Latent growth curve modeling of psychological well-being trajectories

M. Fátima Salgueiro[1, *], Joana Malta[2]

[1]Instituto Universitário de Lisboa (ISCTE-IUL), Department of Quantitative Methods and Business Research Unit, Lisbon, Portugal
[2]Statistics Portugal, Lisbon, Portugal

Email address:

fatima.salgueiro@iscte.pt (M. F. Salgueiro), joana.malta@ine.pt (J. Malta)

Abstract: This paper proposes modeling trajectories of psychological well-being using latent growth curve models (LGCMs). The psychometric scale of the General Health Questionnaire-12 (GHQ-12) is considered. Data from the British Household Panel Survey (BHPS), from years 2003 to 2006 are used. In 1991 Graetz proposed the GHQ-12 as a multidimensional scale, containing three distinct dimensions: anxiety and depression, social dysfunction and loss of confidence. Using such scale, this paper compares a second-order LGCM for the trajectories of a latent factor (measured by these three dimensions) with a LGCM for the trajectories of an overall sum score. Conditional LGCMs are then fitted; sex, age group and perceived health status are considered as the explanatory variables of the growth trajectories. Results show that the model which considers the three dimensions of subjective well-being has a larger explaining capability than the one utilizing the subjective well-being score.

Keywords: British Household Panel Survey, General Health Questionnaire 12, Latent Growth Curve Model, Longitudinal Data Analysis

1. Introduction

The General Health Questionnaire (GHQ) is a psychometric scale, initially composed of 60 items, developed by Goldberg in 1970 to detect psychological dysfunctions in a non clinical environment. Even though the GHQ was developed "as a screening instrument for psychiatric illness (…), it is often used as an indicator of psychological well-being" [1]. Due to the difficulty in applying such a large questionnaire, several other questionnaires were developed from the original 60-item GHQ. One of the most common is the GHQ-12, containing only 12 items for measurement of psychological well-being. One should note that most authors refer to the GHQ-12 as a scale for measuring subjective well-being (SWB). Though the GHQ-12 was initially validated under the assumption that it measured a single dimension, however several authors argue that the scale is multidimensional. Indeed, [2] claims that the GHQ-12 assesses psychological morbidity and is composed of three underlying factors: anxiety and depression, social dysfunction and loss of confidence.

The current paper proposes measuring (and,

consequently, modeling) SWB in two different ways: i) as a one-dimensional concept, using a SWB score derived from the sum of the individual responses to the 12 items of the GHQ-12 scale; ii) as a three-dimensional concept, using a psychological morbidity (PM) latent factor, measured by three score variables: anxiety and depression (AD), social dysfunction (SD) and loss of confidence (LC). The derived scores of these three variables were also computed as sums of item responses.

Two unconditional latent growth curve models (LGCMs) are first fitted to describe the growth trajectories of the SWB score and of the latent PM factor. Finally, sex, age group and health status are included as time-invariant covariates in the conditional LGCMs proposed to explain growth trajectories.

2. The BHPS Data Under Analysis

Data from the British Household Panel Survey (BHPS), developed by the Longitudinal Studies Centre of the Economic and Social Research Council (ESRC) and the Institute for Social and Economic Research of the University of Essex, are used. The BHPS is a panel of

British households, conducted since 1991, and its main objective is to gather information intended to "further our understanding of social and economic change at the individual and household level in Britain, to identify, model and forecast such changes, their causes and consequences in relation to a range of socio-economic variables" [3].

The variables in the GHQ-12 were obtained through a self completion questionnaire, answered by respondents aged 16 or older. Variables were measured on a Likert-type scale, from 0 to 3. Data for four time periods,

corresponding to years 2003 to 2006, are used. The final working sample includes 4562 individuals, 55.8% of whom are women. The mean age of the respondents equals 47.86 years, with a standard deviation of 17.78 years. Regarding perceived health status, almost half of the respondents perceive themselves as having a good health (49.9%), and only 20.8% consider themselves to have an excellent health status.

Table 1 lists the complete wording of the 12 questions of the GHQ-12 scale used in the BHPS.

Table 1. The GHQ-12 scale used in the BHPS

Variable - Label	Question	Likert scale	Construct
GHQ1 - loss of sleep	Have you recently.... lost much sleep over worry?	4 points: Not at all; No more than usual; Rather more than usual; Much more than usual	AD
GHQ2 - constantly under strain	Have you recently.... felt constantly under strain?		AD
GHQ3 - problem overcoming difficulties	Have you recently.... felt you couldn't overcome your difficulties?		AD
GHQ4 - unhappy or depressed	Have you recently.... been feeling unhappy or depressed?		AD
GHQ5 - concentration	Have you recently.... been able to concentrate on whatever you're doing?	4 points: More than usual; Same as usual; Less so; Much less	SD
GHQ6 - playing a useful role	Have you recently.... felt that you were playing a useful part in things?		SD
GHQ7 - capable of making decisions	Have you recently.... felt capable of making decisions about things?		SD
GHQ8 - enjoy day-to-day activities	Have you recently.... been able to enjoy your normal day-to- day activities?		SD
GHQ9 - ability to face problems	Have you recently.... been able to face up to problems?		SD
GHQ10 - general happiness	Have you recently.... been feeling reasonably happy, all things considered?		SD
GHQ11 - losing confidence	Have you recently.... been losing confidence in yourself?	4 points: Not at all; No more than usual; Rather more than usual; Much more than usual	LC
GHQ12 - believe in self-worth	Have you recently.... been thinking of yourself as a worthless person?		LC

3. The Two LGCMs Proposed to Model the Growth Trajectories of the SWB Score and of the PM factor
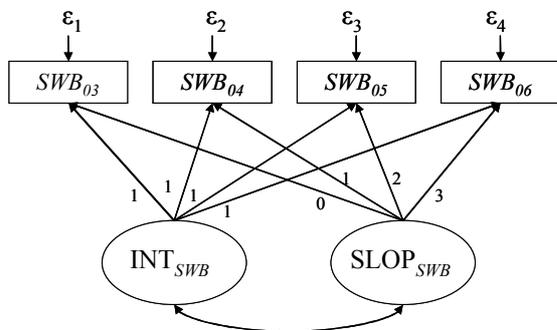
Figure 1. Unconditional LGCM for the 2003-2006 trajectory of the SWB score variable.

Fig. 1 shows the path diagram of the unconditional LGCM proposed to describe growth trajectories of the

SWB score. Recall this score was obtained by summing the responses of each individual to the 12 items of the GHQ-12 scale, assuming that SWB is a one-dimensional concept. The repeated measures of the SWB score variable are considered as continuous variables. Linear growth is assumed, with the conventional codes for the slope factor loadings: 0, 1, 2, 3. For a detailed explanation on LGCMs see [4] and [5]. The parameters of special interest in this model are the means, variances and covariance of the intercept and slope random factors for the SWB score trajectories.

The general form of the unconditional LGCM is given by

$$y_{it} = \alpha_i + \lambda_i \beta_i + \varepsilon_{it} , \qquad (1)$$

where y_{it} is the repeated measure under analysis (in our case the SWB score) for individual i at time t; α_i denotes the intercept for individual i and β_i denotes the slope of the latent trajectory for individual i. The random effects (intercepts and slopes) are given by

$$\alpha_i = \mu_\alpha + \zeta_{\alpha i} \qquad (2)$$

$$\beta_i = \mu_\beta + \zeta_{\beta i} \qquad (3)$$

ζ_{α_i} and ζ_{β_i} represent, respectively, the disturbances around the intercept and around the mean.

The term λ_t represents the passing of time. In our model a linear growth trajectory is considered and $\lambda_t = t - 1$, so that the expected value for the intercept represents the mean trajectory at the initial time point (when $\lambda_t = 0$). The disturbances ζ_{α_i} and ζ_{β_i} have zero mean, variances $\psi_{\alpha\alpha}$ and $\psi_{\beta\beta}$ and covariance $\psi_{\alpha\beta}$.

Fig. 2 represents the path diagram of the second-order LGCM proposed for the growth trajectory of the latent PM factor measured by three indicators. The repeated measures of the indicators AD, SD and LC are also considered as continuous outcomes, and were obtained as sums of item responses. A linear growth trajectory for the latent PM factor is assumed. Because it is plausible to expect some degree of dependency between repeated measures [6], covariances between the measurement errors of the repeated measures are freely estimated.
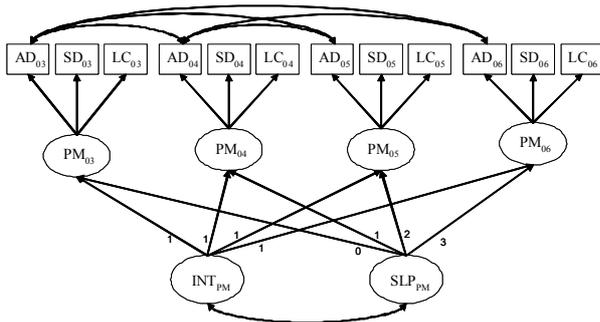
Figure 2. Unconditional LGCM for the PM factor, measured by three score variables: AD, SD and LC.

The general form of the unconditional LGCM with multiple indicators is given by

$$\eta_{it} = \alpha_i + \lambda_t \beta_i + \varepsilon_{it},$$

where η_{it} stands for the multiple indicator latent variable for individual i at moment t; ε_{it} represents the disturbance, which has zero mean and is not correlated with α_i, β_i or λ_t. To consider multiple indicators a third index has to be added: let $j = 1, 2, ..., J$ denote the indicators. Thus,

$$y_{jit} = v_{jt} + \Lambda_{jt} \eta_{it} + \upsilon_{jit}$$

Let v_{jt} denote the intercept for indicator j at the time t, Λ_{jt} represent the weight matrix for indicator j at time t, and υ_{jit} represent the disturbance from the estimated trajectory for the i-th case on indicator j at time t. Note that λ_t and Λ_{jt} denote different parameters: λ_t represents

the trajectory matrix (the shape of the growth trajectory); Λ_{jt} represents the estimated weight matrix of the factor loadings between the outcome variables and the latent variable η_{it}.

Up to this point we have been discussing unconditional latent growth curve models, i.e., models that do not take into account the potential causes for the observed individual variability around the mean trajectory. Including covariates in the model is a common approach to account for this variability, and thus explain individual heterogeneity.

Fig. 3 shows the path diagram of a LGCM with conditional growth. Two time-invariant covariates x_1 and x_2 are used to explain the growth trajectories of the SWB score. Covariates can be introduced in the model either as continuous or as dummy variables, as in a linear regression model.
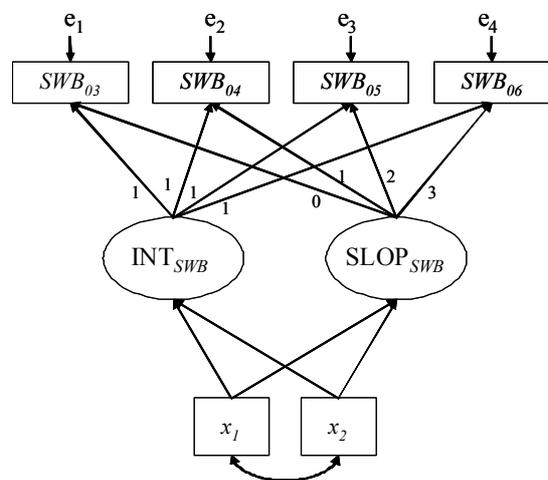
Figure 3. Conditional LGCM for the 2003-2006 growth trajectories of the SWB score with two time-invariant covariates x_1 and x_2

Equations (2) and (3) are now expanded to incorporate the regression effect of each covariate on the random intercept and slope. In a general case with M covariates the random effects can be obtained as

$$\alpha_i = \mu_\alpha + \gamma_{\alpha 1} x_{1i} + \gamma_{\alpha 2} x_{2i} + ... + \gamma_{\alpha M} x_{Mi} + \zeta_{\alpha i} \qquad (1)$$

$$\beta_i = \mu_\beta + \gamma_{\beta 1} x_{1i} + \gamma_{\beta 2} x_{2i} + ... + \gamma_{\beta M} x_{Mi} + \zeta_{\beta i}$$

Here, μ_α and μ_β represent the mean intercept and slope for all individual trajectories when all $x_1, x_2, ..., x_M$ are null. $x_1, x_2, ..., x_M$ denote the M considered covariates; $\gamma_{\alpha 1}, \gamma_{\alpha 2}, ..., \gamma_{\alpha M}$ and $\gamma_{\beta 1}, \gamma_{\beta 2}, ..., \gamma_{\beta M}$ are their corresponding regression coefficients. All covariates are considered as time invariant, assuming that their value is assumed to be constant over all time periods, varying only between individuals.

The variances of the disturbances $\zeta_{\alpha i}$ and $\zeta_{\beta i}$ now represent the conditional variances of the random effects. The assumption that disturbances $\zeta_{\alpha i}$ and $\zeta_{\beta i}$ have zero

mean is still valid. Additionally, they are neither correlated with the estimated trajectory error ε_{it}, nor with the covariates $x_1, x_2, ..., x_M$.

In this study sex, age group and perceived health status are included in the conditional LGCMs as time-invariant covariates. Although it can be argued that health status should be considered as time-varying, descriptive statistics suggested there was little variation over the analyzed time period, thus justifying the use of the initial values of perceived health status as representative of individual state of health.

In order to assess goodness of fit, the following measures are used in this paper: the Tucker-Lewis Index (TLI), the Comparative Fit Index (CFI) and the Root Mean Square Error of Approximation (RMSEA). The Akaike Information Criterion and the Bayesian Information Criterion are used as goodness of fit measures for model comparison. One should note that these two criteria penalize complexity, favoring more parsimonious models. Indeed, [7] suggests the following thresholds for a reasonably good level of fit: TLI > 0.95; CFI > 0.95; RMSEA < 0.06.

4. Results

The LGCM parameterization for a latent trajectory of a single continuous outcome allows for freely estimating the mean of the intercept factor, which is interpreted as the mean value of the outcome variable at the initial time point, i.e., the mean SWB score in 2003. However, in a second-order LGCM for modeling the growth trajectory of a latent variable measured by several indicators, the mean of the intercept factor has to be constrained to zero and is no longer a parameter of interest.

Table 2 presents the estimates for the means, variances and covariances of the random intercept and slope factors, for the unconditional LGCMs fitted to describe the growth trajectories of the SWB score and the PM factor.

Table 2. *Estimates (and t-values) for the unconditional LGCMs for the SWB score and the PM factor*

Parameter	SWB score	PM factor
Mean of the Intercept	11.000	0.000
	(150.333)	----
Mean of the Slope	0.107	0.045
	(3.965)	(4.141)
Variance of the Intercept	14.779	2.306
	(25.983)	(21.852)
Variance of the Slope	0.611	0.086
	(6.262)	(5.273)
Covariance Intercept/Slope	-0.605	-0.073
	(-3.192)	(-2.287)

Recall that the SWB score was obtained by adding the answers of the 12 GHQ items, thus varying from 0 (the least distressed) to 36 (the most distressed). The mean of the intercept of the SWB score equals 11, suggesting individuals have relatively more positive SWB levels. The

estimate for the mean of the slope is significant and positive, both for the SWB score and for the PM factor, indicating mean levels of perceived SWB increase between 2003 and 2006, at an average annual growth rate of 0.107 for the SWB score and 0.045 for the PM factor.

The estimated variances of the intercept and slope factors are both significant, leading to the conclusion that individuals vary in their growth trajectories (both regarding initial values and rate of change). The covariance between the latent intercept and the latent slope is significant and negative, suggesting that higher levels of the subjective well-being score (or higher levels of psychological morbidity) are associated with a slower than average growth rate in SWB (score or PM factor).

The goodness of fit measures presented in Table 3 indicate that the LGCM for the SWB score has a better fit than the model for the PM factor. At this stage this result gives preference to the one-dimensional model, invalidating the hypothesis of gaining extra explaining capability by introducing three dimensions for PM.

Conditional LGCMs were fitted to account for individual variability in SWB growth trajectories. Three covariates were used: sex, age group and perceived health status. Because all three variables are categorical, dummy variables had to be created. The reference categories are male, age group 16-29 years old and excellent perceived health status (as indicated in Table 4).

Significant estimates (at the 5% level) for the regression coefficients of the random intercepts and slopes on the covariates, for the conditional LGCMs for the SWB score and for the PM factor are presented in Table 4. It is possible to conclude that women have significantly higher mean scores of SWB in 2003 than men, suggesting that women have, on average, higher levels of psychological distress. Individuals with ages between 30 and 44 years have, on average, higher levels of psychological morbidity than those aged 16-29. Individuals who report poorer health statuses also have higher levels of psychological morbidity than individuals who consider themselves as being in an excellent state of health. This asymmetry is stronger for individuals with a poor or very poor perceived health status.

Looking at the impact of the covariates on the slope factor, it is possible to conclude that individuals who do not have an excellent health status in 2003 have a significant decrease on their mean rate of change of SWB throughout the four-year period, when compared to those with an excellent health status.

As it can be observed in Table 4, conclusions regarding the impact of the covariates on the intercept and slope factors of the SWB score and of the PM factor are similar. However, two important differences can be noted. One difference is that those aged >=65 have significantly lower mean levels for the PM factor in 2003 than those aged 16-29. This is an interesting result, as it reinforces some of the conclusions drawn by [8] about the cohort and the U-shaped effect of age on psychological well-being. Also

using data from the BHPS, the author concludes that "well-being falls up to middle age, and then rises only little [where] the estimated turning point (…) is at or above the age of retirement", and also that "those born earlier report higher levels of well-being on the GHQ scale, independently of their current age". Another difference one can observe is the absence of a significant impact from the category of perceived health status "Good" on the slope of the PM factor growth trajectory. In fact, only the individuals who declared poor or very poor levels of perceived health have significantly slower rate of change on their psychological morbidity levels, when compared to individuals in an excellent state of health.

When comparing the model fit of the two conditional LGCMs (see Table 3), the AIC and BIC criteria indicate that the conditional LGCM for the SWB score has a better fit than the model for the PM factor, leading to the conclusion that a simpler model is enough to explain subjective well-being trajectories when using the GHQ-12. Yet, it should be stressed that a more complex model, based on the assumption that subjective well-being (when measured by the GHQ-12) is a multidimensional construct, highlights the impact of older age groups in explaining the trajectory of the latent PM factor.

Table 3. Goodness of fit measures for the unconditional and conditional LGCMs for the SWB score and the PM factor

	Unconditional LGCM		Conditional LGCM	
Measures	SWB score	PM factor	SWB score	PM factor
CFI	0.987	0.974	0.986	0.955
TLI	0.985	0.963	0.974	0.942
RMSEA	0.054	0.064	0.033	0.056
AIC	107993.15	197946.31	136475.68	226415.56
BIC	108050.98	198222.61	136610.62	226768.96

Table 4. Estimates (and t-values) for the regression coefficients of the random intercepts and slopes on the covariates, for the conditional LGCMs for the SWB score and the PM factor

	SWB score		PM factor	
Covariates (reference category)	Intercept	Slope	Intercept	Slope
Sex (Male):				
Female	1.314	0.054	0.601	0.016
	(9.510)	(0.966)	(10.387)	(0.750)
Age Group (16-29):				
30-44	0.647	-0.048	0.215	-0.020
	(3.094)	(-0.578)	(2.467)	(-0.598)
45-64	0.224	0.015	0.023	-0.005
	(1.132)	(0.196)	(0.278)	(-0.155)
>= 65	-0.424	0.086	-0.301	0.029
	(-1.875)	(0.971)	(-3.132)	(0.791)
Health Status (Excellent):				
Good	1.383	-0.149	0.586	-0.057
	(7.708)	(-2.108)	(7.834)	(-1.986)
Poor/very poor	4.237	-0.329	1.780	-0.126
	(21.283)	(-4.196)	(21.238)	(-3.988)

5. Discussion

The results of estimating unconditional LGCMs for subjective well-being led to the conclusion, shared by [9], that the gains in considering a more complex structure to understand SWB are very minor.

On the other hand, even if there is evidence that the conditional LGCM for the latent factor measuring psychological morbidity does not fit the data so well, we believe that such a model gives a clear understanding of how covariates impact and explain some of the observed heterogeneity in individual trajectories. Regarding goodness of fit and the comparison between the overall fit of the estimated models, we again note the used criteria favor a more parsimonious model. Therefore, it can be easily understood why a more complex model, with more parameters to estimate, has a slightly worse overall goodness of fit.

It should also be noted that in this paper a model considering the 12 ordered outcome variables of the GHQ-12 scale is not considered. A LGCM of this sort requires a more complex method for estimating parameters, and some fundamental assumptions and restrictions. Nevertheless, it could be an interesting way of analyzing the advantages arising from using the outcome variables as they were originally measured, posing the question of whether it would result in a deeper understanding of how subjective well-being evolves over time. Adding to this consideration, the need for a wider time range can also be pointed out, as subjective well-being is not only a subject with a slow change rate, but is influenced by a large number of independent elements [10].

Regarding the great number of possible variables which can influence subjective well-being, [11] states a wide list of possibilities. Also, [8] and [12] use a vast number of

covariates, other than sex, age or perceived heath status. They come to the conclusion that salary and labor situation strongly influence levels of subjective well-being.

Future research should not only consider models with ordinal outcomes and a wider time range, but also investigate the inclusion of additional covariates, to better the understanding of the external factors which can explain changes in individual perceptions of subjective well-being.

Acknowledgments

References

[1] Jones, A. M. and Widman J.: "Health, income and relative deprivation: Evidence from BHPS". Journal of Health Economics, vol. 27, 2008, pp. 308-324.

[2] Graetz, B.: "Multidimensional properties of the General Health Questionnaire". Social Psychiatry and Psychiatric Epidemiology, vol. 26, 1991, pp. 132-138.

[3] Taylor, M.F. (Ed.), with Brice, J., Buck, N., & Prentice-Lane, E.: British household panel survey user manual volume A: Introduction, technical report and appendices. Colchester: University of Essex, 2008.

[4] Bollen, K.A., Curran, P.J.: Latent Curve Models – A Structural Equation Perspective. John Wiley & Sons, New Jersey, 2006.

[5] Duncan, T.E., Duncan, S.C., Strycker, L.A.: An Introduction to Latent Variable Growth Curve Modeling - Concepts, Issues an Applications. 2nd ed. Lawrence Erlbaum Associates, Inc., New Jersey, 2006.

[6] Berrington, A. and Smith, P.W.F.: "An overview of methods for the analysis of panel data". ESRC National Centre for Research Methods, NCRM Methods Review Paper, 2006.

[7] Muthén, B.O.: Mplus Technical Appendices, Los Angeles: Muthén and Muthén, 1998-2004.

[8] Clark, A.: Born to be mild: cohort effects in subjective well-being. DELTA mimeo, 2002.

[9] Shevlin, M. and Adamson, G.: "Alternative factor models and factorial invariance of the GHQ-12: a large sample analysis using confirmatory factor analysis". American Psychological Association, vol. 17 (2), 2005, pp. 231-236.

[10] Pavot, W.: "The assessment of subjective well-being". In: Eid, Michael and Randy J. Larsen, Eds. The Science of Subjective Well-Being. New York: The Guilford Press, 2008, pp. 124-140.

[11] Diener, E.: "Myths in the science of happiness and directions for future research". In Eid, Michael and Randy J. Larsen, Eds. The Science of Subjective Well-Being, New York: The Guilford Press, 2008, pp. 493-514.

[12] Clark, A. and Oswald, A.: "The curved relationship between subjective well-being and age". Paris-Jourdan Sciences Economiques. Working paper nº 2006-9, Centre National de la Recherche Scientifique, 2006.