Rafael, Robert Mac Culloch and Andrew J. Oswald. (2003) The Macroeconomics of Happiness. Review of Economics and Statistics, 85 :809-827.
Di Tella, Rafael, John Haisken-De New, and Robert MacCulloch. (2010). Happiness Adaptation to Income and to Status in an Individual Panel. Journal of Economic Behavior & Organization, 76:834-852.
Durante, D., Paganin, S., Scarpa, B. and Dunson, D. B. (2017), Bayesian modelling of networks in complex business intelligence problems. Journal Royal Statistical Society C, 66: 555–580.
Easterlin, R. A., McVey, L. A., Switek, M., Sawangfa, O., & Zweig, J. S. (2010). The happiness–income paradox revisited. Proceedings of the National Academy of Sciences, 107(52), 22463-22468.
Ferrer-i-Carbonell, A., and Frijters, P. (2004). How important is methodology for the estimates of the determinants of happiness? The Economic Journal, 114(497), 641-659.
Fearnhead, P., Papaspiliopoulos, O., and Roberts, G. O. (2008). Particle filters for partially observed diffusions. Journal of the Royal Statistical Society, Series B, 70:1–28.
Frey, B.S., & Stutzer, A. (2012). The use of happiness research for public policy. Social Choice and Welfare, 38(4), 659-674.
Fleurbaey Marc and Hannes Schwandt (2015). Do People Seek to Maximize Their Subjective Well-Being? CEP Discussion Paper No 1391, November.
Graham, L. and Oswald, A. J. (2010). Hedonic capital, adaptation and resilience. Journal of Economic Behavior and Organization, 76(2): 372-384.
Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith (Eds.), Bayesian Statistics 4, pp. 169–193. Oxford: Oxford University Press.
Gordon, N. J.; Salmond, D. J. and Smith, A. F. M. (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation. IEEE Proceedings F on Radar and Signal Processing, 140 (2): 107–113.
Gosling, S.D. and W.B.S. Rentfrow, (2003). A very brief measure of the Big-Five personality domains. Journal of Research in Personality, 37, pp. 504–528
Hasegawa, H. (2009). Bayesian Dynamic Panel-Ordered Probit Model and Its Application to Subjective Well-Being, Communications in Statistics: Simulation and Computation, 38(6): 1321-1347.
Jeliazkov, I., J. Graves, and M. Kutzbach (2009). Fitting and comparison of models for multivariate ordinal outcomes. In S. Chib, G. Koop, W. Griffiths, and D. Terrell (Eds.), Advances in Econometrics: Bayesian Econometrics, Volume 23, pp. 115–156. Bingley: Emerald.
Johnson, V. E. and J. H. Albert (1999). Ordinal Data Modeling. New York: Springer.
Kahneman, D. (2003). Maps of bounded rationality: psychology for behavioral economics. American Economic Review, 1449-1475.
Kahneman, D., & Krueger, A. B. (2006). Developments in the measurement of subjective well-being. The Journal of Economic Perspectives, 20(1), 3-24.
Krishnakumar, J. (2007). Going beyond functionings to capabilities: an econometric model to explain and estimate capabilities. Journal of Human Development, 8 (1), pp. 39–63.
Luque, M. O.D. Marcenaro-Gutiérrez, L.A. López-Agudo, (2015). On the potential balance among compulsory education outcomes through econometric and multiobjective programming analysis, European Journal of Operational Research, Volume 241, Issue 2, 1 March, Pages 527-540
Lucas, R. E., & Diener, E. (2015). Personality and subjective well-being: Current issues and controversies. In M. Mikulincer, P. R. Shaver, M. L. Cooper, & R. J. Larsen (Eds.), APA handbooks in psychology. APA handbook of personality and social psychology, Vol. 4. Personality processes and individual differences (pp. 577-599). Washington, DC, US: American Psychological Association.

Liu, J., West, M. (2001). Combined parameter and state estimation in simulation-based filtering. In: Doucet, A, de Freitas, J., Gordon, N.J. (eds.), Sequential Monte Carlo Methods in Practice. Springer-Verlag, New York.
Marcenaro-Gutierrez, O.D., M. Luque, and F. Ruiz, (2010). An application of multiobjective programming to the study of workers' satisfaction in the Spanish labour market, European Journal of Operational Research, Volume 203, Issue 2, 1 June, Pages 430-443.
Mizobuchi, H. Measuring Socio-economic Factors and Sensitivity of Happiness. (2017) Journal of Happiness Studies, 18 (2), pp. 463-504.
Nemeth, C., Sherlock, C., Fearnhead, P. (2014). Particle Metropolis adjusted Langevin algorithms. Working Paper.
Pitt, M.K.; Shephard, N. (1999). Filtering Via Simulation: Auxiliary Particle Filters. Journal of the American Statistical Association, Vol. 94, No. 446) 94 (446): 590–591.
Pitt, M. K., Silva, R., Giordani, P., and Kohn, R. (2012). On some properties of Markov chain Monte Carlo simulation methods based on the particle filter. Journal of Econometrics, 171(2):134–151.
Poyiadjis, G., Doucet, A., and Singh, S. S. (2011). Particle approximations of the score and observed information matrix in state space models with application to parameter estimation. Biometrika, 98(1):65–80.
Powdthavee, N., Lekfuangfu, W. N., and Wooden, M. (2015). What's the good of education on our overall quality of life? A simultaneous equation model of education and life satisfaction for Australia. Journal of Behavioral and Experimental Economics, 54, 10-21.
Rayo, L. & Becker, G. S. (2007). Evolutionary efficiency and happiness. Journal of Political Economy, 115(2): 302-337.
Ristic, B.; Arulampalam, S.; Gordon, N. (2004). Beyond the Kalman Filter: Particle Filters for Tracking Applications. Artech House.
Roberts, G. O., Gelman, A., and Gilks, W. (1997). Weak Convergence and Optimal Scaling of the Random Walk Metropolis Algorithms. The Annals of Applied Probability, 7(1):110–120.
Roberts, G. O. and Rosenthal, J. S. (1998). Optimal scaling of discrete approximations to Langevin diffusions. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 60(1):255–268.
Sewell, D. K. and Chen, Y. (2015). Analysis of the formation of the structure of social networks by using latent space models for ranked dynamic networks. Journal Royal Statistical Society C, 64: 611–633.
Tsurumi, T., Managi, S. Monetary Valuations of Life Conditions in a Consistent Framework: The Life Satisfaction Approach. (2017) Journal of Happiness Studies, 18 (5), pp. 1275-1303.
White, L., Burger, K. and Yearworth, M. (2016). Understanding behaviour in problem structuring methods interventions with activity theory, European Journal of Operational Research, 249, 3, 983-1004.
White, Leroy (2016). Behavioural operational research: towards a framework for understanding behaviour in or interventions, European Journal of Operational Research, 249, 3, 827-841.
White, Leroy (2018). A Cook's tour: Towards a framework for measuring the social impact of social purpose organisations, European Journal of Operational Research, Volume 268, Issue 3, Pages 784-797.

APPENDIX

A1. Markov Chain Monte Carlo

We use a Girolami and Calderhead (2012) algorithm to update draws for θ. The algorithm uses local information about both the gradient and the Hessian of the log-posterior conditional of θ at the existing draw. A Metropolis test is again used for accepting the candidate, but the GC algorithm moves considerably faster relative to our naive scheme previously described. The GC algorithm is started at the first stage GMM estimator and MCMC is run until convergence. Depending on the model and the subsample this takes 1,000 to 3,000 iterations. For safety we run 10,000 iterations. Then we run another 50,000 MCMC iterations to obtain final results for posterior moments and densities of parameters and functions of interest. It has been found that the GC algorithm performs vastly superior relative to the standard MH algorithm and autocorrelations are much smaller.

Suppose $L(q) = \log p(q \mid X)$ is used to denote for simplicity the log posterior of θ. Moreover, define

$$G(q) = \text{est.cov} \frac{¶}{¶q} \log p(X \mid q) \quad \text{(B.1)}$$

the empirical counterpart of

$$G_o(q) = - E_{Y|q} \frac{¶^2}{¶q ¶q¢} \log p(X \mid q) \quad \text{(B.2)}$$

The Langevin diffusion is given by the following stochastic differential equation:

$$dq(t) = \tfrac{1}{2} \tilde{\square}_q L\{q(t)\}\, dt + dB(t), \quad \text{(B.3)}$$

where

$$\tilde{\square}_q L\{q(t)\} = - \mathbf{G}^{-1}\{q(t)\} \times \tilde{\square}_q L\{oldsymbol q(t)\} \quad \text{(B.4)}$$

is the so called ‘*natural gradient*’ of the Riemann manifold generated by the log posterior. The elements of the Brownian motion are

$$\mathbf{G}^{-1}\{\theta(t)\} d\mathbf{B}_i(t) = \mid \mathbf{G}\{\theta(t)\} \mid^{-1/2} \sum_{j=1}^{K_\beta} \tfrac{\partial}{\partial \theta}\left[mbolG^{-1}\{\theta(t)\}_{ij} \mid \mathbf{G}\{\theta(t)\} \mid^{1/2} \right] dt \quad \text{(B.5)}$$

$$+\left[\sqrt{\mathbf{G}\{\theta(t)\}} d\mathbf{B}(t) \right]_i$$

The discrete form of the stochastic differential equation provides a proposal as follows:

$$\tilde{q}_i = q_i^o + \tfrac{e^2}{2} \left(\mathbf{G}^{-1}(q^o) \square_q L(q^o) \right)_i - e^2 \square_{j=1}^{K_q} \left(\mathbf{G}^{-1}(q^o) \frac{\partial \mathbf{G}(q^o)}{\partial q_j} \mathbf{G}^{-1}(q^o) \right)_{ij}$$

$$+ \tfrac{e^2}{2} å_{j=1}^{K_q} \left\{ \mathbf{G}^{-1}(q^o) \right\}_{ij} \text{tr} \left\{ \mathbf{G}^{-1}(q^o) \frac{¶\mathbf{G}(q^o)}{¶q_j} \right\} + \left\{ e \sqrt{\mathbf{G}^{-1}(q^o)} x^o \right\}_i$$

$$= \mu\left(\theta^o, \varepsilon\right)_i + \left\{\varepsilon\sqrt{\mathbf{G}^{-1}\left(\theta^o\right)}\xi^o\right\}_i$$

where β^o is the current draw.

The proposal density is

$$q\left(\tilde{\theta} \mid \theta^o\right) = N_{K_\theta}\left(\tilde{\theta}, \varepsilon^2 \mathbf{G}^{-1}\left(\theta^o\right)\right) \qquad \text{(B.6)}$$

and convergence to the invariant distribution is ensured by using the standard form Metropolis-Hastings probability

$$\min\left\{1, \frac{p\left(\tilde{\theta} \mid \cdot, Y\right) q\left(\theta^o \mid \tilde{\theta}\right)}{p\left(\theta^o \mid \cdot, Y\right) q\left(\tilde{\theta} \mid \mathbf{a}^o\right)}\right\} \qquad \text{(B.7)}$$

A2. Particle Filtering

The particle filter methodology can be applied to state space models of the general form

$$y_T \sim p(y_t \mid x_t),\, s_t \sim p(s_t \mid s_{t-1}), \quad \text{(B.8)}$$

where s_t is a state variable.

For general introductions see Gordon et al. (1993), Doucet et al. (2000, 2001), Pitt and Shephard (1999) and Ristic et al. (2001).

Given the data Y_t the posterior distribution $p(s_t \mid Y_t)$ can be approximated by a set of (auxiliary) particles $\left\{s_t^{(i)}, i = 1,...,.N\right\}$ with probability weights $\left\{w_t^{(i)}, i = 1,..., N\right\}$ where $\sum_{i=1}^N w_t^{(i)} = 1$. The predictive density can be approximated by

$$p(s_{t+1} \mid Y_t) = \int p(s_{t+1} \mid s_t)\, p(s_t \mid Y_t)\, ds_t \simeq \sum_{i=1}^{N} p(s_{t+1} \mid s_t^{(i)}) w_t^{(i)} \quad \text{(B.9)}$$

and the final approximation for the filtering density is

$$p(s_{t+1} \mid Y_t) \propto p(y_{t+1} \mid s_{t+1})\, p(s_{t+1} \mid Y_t) \simeq p(y_{t+1} \mid s_{t+1}) \sum_{i=1}^{N} p(s_{t+1} \mid s_t^{(i)}) w_t^{(i)} \quad \text{(B.10)}$$

The basic mechanism of particle filtering rests on propagating $\left\{s_t^{(i)}, w_t^{(i)}, i = 1,..., N\right\}$ to the next step, viz. $\left\{s_{t+1}^{(i)}, w_{t+1}^{(i)}, i = 1,..., N\right\}$ but this often suffers from the weight degeneracy problem. If parameters $\theta \in \Theta \in \Re^k$ are available, as is often the case, we follow Liu and West (2001) parameter learning takes place via a mixture of multivariate normals:

$$p(\theta | Y_t) \simeq \sum_{i=1}^{N} w_t^{(i)} N(\theta | a\theta_t^{(i)} + (1-a)\bar{\theta}_t, b^2 V_t) \quad \text{(B.11)}$$

where $\bar{\theta}_t = \sum_{i=1}^{N} w_t^{(i)} \theta_t^{(i)}$, and $V_t = \sum_{i=1}^{N} w_t^{(i)} (\theta_t^{(i)} - \bar{\theta}_t)(\theta_t^{(i)} - \bar{\theta}_t)'$. The constants a and b are related to shrinkage and are determined via a discount factor $\delta \in (0,1)$ as $a = (1-b^2)^{1/2}$ and $b^2 = 1 - [(3\delta - 1)/2\delta]^2$. See also Casarin and Martin (2007).

Andrieu and Roberts (2009), and Pitt et al. (2012) provide the Particle Metropolis-Hastimgs (PMCMC) technique which uses an unbiased estimator of the likelihood function $\hat{p}_N(Y | \theta)$ as $p(Y | \theta)$ is often not available in closed form.

Given the current state of the parameter $\theta^{(j)}$ and the current estimate of the likelihood, say $L^j = \hat{p}_N(Y | \theta^{(j)})$, a candidate θ^c is drawn from $q(\theta^c | \theta^{(j)})$ yielding $L^c = \hat{p}_N(Y | \theta^c)$. Then, we set $\theta^{(j+1)} = \theta^c$ with the Metropolis - Hastings probability:

$$A = \min\left\{1, \frac{p(\theta^c) L^c}{p(\theta^{(j)}) L^j} \frac{q(\theta^{(j)} | \theta^c}{q(\theta^c | \theta^{(j)})}\right\}, \quad \text{(B.12)}$$

otherwise, we repeat the current draws: $\left\{\theta^{(j+1)}, L^{j+1}\right\} = \left\{\theta^{(j)}, L^j\right\}$.

A3. Particle Metropolis adjusted Langevin filters

Nemeth, Sherlock and Fearnhead (2014) provide a particle version of a Metropolis adjusted Langevin algorithm (MALA). See also Poyiadjis et al. (2011). In Sequential Monte Carlo we are interested in approximating $p(s_t | Y_{1:t}, \theta)$. Given that

$$p(s_t | Y_{1:t}, \theta) \propto g(y_t | x_t, \theta) \int f(s_t | s_{t-1}, \theta) p(s_{t-1} | y_{1:t-1}, \theta) ds_{t-1}, \text{ (B.13)}$$

where $p(s_{t-1} | y_{1:t-1}, \theta)$ is the posterior as of time $t-1$. If at time $t-1$ we have a set of particles $\left\{s_{t-1}^i, i = 1, \ldots, N\right\}$ and weights $\left\{w_{t-1}^i, i = 1, \ldots. N\right\}$, which form a discrete approximation for $p(s_{t-1} | y_{1:t-1}, \theta)$ then we have the approximation:

$$\hat{p}(s_{t-1} | y_{1:t-1}, \theta) \propto \sum_{i=1}^{N} w_{t-1}^i f(s_t | s_{t-1}^i, \theta) \qquad \text{(B.14)}$$

See Doucet et al. (2000) and Cappe at al. (2007) for reviews. From (A.14) Fernhead et al. (2008) make the important observation that the joint probability of sampling particle s_{t-1}^i and state s_t is:

$$\omega_t = \frac{w_{t-1}^i g(y_t | s_t, \theta) f(s | s_{t-1}^i, \theta)}{\xi_t^i q(s_t | s_{t-1}^i, y_t, \theta)}, \quad \text{(B.15)}$$

where $q(s_t \mid s_{t-1}^i, y_t, \theta)$ is a density function amenable to simulation and

$$\chi_t^i q(s_t \mid s_{t-1}^i, y_t, \theta) \simeq c g(y_t \mid s_t, \theta) f(s_t \mid s_{t-1}^i, \theta), \quad \text{(B.16)}$$

and c is the normalizing constant in (B.13).

In the MALA algorithm of Roberts and Rosenthal (1998)[11] we form a proposal:

$$\theta^c = \theta^{(s)} + \lambda z + \tfrac{\lambda^2}{2} \nabla \log p(\theta^{(s)} \mid Y_{1:T}), \quad \text{(B.17)}$$

where $z \sim N(0, I)$ which should result in larger jumps and better mixing properties, plus lower autocorrelations for a certain scale parameter λ. Acceptance probabilities are:

$$a(\theta^c \mid \theta^{(s)}) = \min\left\{1, \frac{p(Y_{1:T} \mid \theta^c) q(\theta^{(s)} \mid \theta^c)}{p(Y_{1:T} \mid \theta^{(s)}) q(\theta^c \mid \theta^{(s)})}\right\}, \quad \text{(B.18)}$$

Using particle filtering it is possible to create an approximation of the score vector using Fisher's identity:

$$\nabla \log p(Y_{1:T} \mid \theta) = E\left[\nabla \log p(s_{1:T}, Y_{1:T} \mid \theta) \mid Y_{1:T}, \theta\right], \quad \text{(B.19)}$$

which corresponds to the expectation of

$$\nabla \log p(s_{1:T}, Y_{1:T} \mid \theta) = \nabla \log p(\mid s_{1:T-1}, Y_{1:T-1} \mid \theta) + \nabla \log g(y_T \mid s_T, \theta) + \nabla \log f(s_T \mid s\mid_{T-1}, \theta),$$

over the path $s_{1:T}$.

The particle approximation to the score vector results from replacing $p(s_{1:T} \mid Y_{1:T}, \theta)$ with a particle approximation $\hat{p}(s_{1:T} \mid Y_{1:T}, \theta)$. With particle i at time t-1 we can associate a value $\alpha_{t-1}^i = \nabla \log p(s_{1:t-1}^i, Y_{1:t-1} \mid \theta)$, which can be updated recursively. As we sample κ_i in the APF (the index of particle at time $t-1$ that is propagated to produce the i^{th} particle at time t) we have the update:

$$\alpha_t^i = a_{t-1}^{\kappa_i} + \nabla \log g(y_t \mid s_t^i, \theta) + \nabla \log f(s_t^i \mid s_{t-1}^i, \theta). \quad \text{(B.20)}$$

To avoid problems with increasing variance of the score estimate $\nabla \log p(Y_{1:t} \mid \theta)$ we can use the approximation:

$$\alpha_{t-1}^i \sim N(m_{t-1}^i, V_{t-1}). \quad \text{(B.21)}$$

The mean is obtained by shrinking α_{t-1}^i towards the mean of α_{t-1} as follows:

$$m_{t-1}^i = \varpi \alpha_{t-1}^i + (1-\varpi) \sum_{i=1}^{N} w_{t-1}^i \alpha_{t-1}^i, \quad \text{(B.22)}$$

[11] The benefit of MALA over Random-Walk-Metropolis arises when the number of parameters n is large. This happens because the scaling parameter λ is $O(n^{-1/2})$ for Random-Walk-Metropolis but it is $O(n^{-1/6})$ for MALA, see Roberts et al. (1997) and Roberts and Rosenthal (1998)

where $\varpi \in (0,1)$ is a shrinkage parameter.

Using Rao-Blackwellization one can avoid sampling α_t^i and instead use the following recursion for the means:

$$m_t^i = \varpi m_{t-1}^{\kappa_i} + (1-\varpi)\sum_{i=1}^{N} w_{t-1}^i m_{t-1}^i + \nabla \log g(y_t \mid s_t^i, \theta) + \nabla \log f(s_t^i \mid s_{t-1}^{\kappa_i}, \theta), \quad \text{(B.23)}$$

which yields the final score estimate:

$$\nabla \log \hat{p}(Y_{1:t} \mid \theta) = \sum_{i=1}^{N} w_t^i m_t^i, \quad \text{(A.24)}$$

As a rule of thumb Nemeth, Sherlock and Fearnhead (2014) suggest taking $\delta = 0.95$. Furthermore, they show the important result that the algorithm should be tuned to the asymptotically optimal acceptance rate of 15.47% and the number of particles must be selected so that the variance of the estimated log-posterior is about 3. Additionally, if measures are not taken to control the error in the variance of the score vector there is no gain over a simple random walk proposal.

Of course, the marginal likelihood is

$$p(Y_{1:T} \mid \theta) = p(y_1 \mid \theta)\prod_{t=2}^{T} p(y_t \mid Y_{1:t-1}, \theta), \quad \text{(B.25)}$$

where

$$p(y_t \mid Y_{1:t-1}, \theta) = \int g(y_t \mid s_t) \int f(s_t \mid s_{t-1}, \theta) p(s_{t-1} \mid Y_{1:T-1}, \theta) ds_{t-1} ds_t, \quad \text{(B.26)}$$

provides, in explicit form, the predictive likelihood.