

Instituto Superior de Ciências do Trabalho e da Empresa



**MODELAÇÃO LONGITUDINAL
DO BEM-ESTAR PSICOLÓGICO COM
MODELOS DE TRAJECTÓRIA LATENTE**

Joana Veiga Malta Correia Guedes

Tese submetida como requisito parcial para obtenção do grau de

Mestre em Prospecção e Análise de Dados

Orientadora:

**Prof. Doutora Maria de Fátima Salgueiro, Prof. Auxiliar, ISCTE Business School,
Departamento de Métodos Quantitativos**

Junho 2009

Resumo

O trabalho que se apresenta modela as trajectórias do bem-estar psicológico a partir da escala psicométrica do *General Health Questionnaire – 12 (GHQ-12)*, utilizando dados do *British Household Panel Survey (BHPS)* relativos a quatro anos consecutivos, de 2003 a 2006. Utilizam-se os métodos de estimação desenvolvidos no âmbito do *Latent Growth Curve Models*, que estimam trajectórias latentes. A aproximação à modelação da escala do *GHQ-12* é feita a partir do pressuposto, assente na proposta de Graetz (1991), de que o *GHQ-12* é uma escala multidimensional que contém três dimensões distintas. Considera-se que estas três dimensões (ansiedade e depressão, disfunção social e perda de confiança) têm subjacente um factor de segunda ordem. De forma a comparar a trajectória de um indicador único com trajectórias dos três factores, no primeiro passo modelou-se a trajectória dos três factores considerando os *scores* das respostas dos indivíduos, obtidos a partir das somas das respostas aos respectivos itens. A comparação dos modelos permitiu concluir que o ganho em capacidade explicativa com um modelo considerando três dimensões é reduzido. O mesmo acontece quando modelamos as trajectórias considerando as escalas ordinais das variáveis de *input*. Inversamente, a estimação dos modelos de trajectória latente condicionada permitiu concluir que o modelo que considera as medidas repetidas ordinais e um factor de segunda ordem tem uma maior capacidade explicativa do que os modelos de LGCM condicionados quer do indicador único de bem-estar subjectivo, quer do factor latente medido pelos *scores* nas dimensões de ansiedade e depressão, disfunção social e perda de confiança.

Classificação JEL: C33, I19

Palavras-chave: *General Health Questionnaire*, dados longitudinais, *British Household Panel Survey*, modelos de trajectória latente

Abstract

This work intends to model the trajectories of psychological well-being based on the psychometric scale of the *General Health Questionnaire – 12 (GHQ-12)*, using four consecutive years' data from the *British Household Panel Survey (BHPS)*, from 2003 to 2006. In order to achieve this, estimation methods developed in the context of *Latent Growth Curve Models* – that estimate latent trajectories – were used. The approach to the *GHQ-12* scale modelling was carried out based on the assumption, supported by Graetz's (1991) proposal, that the *GHQ-12* is a multidimensional scale containing three distinct dimensions (anxiety and depression, social dysfunction and loss of confidence) which are considered to have an underlying second order factor. In order to compare the trajectory of a single indicator to the trajectories of these three factors, the first step was to model the three factors' trajectory considering the individuals' answers scores, obtained by adding the answers to the respective items. The models' comparison led to the conclusion that the potential explicative gain from a tripartite model is slim. The same occurs when modelling the trajectories considering the ordinal scales of the input variables. On the contrary, the estimation of the conditioned latent growth models led to the conclusion that the model which considers the repeated ordinal measures and a second order factor has a larger explicative ability than both the conditioned LGCM models: one regarding the subjective well-being single indicator and another the latent factor measured by the scores in the anxiety and depression, social dysfunction and loss of confidence dimensions.

JEL Classification: C33, I19

Keywords: General Health Questionnaire, longitudinal data, British Household Panel Survey, Latent Growth Curve Models

Agradecimentos

O primeiro agradecimento é devido à professora Fátima Salgueiro, pela sua ajuda, paciência e, nalguns momentos decisivos, de pânico, presença de espírito e clarividência para voltar a endireitar a “espinha” deste trabalho. Agradeço também a todos os professores de mestrado. Não posso deixar de evidenciar os professores que mais me ensinaram, e que deram valor acrescido à qualidade do mestrado: Prof^a. Margarida Cardoso, Prof^a Helena Carvalho, Prof. José Dias, Prof^a. Manuela Hill, Prof^a. Elizabeth Reis e, mais uma vez, Prof^a. Fátima Salgueiro. Agradeço ainda ao Miguel Tecedeiro por me ter posto em contacto com o Prof. João Maroco, a quem agradeço o tempo ao telefone, e as sugestões e referências úteis.

Ainda... agradeço, com toda a sinceridade, à Ana Lúcia. De facto, é um prazer trabalhar contigo, e a tua ajuda é sempre, sempre importante. A “dupla maravilha” dos trabalhos de mestrado estendeu-se inequivocamente, ao longo destes anos, ao campo da amizade. Agradeço também à minha família (à minha mãe em especial, porque hoje faço anos!), pela paciência, prometendo que a partir de agora estarei mais presente e disponível. Por fim, ao Ped’Nuno, para quem já não tenho palavras de agradecimento possíveis... Faço-o na partilha diária da minha vida, das minhas tristezas e alegrias, há tempo que nem sei contar. É contigo em especial que partilho este “filho(a)”.

ÍNDICE

1. INTRODUÇÃO.....	1
2. GENERAL HEALTH QUESTIONNAIRE – 12.....	3
3. AMOSTRA EM ANÁLISE	7
4. ESTRATÉGIA ADOPTADA PARA A ANÁLISE LONGITUDINAL DO GHQ-12.....	15
5. MODELOS DE TRAJECTÓRIA LATENTE.....	19
5.1. Modelos de trajectória latente não condicionada.....	20
5.1.1. Estrutura matricial do modelo de trajectória latente não condicionada	26
5.2. Modelos de trajectória latente condicionada	28
5.2.1. Estrutura matricial do modelo de trajectória latente condicionada	31
5.3. Modelos de trajectória latente de variáveis ordinais.....	34
5.3.1. Violação de pressupostos.....	35
5.3.2. Procedimentos de correcção e estimação por WLS	37
5.3.3. Parametização de LGCM com variáveis ordinais	42
5.4. Modelos de trajectória latente considerando múltiplos indicadores	44
5.5. Medidas de bondade do ajustamento	47
6. ANÁLISE DOS RESULTADOS.....	52
6.1. Modelos de trajectória latente não condicionada do GHQ-12.....	52
6.2. Modelos de trajectória latente condicionada do GHQ-12.....	62
7. DISCUSSÃO DOS RESULTADOS.....	66
8. CONCLUSÃO	68
9. BIBLIOGRAFIA.....	71
10. ANEXOS	73
10.1. Lista de variáveis da escala GHQ-12 e indicador único de SWB	73
10.2. Modelo conceptual em análise.....	74
10.3. Equações, definições e pressupostos de um modelo de trajectória latente não condicionada.....	74
10.4. Diagramas dos modelos estimados.....	75
10.4.1. Diagrama do modelo I: trajectória latente não condicionada do índice agregado de bem-estar subjectivo (SWB)	75
10.4.2. Diagrama do modelo AD: trajectória latente não condicionada do índice de ansiedade e depressão (AD)	76
10.4.3. Diagrama do modelo SD: trajectória latente não condicionada do índice de disfunção social (SD)76	
10.4.4. Diagrama do modelo LC: trajectória latente não condicionada do índice de perda de confiança (LC)	76

10.4.5. Diagrama do modelo 2: relacionando as trajectórias latentes dos três processos (AD, SD e LC)..	77
10.4.6. Diagrama do modelo 3: trajectória latente não condicionada do factor de debilidade psicológica (PM), modelado a partir dos índices AD, SD e LC	77
10.4.7. Diagrama do modelo 4: trajectória latente não condicionada dos factores latentes AD, SD e LC .	78
10.4.8. Diagrama do modelo 5: trajectória latente do factor de segunda ordem de PM	79
10.5. Comandos de Mplus utilizados para a estimação do modelo 4	79
10.6. Comandos de Mplus utilizados para a estimação do modelo 5 e 5c (com variáveis explicativas) comentados	82
10.7. Comandos de LISREL utilizados para a estimação do modelo 4	85
10.7.1. Passo 1: comandos para a estimação dos thresholds e da matriz de correlações policóricas comentados	85
10.7.2. Passo 2: comandos para a estimação da matriz de covariâncias assintóticas comentados.....	85
10.7.3. Passo 3: comandos para a estimação do modelo 4	86
10.8. Comandos de LISREL utilizados para a estimação do modelo 5 comentados	87

Índice de Tabelas

Tabela 1 – Médias e desvios-padrão do indicador único <i>SWB</i> nos quatro períodos considerados para indivíduos do sexo masculino	12
Tabela 2 – Médias e desvios-padrão do indicador único <i>SWB</i> nos quatro períodos considerados para indivíduos do sexo feminino	12
Tabela 3 – Estimativas estandardizadas dos pesos factoriais no modelo unifactorial e trifactorial do <i>GHQ-12</i>	15
Tabela 4 – Medidas de bondade do ajustamento dos modelos unifactorial e trifactorial.....	16
Tabela 5 – Estimativas estandardizadas da estrutura de relações entre os factores de primeira primeira ordem e o factor de 2ª ordem.....	17
Tabela 6 – Estimativas para os modelos 1, 2 e 3, com trajectórias não condicionadas.....	53
Tabela 7 – Estimativas para os modelos de trajectória dos <i>scores</i> ansiedade e depressão (<i>AD</i>), disfunção social (<i>SD</i>) e perda de confiança (<i>LC</i>).....	54
Tabela 8 – Medidas de bondade do ajustamento dos modelos de trajectória dos <i>scores</i> ansiedade e depressão (<i>AD</i>), disfunção social (<i>SD</i>) e perda de confiança (<i>LC</i>)	54
Tabela 9 – Medidas de bondade do ajustamento dos modelos de trajectória 1, 2 e 3	57
Tabela 10 – Estimativas dos modelos 4 e 5, com trajectórias não condicionadas.....	57
Tabela 11 – Medidas de bondade do ajustamento dos modelos de trajectória 4 e 5	61
Tabela 12 – Impacto das variáveis explicativas nos modelos 1, 3 e 5 (modelos 1c, 3c e 5c) de trajectória latente condicionada.....	62
Tabela 13 – Medidas de bondade do ajustamento dos modelos de trajectória latente condicionada (modelos 1c, 3c e 5c).....	65

Índice de Figuras

Figura 1 – Distribuição dos indivíduos por sexo.....	9
Figura 2 – Distribuição dos indivíduos por escalões etários	9
Figura 3 – Perda de sono	9
Figura 4 – Constantemente sob pressão	9
Figura 5 – Dificuldade em superar problemas	10
Figura 6 – Triste ou deprimido.....	10
Figura 7 – Concentração.....	10
Figura 8 – Sente que tem desempenhado papel útil	10
Figura 9 – Capacidade de tomar decisões	10
Figura 10 – Capacidade de apreciar actividades diárias.....	10
Figura 11 – Capacidade de encarar problemas.....	11
Figura 12 – Felicidade em geral.....	11
Figura 13 – Perda de confiança	11
Figura 14 – Auto-estima.....	11
Figura 15 – Médias do indicador de bem-estar psicológico e das dimensões de ansiedade e depressão, disfunção social e perda de confiança para indivíduos do sexo masculino.....	13
Figura 16 – Médias do indicador de bem-estar psicológico e das dimensões de ansiedade e depressão, disfunção social e perda de confiança para indivíduos do sexo feminino.....	13
Figura 17 – Distribuição do estado de saúde dos homens ao longo dos períodos considerados.....	14
Figura 18 – Distribuição do estado de saúde das mulheres ao longo dos períodos considerados.....	14
Figura 19 – Modelo de trajectória linear não condicionada da variável y.....	21
Figura 20 – Modelo de trajectória linear latente condicionada da variável y ao longo de 4 períodos consecutivos	29
Figura 21 – Modelo de trajectória linear latente não condicionada para uma variável ordinal observada em 4 momentos consecutivos.....	34
Figura 22 – Variável contínua subjacente a uma variável ordinal com quatro categorias	37
Figura 23 – Modelo de trajectória linear latente de um factor latente medido por três indicadores, em três momentos consecutivos.....	45

SUMÁRIO EXECUTIVO

Nas sociedades ocidentais actuais o estudo das percepções de bem-estar dos indivíduos é cada vez mais um assunto que preocupa cientistas das mais variadas áreas. Estudos na área da psicologia e psicometria têm desenvolvido várias escalas que procuram medir o bem-estar nas suas múltiplas dimensões. A escala de *General Health Questionnaire (GHQ)*, proposta por Goldberg em 1970, inicialmente constituída por 60 itens, foi desenvolvida para avaliar a saúde mental de pacientes de cuidados primários de saúde. A partir desta escala psicométrica, foram aplicadas e validadas escalas mais reduzidas, com o objectivo de medir dimensões específicas da saúde psicológica. É neste contexto que surge a escala *GHQ-12*, uma subescala do questionário original constituída por 12 itens, inicialmente concebida como medindo uma única dimensão de *psychological morbidity*, aqui traduzida por debilidade psicológica. No entanto, esta abordagem tem sido largamente posta em causa por inúmeros autores, onde se defende que estes 12 itens estruturam mais de uma dimensão de bem-estar psicológico.

O trabalho que se apresenta considera, de acordo com vários autores, que a escala *GHQ-12* estrutura três dimensões distintas: *ansiedade e depressão*, medida por quatro indicadores; *disfunção social*, medida por seis indicadores; e *perda de confiança*, medida por dois indicadores. Note-se que todas as variáveis são medidas em escalas de tipo Likert com quatro pontos, que são os mesmos para as variáveis que estruturam cada um dos constructos. Tendo em conta as propostas teóricas discutidas, opta-se ainda por considerar um factor de segunda ordem, tendo subjacente a ideia de que os três constructos considerados constituem uma única dimensão coerente de debilidade psicológica.

Propõe-se a modelação longitudinal do factor de segunda ordem, utilizando dados do painel *British Household Panel Survey (BHPS)*, aplicado desde 1991 e representativo dos agregados familiares Britânicos. A presente modelação longitudinal considera apenas dados de 2003 a 2006 (*waves m a p*) e indivíduos que tenham respondido ao questionário desde a sua primeira aplicação. Ainda, consideram-se apenas indivíduos com respostas válidas nos 12 itens. A amostra final é constituída por 4562 indivíduos, dos quais 2548 (55,8%) são mulheres e 2014 (44,2%) são homens.

Geralmente, os estudos com base no *GHQ-12* utilizam dados seccionais, aplicando uma variedade de métodos, desde modelos de equações estruturais, análises factoriais, modelos *logit* ou *probit*, entre tantos outros. Aqueles que utilizam dados longitudinais consideram por

regra dois indicadores disponibilizados no BHPS, que resultam da soma dos *scores* nos 12 itens, e utilizam modelos *ordered probit* (e.g. Clark, 2002; Clark e Oswald, 2002; Clark e Oswald, 2006). Aqui, a modelação longitudinal é feita com recurso aos métodos dos *Latent Growth Curve Models (LGCM)*, cuja abordagem tem subjacente a ideia de que a trajectória do fenómeno em estudo é latente, não observada. Este método, por se desenvolver no quadro dos modelos de equações estruturais, partilha muitas das suas vantagens. Por um lado, permite uma fácil articulação entre um modelo teórico e a sua modelação estatística. Permite ainda obter medidas de bondade do ajustamento para o modelo global. Ainda, tem a grande vantagem de, para além de estimar trajectórias globais dos fenómenos em análise, ter também em conta a heterogeneidade individual.

A abordagem à escala *GHQ-12* segue um percurso a partir do modelo mais simples de trajectória do indicador único de bem-estar subjectivo, comparando os resultados, num primeiro momento, com as estimativas obtidas para dois modelos de trajectória dos indicadores de ansiedade e depressão, disfunção social e perda de confiança construídos a partir da soma das respostas dos indivíduos aos itens que os constituem. O primeiro destes modelos considera três processos correlacionados e o segundo considera um factor que se denominou de debilidade psicológica, medido pelas três variáveis construídas.

Num segundo momento estimam-se os modelos de trajectória dos factores medidos pelas variáveis ordinais. Consideram-se também dois modelos distintos: um com três processos correlacionados e outro com um factor de segunda ordem, que agrega os factores (agora latentes) de ansiedade e depressão, disfunção social e perda de confiança.

A comparação dos modelos permitiu concluir que, considerando modelos de trajectória latente não condicionada, a trajectória do indicador único reflecte os resultados obtidos para os demais modelos.

Com a introdução de variáveis explicativas, as conclusões alteram-se de forma inequívoca. Consideraram-se o sexo, o escalão etário e o estado de saúde percebido como possíveis variáveis explicativas. Todas as variáveis foram tratadas como invariantes ao longo dos quatro períodos em análise. Relativamente ao sexo, concluiu-se que as mulheres têm níveis médios de debilidade psicológica mais altos do que os homens em 2003. Esta conclusão é transversal a todos os modelos testados. Quanto ao estado de saúde percebido pelos indivíduos, foi possível concluir que, relativamente aos indivíduos que consideram estar em excelente estado de saúde, aqueles com piores estados de saúde apresentam graus médios de debilidade psicológica no primeiro momento mais elevados. Por fim, a relação entre idade e debilidade

psicológica encontrada sugere um conjunto de conclusões interessantes. Por um lado, quando consideramos modelos de trajectória da debilidade psicológica com uma estrutura tripartida da escala *GHQ-12* (quer na modelação da trajectória a partir dos indicadores métricos, quer a partir das variáveis ordinais), percebe-se existir uma relação não linear entre a idade e a trajectória, de forma semelhante ao que conclui Clark (2002) e Clark e Oswald (2006). De facto, é possível concluir que, relativamente a indivíduos com idades entre os 16 e os 29 anos (esta é a categoria de referência), indivíduos com idades entre os 30 e os 49 anos têm valores médios, em 2003, que apontam para uma maior debilidade psicológica, ao passo que os indivíduos com 65 ou mais anos têm, também em 2003, menores graus médios de debilidade. Aponta-se como possível causa o efeito de *coorte* evidenciado por aqueles autores, que sugere que o ano de nascimento está correlacionado com o bem-estar dos inquiridos.

Concluiu-se ainda que um modelo de trajectória do factor de segunda ordem permite evidenciar o impacto de variáveis (ou categorias de variáveis) explicativas, que antes não tinham significância estatística.

A modelação apresentada foi feita com recurso ao *software* estatístico *Mplus*. Sempre que possível, a estimação foi também feita em LISREL. De forma a criar um instrumento de trabalho potencialmente útil a outros investigadores, apresentam-se em anexo os comandos para a estimação dos modelos nos dois *softwares* utilizados.

“The fact that we live at the bottom of a deep gravity well, on the surface of a gas covered planet going around a nuclear fireball 90 million miles away and think this to be normal is obviously some indication of how skewed our perspective tends to be.”

Douglas Adams, discurso no *Second Annual Conference on Cyberbiology*, Cambridge, 1998

1. INTRODUÇÃO

O estudo que se apresenta incide sobre trajectórias de bem-estar psicológico. Uma análise seccional do fenómeno num dado momento permitiria perceber, para aquele momento, qual seria a estrutura de relações entre os factores considerados. Sublinha-se que os resultados diriam apenas respeito ao bem-estar dos indivíduos num dado momento. Uma análise longitudinal permite não só perceber o nível de bem-estar dos indivíduos, como ainda conhecer a sua evolução e quais os factores que interferem na trajectória. Tem em linha de conta uma característica que é inerente aos fenómenos em estudo: a mudança. A realidade não é estática, e os fenómenos estão em permanente evolução, numa teia de interdependências que uma análise estatística não permite apreender na íntegra. No entanto, assumindo que estamos sempre perante uma simplificação da realidade, podemos (e devemos) procurar reduzir a margem do que fica por explicar, do que ultrapassa o nosso modelo de análise. Assim, considerando a natureza dinâmica, é possível reduzir a distância entre a realidade e o modelo. Esta é sem dúvida uma das mais notáveis vantagens de uma análise de dados longitudinal.

A abordagem aos dados longitudinais será feita no quadro dos modelos de trajectória latente (*Latent Growth Curve Models – LGCM*). Por se enquadrar na estrutura metodológica dos modelos de equações estruturais, encerra também muitas das suas vantagens, como a fácil adaptação de um modelo conceptual, articulando de forma simples diversas hipóteses em estudo. Na introdução ao livro *Latent Growth Curve Modeling* (Preacher *et al.*, 2008) o editor enuncia as vantagens da utilização de métodos de análise longitudinal. Mais ainda, afirma que “*even more appropriate can be the method of latent growth curve models, with which the data analyst can not only deal with unobserved heterogeneity but also model tendencies (...) over the life course*”. É certamente uma vantagem destes métodos a capacidade de analisar não só as trajectórias de fenómenos sociais, económicos, ou de outra natureza, como considerar ainda as diferenças de variabilidade ao nível individual. Permitindo a introdução de variáveis explicativas é possível perceber alguma da heterogeneidade individual.

Os dados utilizados neste estudo pertencem ao *British Household Panel Survey* (BHPS), um painel de dados representativos dos agregados familiares da Grã-Bretanha, activo desde 1991. São considerados dados relativos aos anos consecutivos de 2003 a 2006. Para o estudo das trajectórias de bem-estar psicológico é utilizada a escala psicométrica *General Health Questionnaire* (GHQ-12), aplicada através de um questionário de auto-preenchimento aos

indivíduos pertencentes a cada agregado familiar, com idades iguais ou superiores a 16 anos. O objectivo central do trabalho que se apresenta é a utilização de *LGCM* para modelar a trajectória do bem-estar psicológico, com base numa amostra constituída por 4562 indivíduos, que participam no painel do BHPS desde o seu início.

No capítulo 2 do presente documento é feita uma apresentação e discussão das principais propostas de abordagem à escala *GHQ-12*, sobre as quais assenta o presente estudo.

No capítulo 3 faz-se uma caracterização da amostra em estudo, apresentando com algum detalhe o painel BHPS. Faz-se uma caracterização breve das variáveis que constituem a escala *GHQ-12*, bem como daquelas a considerar como explicativas de alguma heterogeneidade individual.

O capítulo 4 expõe, de forma estruturada, a abordagem seguida para a análise das trajectórias de bem-estar psicológico, apresentando os vários modelos que se pretendem estimar. Apresentam-se também resultados de análises factoriais confirmatórias prévias à modelação das trajectórias, de forma a estabelecer o modelo conceptual em estudo.

O capítulo 5 dedica-se à exposição metodológica dos métodos aplicados. Aqui apresentam-se com maior detalhe as técnicas, pressupostos, estrutura conceptual, parametrização e métodos de estimação de modelos de trajectória latente não condicionada e condicionada, considerando medidas repetidas ordinais e modelos com múltiplos indicadores. Apresentam-se ainda as medidas de bondade do ajustamento utilizadas.

O capítulo 6 apresenta as conclusões a retirar dos vários modelos testados, comparando as diversas soluções. De forma a facilitar a análise e a leitura do documento, dividiu-se o capítulo 6 em duas secções distintas, dedicadas, respectivamente, aos modelos estimados de trajectória latente não condicionada e aos modelos considerando variáveis explicativas.

No capítulo 7 apresenta-se a discussão dos resultados, articulando as hipóteses consideradas inicialmente e os resultados obtidos, procurando ter em vista as propostas apresentadas pelos diversos autores, e discutidas anteriormente.

No capítulo 8 apresentam-se algumas conclusões gerais a retirar, enunciando pontos fracos e pontos fortes do trabalho apresentado, e propondo possíveis estratégias futuras com vista à resolução de questões deixadas em aberto ou apenas afloradas neste estudo.

No capítulo 10, que se segue à lista da bibliografia utilizada, apresenta-se um conjunto de anexos potencialmente úteis ao leitor, que facilitam a compreensão de alguns dos argumentos utilizados.

2. GENERAL HEALTH QUESTIONNAIRE – 12

O *General Health Questionnaire (GHQ)* é uma escala psicométrica já vastamente validada, desenvolvida para detectar disfunções psíquicas fora do contexto clínico. A escala inicial, constituída por 60 itens, foi desenvolvida por David P. Goldberg em 1970 como escala de medição da saúde mental de pacientes de cuidados primários de saúde. Apesar de a escala do *GHQ* ter sido desenvolvida “*as a screening instrument for psychiatric illness (...), it is often used as an indicator of psychological well-being*” (Jones e Widman, 2008: 311).

Por ser de difícil aplicação e relativamente redundante, a escala *GHQ-60* foi posteriormente reduzida para escalas contendo um menor número de itens. A versão do *GHQ* constituída por 12 itens é a mais curta escala resultante do *GHQ-60*, e tem a vantagem indiscutível de ser de mais fácil aplicação, mantendo algumas das mais relevantes propriedades da escala psicométrica alargada (Shevlin e Adamson, 2005: 231).

A versão reduzida do *GHQ* considerando 12 itens foi também desenvolvida com vista à detecção de “*psychiatric disorders (...) in community settings and non-psychiatric settings*” (Goldberg citado por Hankins, 2008a: 2). Esta escala foi validada sob o pressuposto de ser uma medida unidimensional de debilidade psicológica. De facto, a dimensionalidade da escala *GHQ-12* tem sido um tema de debate, e encontram-se estudos que sugerem estruturas bastante diferentes. Apesar de estes 12 itens terem sido desenvolvidos com o objectivo de medir uma dimensão do distúrbio psicológico dos inquiridos, multiplicam-se os artigos que contestam esta abordagem, apresentando resultados que mostram que modelos com dois ou três factores resultam num melhor ajustamento (*e.g.* Graetz, 1991; Cheung, 2002; Shevlin e Adamson, 2005; Hankins, 2008a; Ye, 2008). Um dos argumentos mais utilizados na discussão da multidimensionalidade da escala remete para a formulação das questões que a constituem.

Como nota Hankins (2008a, 2008b), a multidimensionalidade da escala do *GHQ-12* pode ser resultante de um efeito vastamente conhecido em psicologia, decorrente da utilização de escalas que misturam formulações positivas e negativas. Este efeito, denominado por ‘efeito de método’, pode resultar efectivamente numa divisão dos itens em factores, relativamente coerentes, com a divisão entre formulação positiva e negativa. O *GHQ-12* contém, de facto, seis descrições positivas de estados de espírito ou disposições (*e.g.* “*...been feeling reasonably happy, all things considered?*”), e outras seis negativas (*e.g.* “*...been thinking of*

yourself as a worthless person?”)¹. O autor explica ainda que itens formulados pela negativa são conhecidos por provocarem algum tipo de enviesamento, seja pela posse diferenciada de níveis de instrução, ou pela aversão à formulação de estados emocionais pela negativa. Estes factores são conhecidos por causar uma maior dificuldade por parte dos inquiridos em responderem a estas questões.

Outro factor que leva a considerar uma estrutura multidimensional do *GHQ-12* é o facto de, embora os itens sejam todos medidos em escalas de tipo Likert com quatro pontos, as escalas não serem sempre iguais. Este facto levanta um problema fundamental, que põe em causa o rigor de um modelo com um factor medido por todas as variáveis. É certamente contestável a opção de considerar estas variáveis como sendo medidas na mesma escala. Recordamos que, embora as variáveis sejam medidas por um mesmo número de pontos, isto não significa que a escala seja a mesma. De facto, as escalas de tipo Likert não permitem averiguar a distância a que estão os seus pontos. Se estes pontos são substantivamente diferentes, também poderão ser as distâncias que os separam.

O primeiro autor a sugerir um modelo de medida do *GHQ-12* constituído por três factores ou dimensões foi Graetz (1991). No seu artigo *Multidimensional Properties of the General Health Questionnaire*, o autor reexamina a estrutura factorial do *GHQ-12* a partir de uma grande amostra de jovens australianos e conclui, através de uma análise simples em componentes principais com rotação oblíqua, que há vantagens claras na utilização do *GHQ* numa perspectiva multidimensional. Graetz identifica três dimensões distintas: “*Factor I includes elements of anxiety and depression (...) factor II includes aspects of anhedonia and social dysfunction [and] factor III contains just two items that encompass loss of confidence or self-esteem*” (Graetz, 1991: 135). A partir destes factores, o autor denomina as três dimensões de *ansiedade e depressão, disfunção social e perda de confiança*².

A partir deste estudo, vários autores aplicaram métodos diferentes para a análise factorial da escala do *GHQ-12*. Gao *et al.* (2004) compararam estruturas com um, dois e três factores, e concluíram que o modelo com melhor ajustamento era o modelo proposto por Graetz (1991). Concluíram ainda existirem fortes correlações entre os três factores enunciados.

Também Campbell e Knowles (2007) comparam modelos com diferentes estruturas factoriais. Tendo ainda em conta diferentes codificações nas respostas aos itens (considerando quer as escalas de tipo Likert, quer dicotomizando as respostas ao itens de diferentes formas, de

¹ Para a consulta das 12 variáveis que constituem o *GHQ-12* e respectivas escalas de medida, ver anexo 10.1.

² Vários autores atribuem a designação de *psychological morbidity* ao conceito subjacente às três dimensões enunciadas. Substituiu-se esta denominação por *debilidade psicológica*, embora se mantenha a sigla *PM*.

maneira a incorporar os efeitos de nomeação positiva e negativa referidos anteriormente), os autores concluem que, considerando as variáveis medidas em escalas de tipo Likert, o modelo proposto por Graetz era o que resultava num melhor ajustamento em todos os indicadores. No entanto, sublinham também a forte correlação entre os três factores.

Cheung (2002) utiliza dados do *British Household Panel Survey* para uma análise seccional da estrutura do *GHQ-12* em indivíduos com mais de 60 anos. As conclusões retiradas apontam também para um melhor ajustamento de um modelo considerando uma estrutura tripartida da escala em análise. No entanto, a estrutura de correlações evidenciada entre os três factores é, neste estudo, bastante mais moderada.

Também Shevlin e Adamson (2005) testaram modelos factoriais concorrentes no estudo do *GHQ-12*. Estes autores concluíram que, embora de facto o modelo trifactorial resultasse num melhor ajustamento, *“the predictive utility of the three subscales, as opposed to the total GHQ-12 score, is questioned as the three subscales do not appear to contribute any additional information, at least with regard to the outcome variables considered in this article”* (Shevlin e Adamson, 2005: 235). Embora os autores chamem a atenção para o facto de a estrutura factorial tripartida ter sido validada em estudos independentes, que analisam dados de diferentes fontes (os autores notam, aliás, que o modelo trifactorial estimado é muito semelhante ao apresentado por Cheung, 2002), consideram a hipótese de esta ser uma estrutura desnecessária, obtendo-se resultados semelhantes pela utilização do indicador de bem-estar subjectivo construído a partir da soma das respostas dos inquiridos aos 12 itens do *GHQ-12*.

A discussão das propostas de abordagem ao *GHQ-12* apresentada leva a que se pondere a utilidade de um modelo factorial de segunda ordem, construído sob o pressuposto de que os factores de ansiedade e depressão, disfunção social e perda de confiança medem dimensões distintas de um mesmo constructo de debilidade psicológica, resultante da constatação quase transversal a todos os estudos apresentados de que os três constructos enunciados por Graetz (1991), fazendo sentido do ponto de vista conceptual, estão fortemente correlacionados.

A partir das várias propostas apresentadas, e da ideia de que os três factores (ansiedade e depressão, disfunção social e perda de confiança) podem ser explicados por um factor de ordem superior que se denominou de debilidade psicológica, a abordagem longitudinal ao *GHQ-12* apresentada no presente trabalho parte da hipótese de que uma estrutura factorial mais complexa, reproduzindo um modelo conceptual elaborado, permite evidenciar trajectórias de debilidade psicológica diferenciadas a partir de algumas variáveis explicativas.

Pretende-se testar também se, de facto, e como sugerem Shevlin e Adamson (2005), o indicador único de bem-estar subjectivo é suficiente para perceber a evolução do bem-estar dos indivíduos, ou se uma estrutura tripartida ou, ainda, de segunda ordem, evidenciam especificidades naquela trajectória.

3. AMOSTRA EM ANÁLISE

O *British Household Panel Survey*, desenvolvido pelo *Longitudinal Studies Center* do *Economic and Social Research Council* (ESRC) e pelo *Institute for Social and Economic Research* da Universidade de Essex, é um painel representativo dos agregados familiares da Grã-Bretanha, activo desde 1991. Actualmente um dos maiores painéis existentes³, o seu principal objectivo é o de recolher informação de forma a “*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*” (Taylor *et al.*, 2008: A2-1).

Para além da recolha de informação ao nível dos agregados familiares, o BHPS recolhe ainda informação relativa aos indivíduos que o constituem, com idades superiores a 16 anos, procurando manter a sua participação ao longo do tempo⁴. A partir da amostra recolhida no primeiro momento de observação (ano de 1991), da qual constavam 8167 agregados familiares, foram sendo incluídos na amostra indivíduos recém-nascidos nesses agregados, indivíduos que se juntassem aos agregados originalmente considerados, ou novos indivíduos que passassem a habitar com aqueles que estavam em agregados inicialmente seleccionados (*i.e.*, todos os indivíduos que partilham habitação com elementos da amostra inicial)⁵.

As variáveis que constituem o *GHQ-12* foram recolhidas a partir de um questionário de auto-preenchimento (*Self Completion Questionnaire*). Este questionário, aplicado desde o primeiro momento de observação, contém, para além das 12 variáveis da escala de saúde psicológica, outras questões relativas ao bem-estar psicológico dos inquiridos.

Para o estudo apresentado, por se pretender essencialmente avaliar a adequabilidade e mais-valia de modelos de trajectória latente construídos a partir das variáveis tal como foram recolhidas (medidas em escalas de tipo Likert) relativamente a modelos que consideram apenas índices que agregam a informação deste 12 indicadores, e por estarmos a considerar modelos de relativa complexidade, optou-se por utilizar apenas informação recolhida em quatro momentos temporais. Assim, consideram-se para análise os dados recolhidos pelo

³ Em 2009, o BHPS conterà informação recolhida ao longo de 18 momentos de observação.

⁴ Desde 1994 é ainda aplicado um questionário de auto-preenchimento aos jovens entre os 11 e 15 anos (*Youth Questionnaire*).

⁵ Para uma leitura pormenorizada dos procedimentos de amostragem e seguimento dos indivíduos, ver Taylor *et al.*, 2008: A4-1 a A4-49.

BHPS nas *waves* m a p, relativas aos anos de 2003 a 2006. Houve ainda a preocupação de utilizar a informação mais recentemente disponibilizada⁶.

A construção da base de dados para análise obedeceu a alguns critérios de selecção dos indivíduos. De uma amostra inicial de 10 606 indivíduos com respostas válidas nos 12 itens que constituem o *GHQ-12*⁷, consideraram-se apenas indivíduos que participaram no estudo desde o primeiro momento, *i.e.*, que estavam incluídos no painel desde o seu início⁸. Ao longo do processo de junção das bases de dados, tomou-se a opção de deixar em aberto a possibilidade de considerar o desenho amostral na modelação a efectuar. Embora esta linha de investigação tenha sido abandonada no presente estudo, optou-se por manter a amostra tal como tinha sido inicialmente considerada.

A selecção dos indivíduos que participaram no estudo desde o início foi feita a partir da variável relativa aos pesos longitudinais dos indivíduos na amostra constante do painel. Porque “*only cases who have responded at each wave up to and including the latest wave will have positive longitudinal weights at that wave*” (Taylor *et al.*, 2008: A5-1), incluíram-se na amostra os indivíduos com pesos longitudinais no último momento temporal considerado (*wave* p, correspondente ao ano de 2006) superiores a zero⁹. Esta condição reduziu a amostra inicial para 4 562 indivíduos. São estes os casos considerados na amostra em estudo.

Embora a amostra seja constituída por mais mulheres do que homens (55,8% de mulheres e 44,2% de homens), a proporção é muito próxima (ver figura 1).

⁶ A *wave* q, relativa a informação recolhida no ano de 2007, só foi disponibilizada no decurso deste estudo, não tendo sido possível, por isso, considerá-la para as análises realizadas.

⁷ Para este estudo não se considerou a problemática em torno da modelação de LGCM com casos omissos, o que implicaria diferentes estratégias para a estimação e parametrização. Para uma análise mais detalha *cf. e.g.* Bollen e Curran (2006: 58-87), Duncan *et al.* (2006: 179-193), Preacher *et al.* (2008: 15-18).

⁸ Note-se que uma criança nascida em 1991, ano do início do painel BHPS, num agregado familiar original do painel, é seguida desde essa data e incluída na amostra.

⁹ Para uma leitura detalhada do desenho amostral do BHPS, consultar as secções IV e V do manual de utilização (Taylor *et al.*, 2008).

Figura 1 – Distribuição dos indivíduos por sexo

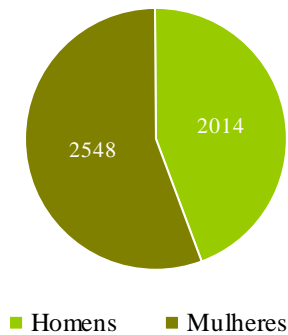
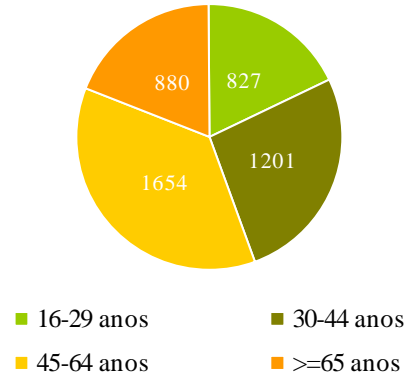


Figura 2 – Distribuição dos indivíduos por escalões etários



Em todos os momentos de observação foi recolhida a idade dos indivíduos a 1 de Dezembro do ano anterior ao da recolha. Assim, de forma a permitir avaliar a coerência dos dados (sabemos que, à mesma data do ano seguinte, os indivíduos terão sempre mais um ano de idade), considerou-se a idade dos inquiridos a 1 de Dezembro de 2002. A média de idades na amostra é de cerca de 48 anos, com um desvio-padrão de aproximadamente 18 anos¹⁰. A idade mínima observada é de 16 anos, e o indivíduo mais velho tinha, à data referida, 96 anos. Relativamente aos 12 itens que constituem o *GHQ-12*, os gráficos das figuras 3 a 14 permitem perceber que as respostas dos indivíduos não se alteram muito ao longo do tempo. Na formulação de todas as questões perguntava-se aos inquiridos se recentemente tinham sentido/percepcionado algum dos sintomas psicológicos, de cansaço ou mau-estar. As figuras 3 a 6 representam as respostas dos indivíduos aos quatro itens que constituem a primeira dimensão de debilidade psicológica: ansiedade e depressão. Na escala em que as variáveis foram medidas, o primeiro ponto representa a ausência do sintoma, e o último ponto um aumento do sintoma sentido nos últimos tempos¹¹.

Figura 3 – Perda de sono

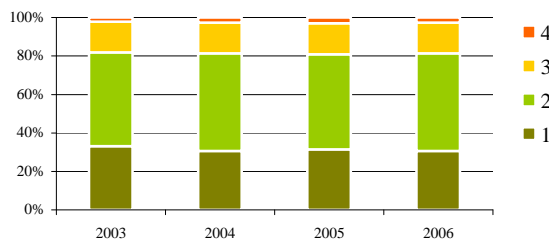
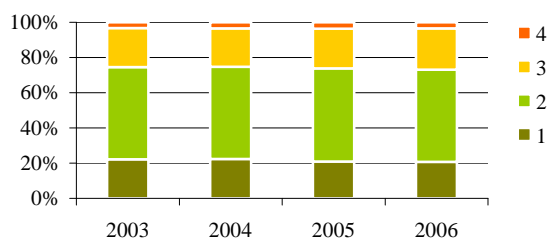


Figura 4 – Constantemente sob pressão



¹⁰ Mais exactamente, a média é de 47,86 anos e o desvio-padrão de 17,78 anos.

¹¹ Os primeiros quatro gráficos representam a distribuição das seguintes variáveis, ao longo dos quatro períodos considerados: “Have you recently lost much sleep over worry?”, “Have you recently felt constantly under strain?”, “Have you recently felt you couldn’t overcome your difficulties?” e “Have you recently been feeling unhappy or depressed?”. As quatro variáveis foram medidas com os seguintes pontos: (1) Not at all; (2) No more than usual; (3) Rather more than usual; (4) Much more than usual.

Figura 5 – Dificuldade em superar problemas

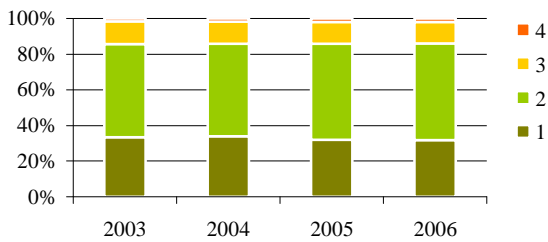
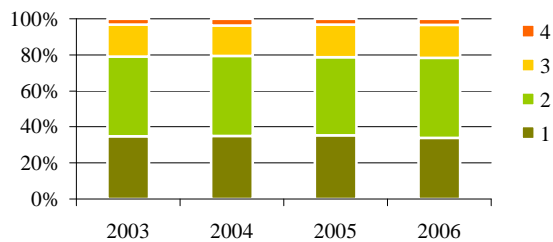


Figura 6 – Triste ou deprimido



A maior parcela dos indivíduos afirma não ter sentido qualquer um dos sintomas mais do que o habitual, e esta tendência mantém-se ao longo dos quatro períodos observados.

As figuras 7 a 12 representam a distribuição dos itens que anteriormente referimos constituírem a dimensão de disfunção social de debilidade psicológica. Estas seis variáveis foram igualmente medidas em escalas de quatro pontos, mas agora, porque os itens são, ao contrário dos restantes, de formulação positiva, a escala é inversa à anterior. O primeiro ponto representa um maior bem-estar psicológico, decrescendo ao longo da escala¹².

Figura 7 – Concentração

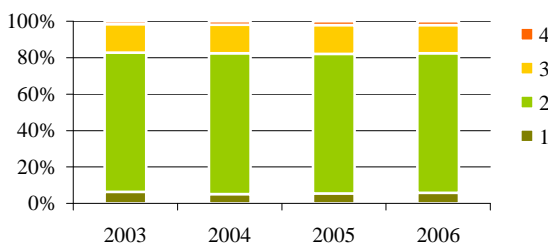


Figura 8 – Sente que tem desempenhado papel útil

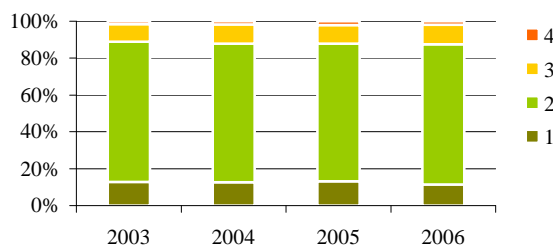


Figura 9 – Capacidade de tomar decisões

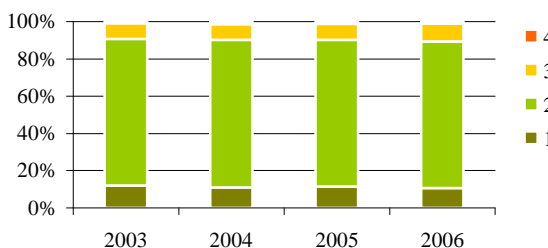
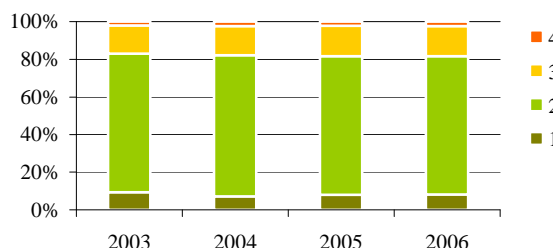


Figura 10 – Capacidade de apreciar actividades diárias



¹² Os seis gráficos representam a distribuição das seguintes variáveis, ao longo dos quatro períodos considerados: “Have you recently been able to concentrate on whatever you’re doing?”, “Have you recently felt that you were playing a useful part in things?”, “Have you recently felt capable of making decisions?”, “Have you recently been able to enjoy your normal day-to-day activities?”, “Have you recently been able to face up to problems?” e “Have you recently been feeling reasonably happy, all things considered?”. Os pontos da escala em que foram medidas são os seguintes: (1) More than usual; (2) Same as usual; (3) Less than usual; (4) Much less than usual.

Figura 11 – Capacidade de encarar problemas

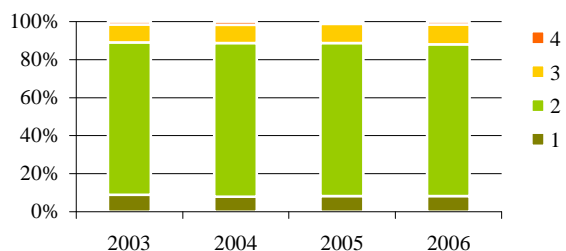
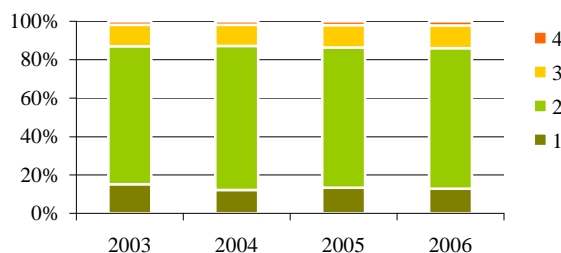


Figura 12 – Felicidade em geral



Também neste conjunto de variáveis se observa que a maior parcela de inquiridos considera que se sente, relativamente a cada um destes factores, como habitualmente. Refira-se no entanto que, relativamente ao conjunto de variáveis observadas anteriormente, esta parcela é bastante maior. A redução desta parcela nos itens que enunciam factores negativos de saúde psicológica poderá ser devida à forma como são colocadas as questões. De facto, Hankins (2008a: 2) sugere que “*negatively phrased items are more difficult to process because of inattention, variation in education, or an aversion to negative emotional content*”.

Por fim, representa-se a distribuição das variáveis que constituem a dimensão de perda de confiança (figuras 13 e 14). Estas variáveis, que enunciam estados psicológicos negativos, têm escalas de medida iguais às variáveis que constituem a dimensão de ansiedade e depressão¹³.

Figura 13 – Perda de confiança

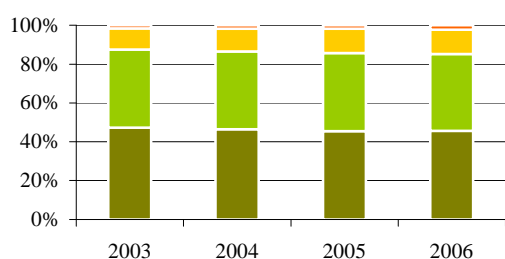
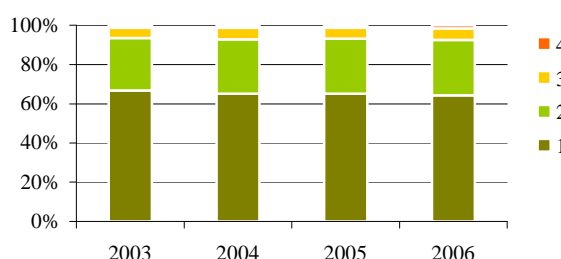


Figura 14 – Auto-estima



Ao contrário do que se observou relativamente às restantes variáveis, uma percentagem considerável dos inquiridos considera não ter falta de auto-estima ou de confiança.

Para além dos indicadores que constituem o GHQ-12, o BHPS fornece ainda dois indicadores de bem-estar psicológico, construídos a partir dos 12 itens. Apenas o primeiro destes indicadores será utilizado neste trabalho¹⁴. Este indicador é construído a partir da soma das

¹³ As questões colocadas aos inquiridos foram: “*Have you recently been losing confidence in yourself?*” e “*Have you recently been thinking of yourself as a worthless person?*”.

¹⁴ O segundo indicador derivado da escala GHQ-12 varia entre 0 (menos perturbado(a)) e 12 (mais perturbado(a)), e é construído também a partir da soma dos 12 itens, agora recodificados em dois pontos, onde o primeiro agrega as respostas no primeiro e segundo pontos da escala de medida das variáveis originais, e o

respostas dos indivíduos à escala do *GHQ-12*, recodificadas de forma a variarem entre 0 e 3. Este indicador varia entre 0 (menos perturbado(a)) e 36 (mais perturbado(a))¹⁵. As médias e desvios-padrão desta variável nos quatro períodos em análise apresentam-se nas tabelas 1 e 2.

Tabela 1 – Médias e desvios-padrão do indicador único SWB nos quatro períodos considerados para indivíduos do sexo masculino

	Média	Desvio-padrão
2003	10,117	4,721
2004	10,335	4,884
2005	10,281	4,875
2006	10,416	4,945

Tabela 2 – Médias e desvios-padrão do indicador único SWB nos quatro períodos considerados para indivíduos do sexo feminino

	Média	Desvio-padrão
2003	11,640	5,622
2004	11,821	5,641
2005	11,933	5,609
2006	12,017	5,704

Na análise dos modelos de *LGCM* que se segue procuraremos perceber se, por um lado, a evolução do indicador único de bem-estar psicológico é semelhante entre os indivíduos que constituem a nossa amostra, e, por outro, se uma estrutura que considera três dimensões de bem-estar se ajusta melhor. Neste sentido, para além da análise a partir das variáveis tal como foram medidas (em escalas ordinais), faremos ainda a modelação a partir de uma codificação dos três indicadores segundo o mesmo critério adoptado para a construção do indicador único. Assim, construíram-se três novas variáveis – ansiedade e depressão, disfunção social e perda de confiança – a partir da soma das respostas dos indivíduos aos índices, recodificadas numa escala de 0 a 3. Pela estrutura que se apresentará no capítulo 4¹⁶, o indicador de ansiedade e depressão foi construído somando as respostas às questões relativas à perda de sono, constantemente sob pressão, dificuldade em superar problemas e sentimento de tristeza ou depressão. O indicador relativo à disfunção social resulta da soma das variáveis relativas à concentração, desempenho de papel útil, capacidade em tomar decisões, capacidade em apreciar actividades diárias, capacidade em encarar problemas, e felicidade em geral. Por fim, o indicador de perda de confiança resulta da soma das respostas às questões relativas à perda de confiança e à auto-estima.

segundo as respostas nos dois últimos pontos. Cf. Taylor *et al.* (2008: App2-7). Tradução nossa de “*the less distressed*” e “*the most distressed*”.

¹⁵ Cf. Taylor *et al.* (2008: App2-7). Tradução idêntica.

¹⁶ No diagrama do modelo conceptual utilizado, apresentado no ponto 10.2 dos anexos, facilmente se percebe que variáveis foram somadas para obter cada um dos indicadores.

Figura 15 – Médias do indicador de bem-estar psicológico e das dimensões de ansiedade e depressão, disfunção social e perda de confiança para indivíduos do sexo masculino

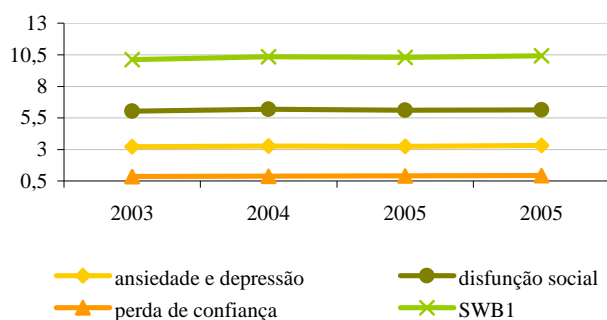
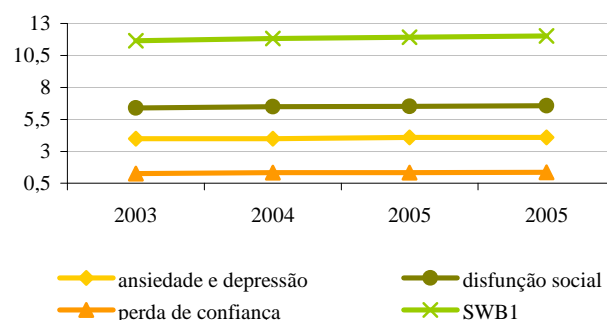


Figura 16 – Médias do indicador de bem-estar psicológico e das dimensões de ansiedade e depressão, disfunção social e perda de confiança para indivíduos do sexo feminino



Assim, o indicador relativo à dimensão de ansiedade e depressão varia entre 0 e 12, o indicador de disfunção social entre 0 e 18, e o indicador de perda de confiança entre 0 e 6¹⁷.

As figuras 15 e 16 representam as médias em cada uma das dimensões, bem como no indicador único de bem-estar ao longo do período em análise. A leitura dos gráficos mostra, por um lado, que as mulheres manifestam maiores graus de insatisfação média. Observamos ainda, tanto relativamente aos homens como às mulheres, uma ligeira tendência para o aumento da insatisfação média ao longo dos períodos considerados.

Clark (2002) e Clark e Oswald (2002, 2006) apresentam estudos longitudinais do *subjective well-being* utilizando o indicador único já referido (SWB). Nestes estudos, os autores sugerem a utilização de um vasto conjunto de variáveis explicativas para as variações do bem-estar psicológico observadas, como o sexo, o estado de saúde, o rendimento, as habilitações literárias, entre várias outras. Embora tendo consciência de que todas aquelas variáveis seriam potencialmente relevantes para uma melhor compreensão do fenómeno em análise, neste estudo optou-se por considerar apenas variáveis explicativas não variantes no tempo¹⁸. Assim, procuraremos perceber se o sexo e a idade contribuem de alguma forma para explorar as diferenças no bem-estar manifestado pelos indivíduos. Para além do sexo e da idade, consideraremos ainda o indicador relativo ao estado de saúde percebido.

As figuras 17 e 18 representam o estado de saúde percebido por homens e mulheres ao longo dos quatro anos em análise.

¹⁷ Também aqui o valor mais baixo indica menor grau de perturbação.

¹⁸ Na secção 5.2., dedicada à exposição teórica dos modelos de trajectória latente condicionados, aprofundaremos as questões relativas à introdução de variáveis explicativas variantes e não variantes no tempo.

Figura 17 – Distribuição do estado de saúde dos homens ao longo dos períodos considerados

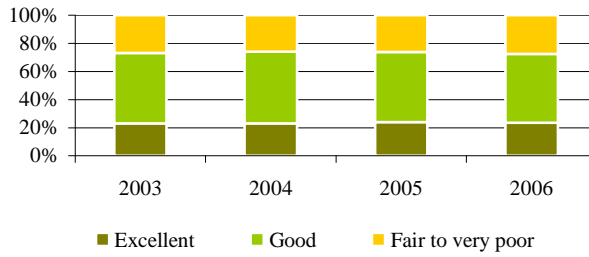


Figura 18 – Distribuição do estado de saúde das mulheres ao longo dos períodos considerados



Comparando os estados de saúde percebidos de homens e mulheres, verificamos que as mulheres parecem considerar-se menos saudáveis. Embora nos dois casos a maior parcela dos indivíduos afirme considerar ter um bom estado de saúde, mais mulheres avaliam o seu estado de saúde negativamente (ver figuras 17 e 18).

Apesar de o estado de saúde ser uma variável que pode variar ao longo do tempo, a sua distribuição, quer para homens quer para mulheres, aponta para alguma estabilidade nas respostas. Assim, e para efeitos de modelação, consideraremos o estado de saúde como invariante ao longo do tempo.

4. ESTRATÉGIA ADOPTADA PARA A ANÁLISE LONGITUDINAL DO *GHQ-12*

Antes de delinear a estratégia de abordagem à modelação longitudinal do *GHQ-12*, apresentamos a solução dos modelos de análise factorial confirmatória, comparando uma estrutura unifactorial com uma estrutura trifactorial. O objectivo é o de perceber se a estrutura factorial proposta por Graetz (1991) e reproduzida por outros autores tem de facto um melhor ajustamento do que uma estrutura mais simples do *GHQ-12*.

As análises factoriais confirmatórias foram realizadas para o último momento de observação, correspondente ao ano de 2006. A estimação dos modelos foi realizada em LISREL, utilizando os métodos adequados ao tratamento de variáveis ordinais¹⁹. Na tabela 3 apresentam-se as estimativas dos pesos factoriais num modelo considerando um factor e num modelo que reproduz a estrutura proposta por Graetz (1991), que considera três dimensões de debilidade psicológica: ansiedade e depressão; disfunção social; e perda de confiança.

Tabela 3 – Estimativas estandardizadas dos pesos factoriais no modelo unifactorial e trifactorial do *GHQ-12*

	Modelo unifactorial	Modelo trifactorial		
		Ansiedade e depressão	Disfunção social	Perda de confiança
Perda de sono por preocupações	0,69	0,75	---	---
Constantemente sob pressão	0,71	0,80	---	---
Problema em superar dificuldades	0,80	0,85	---	---
Sentimento de tristeza ou depressão	0,84	0,88	---	---
Facilidade de concentração	0,70	---	0,75	---
Sentimento de utilidade	0,61	---	0,72	---
Capacidade de tomar decisões	0,67	---	0,80	---
Capacidade de encarar problemas	0,71	---	0,77	---
Capacidade de apreciar actividades diárias	0,72	---	0,82	---
Sentimento de felicidade	0,77	---	0,81	---
Perda de confiança em si próprio	0,88	---	---	0,93
Perda de auto-estima	0,80	---	---	0,89
Correlações entre os factores no modelo trifactorial				
Ansiedade e depressão		---		
Disfunção social		0,72	---	
Perda de confiança		0,84	0,73	---

¹⁹ Estes métodos serão devidamente expostos no capítulo 5, relativo aos métodos aplicados.

Verificamos que as estimativas estandardizadas são superiores no modelo com uma estrutura trifactorial. Verificamos ainda que a estrutura de correlações entre os três factores é bastante forte, tal como sugeriam Cheung (2002), Gao *et al.* (2004), Shevlin e Adamson (2005) e Campbell e Knowles (2007).

Na tabela 4 apresentam-se as medidas de bondade do ajustamento dos dois modelos de análise factorial confirmatória. As conclusões a retirar são idênticas àquelas apresentadas por vários dos autores referidos: um modelo considerando uma estrutura tripartida do *GHQ-12* resulta num melhor ajustamento. A leitura dos valores obtidos para o critério de informação *AIC* nos dois modelos, que permite a sua comparação, reforça a ideia de que um modelo com três factores é um melhor modelo²⁰. Todavia, as correlações entre os três factores são bastante elevadas.

Tabela 4 – Medidas de bondade do ajustamento dos modelos unifactorial e trifactorial

	Modelo unifactorial	Modelo trifactorial
CFI	0,97	0,99
NNFI²¹	0,97	0,98
RMSEA	0,100	0,077
AIC	2623,100	1494,040

O modelo de análise factorial de segunda ordem, concebendo um factor latente subjacente à ansiedade e depressão, à disfunção social e à perda de confiança, parece fazer todo o sentido do ponto de vista conceptual. Sendo três os factores de primeira ordem, as estruturas de pesos factoriais e as medidas de bondade do ajustamento são idênticas para os modelos de primeira e de segunda ordem. As estimativas estandardizadas obtidas para a estrutura da relação entre o factor de segunda ordem e os factores de primeira ordem apresentam-se na tabela 5:

²⁰ Apresentaremos detalhadamente os critério para a comparação de modelos no ponto 5.5.

²¹ A medida de ajustamento *Non-Normed Fit Index* é equivalente ao índice TLI (*Tucker-Lewis Index*), apresentado na secção 5.5.

Tabela 5 – Estimativas estandardizadas da estrutura de relações entre os factores de primeira primeira ordem e o factor de 2ª ordem.

Factor de segunda ordem: debilidade psicológica	
Ansiedade e depressão	0,91
Disfunção social	0,79
Perda de confiança	0,92

Verificamos que as estimativas são bastante elevadas, espelhando uma estrutura de fortes correlações entre os factores de primeira ordem.

A abordagem adoptada à modelação longitudinal do *GHQ-12*, utilizando modelos de trajectória latente, parte da hipótese de que, tal como tem sido defendido por alguns dos autores referidos, um modelo teórico com três componentes é mais adequado para a análise da evolução do bem-estar psicológico dos indivíduos. A aproximação à modelação dos três factores segue um percurso com início num modelo considerando o índice agregado de bem-estar disponibilizado pelo BHPS, que pressupõe a unidimensionalidade do *GHQ-12*, até um modelo complexo, que modela a trajectória de um factor de segunda ordem que dá coerência ao modelo tripartido, a partir dos itens da escala tal como foram medidos. Assim, procuraremos comparar cinco estruturas diferentes de trajectória do *GHQ-12*.

O primeiro modelo a considerar é o modelo de trajectória do indicador obtido pela soma das respostas dos inquiridos aos 12 itens do *GHQ-12*. De forma a comparar este modelo com um modelo tripartido, apresentaremos modelos de trajectória dos três factores de bem-estar construídos a partir da mesma lógica aplicada ao indicador *SWB*, *i.e.*, somando as respostas dos indivíduos aos itens que estão associados a cada um dos constructos. Considerando as três medidas repetidas relativas às dimensões do *GHQ-12* de *ansiedade e depressão (AD)*, *disfunção social (SD)* e *perda de confiança (LC)*, recodificadas a partir dos *scores*, analisaremos duas estruturas distintas: a primeira considerando três processos correlacionados, *i.e.*, um processo para cada uma das dimensões enunciadas, correlacionando os constructos que definem as diferentes trajectórias; e um segundo considerando o factor latente de *debilidade psicológica (PM)* como dimensão subjacente às três variáveis de bem-estar.

Com o objectivo de comparar os modelos que consideram os *scores* quer do factor único de bem-estar, quer das três dimensões que o constituem, modelaremos ainda as trajectórias considerando as medidas repetidas ordinais. Apresentaremos um modelo com os três

processos correlacionados, *i.e.*, modelando as trajectórias de cada um dos factores e correlacionando os respectivos interceptos e declives, e outro concebido sob o pressuposto de que as três dimensões de bem-estar estruturam um factor de ordem superior. Este último modelo, inovador no contexto dos modelos de trajectória latente, considera um intercepto e um declive para a trajectória do factor de *debilidade psicológica (PM)*, modelando a trajectória deste factor de segunda ordem.

Para os modelos mais relevantes do ponto de vista dos objectivos traçados, serão introduzidas variáveis explicativas invariantes ao longo dos períodos analisados, procurando com isso perceber se características como o sexo, a idade ou o estado de saúde influenciam a trajectória do bem-estar manifestado pelo indivíduos²².

A análise dos resultados dos modelos estimados será apresentada no capítulo 6. Antes, no capítulo 5, apresenta-se a exposição teórica detalhada dos modelos de trajectória latente.

²² Consideram-se invariantes variáveis com variação constante ao longo do tempo. Quanto à variável relativa ao estado de saúde percebido, vimos já que a sua evolução é bastante estável.

5. MODELOS DE TRAJECTÓRIA LATENTE

A análise das trajectórias de bem-estar psicológico é aqui tratada a partir do quadro metodológico dos modelos conhecidos como *Latent Growth Curve Models (LGCM)*.

A ideia subjacente a este método de modelação estatística é a de que a verdadeira trajectória ou evolução temporal dos indicadores em análise (percepções, valores, etc.) não é observada, mas latente. Por outras palavras, considera-se que, a cada variável observada, está associada uma trajectória latente ao longo de um determinado período. Estas trajectórias são estimadas a partir da estrutura de médias e de associações entre as variáveis medidas em vários momentos. Assim, não modelamos a evolução das variáveis observadas, mas antes os processos que lhes estão subjacentes.

Por assentarem a sua estrutura e método na lógica dos modelos de equações estruturais, os *LGCM* partilham algumas das suas vantagens e desvantagens (*cf.* Duncan *et al.*, 2006). Como vantagens, salientam-se a capacidade para testar a adequabilidade global do modelo em estudo, permitindo avaliar a qualidade do ajustamento modelo-dados, bem como a utilização de variáveis latentes medidas por múltiplos indicadores, permitindo a modelação de fenómenos multidimensionais e complexos. Por outro lado, e no caso de variáveis métricas, este método pressupõe a normalidade multivariada das variáveis em estudo, e requer a utilização de grandes amostras²³.

Como principal vantagem dos modelos de trajectória latente devemos salientar a capacidade para perceber a mudança e evolução dos fenómenos em estudo, quer ao nível individual, permitindo medir as alterações nas trajectórias individuais, quer ao nível global, permitindo apreender diferenças entre aquelas trajectórias. Ainda, pela possibilidade de se introduzirem variáveis explicativas variantes e invariantes no tempo, é possível explicar a mudança neste dois níveis – intra e inter individual.

Bollen e Curran (2006: 8-9) sugerem que, recorrendo à modelação no quadro dos *LGCM*, o investigador dê resposta a três questões fundamentais: (1) qual a trajectória do fenómeno para o conjunto dos indivíduos; (2) existe heterogeneidade nas trajectórias individuais, implicando que se considerem trajectórias distintas; (3) existindo heterogeneidade, é possível identificar

²³ A este respeito Duncan *et al.* (2006: 4-5) referem estudos recentes onde se demonstra que, em modelos básicos de *LGCM*, a utilização de grandes amostras não é condição necessária, mas antes dependente da complexidade do fenómeno em estudo e do número de momentos temporais considerado.

variáveis que permitam explicar as diferenças entre os indivíduos. A primeira questão remete para a mudança ao nível do conjunto dos indivíduos, avaliada a partir da variação da trajectória média. A segunda questão aponta para a necessidade de avaliar se existe ou não variabilidade significativa das trajectórias individuais em torno da trajectória média, *i.e.*, se a trajectória média reflecte todas as trajectórias individuais. Caso tal não se verifique, é necessário perceber as diferenças. A terceira questão enunciada remete exactamente para a capacidade de explorar, explicar e prever as diferenças das trajectórias individuais relativamente à trajectória média. Esta é a ideia subjacente aos modelos de trajectória latente condicionada.

5.1. MODELOS DE TRAJECTÓRIA LATENTE NÃO CONDICIONADA

Conceptualmente, os *LGCM* assentam na ideia de que a trajectória dos fenómenos em estudo é latente, não observada, mas antes obtida a partir das variáveis recolhidas. Assim, os parâmetros que definem uma trajectória de evolução de um qualquer fenómeno (o intercepto e o declive, no caso de uma trajectória linear) são variáveis latentes que, com a adequada parametrização, espelham a mudança subjacente aos indicadores medidos em diversos pontos temporais.

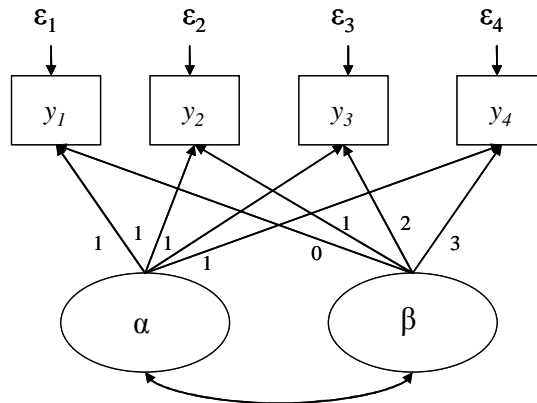
A perspectiva que subjaz aos modelos com variáveis latentes é a de que existem uma ou mais dimensões, constructos, factores, que são evidenciados pela estrutura de associações entre as variáveis observadas. Se tomarmos a trajectória das variáveis observadas ao longo do tempo como um factor não observado, facilmente se percebe o paralelismo na conceptualização dos *LGCM* e dos modelos de equações estruturais (*SEM*²⁴).

Como veremos, uma importante diferença entre os *SEM* e os modelos de trajectória latente está na matriz dos pesos – Λ – que estruturam a relação entre os factores latentes e as variáveis manifestas.

Far-se-á agora uma abordagem mais teórica e metodológica dos modelos de trajectória latente, ainda sem considerar a inclusão de variáveis explicativas. Na figura 19 apresenta-se o diagrama de uma trajectória linear da variável *y*, medida em quatro momentos consecutivos.

²⁴ A sigla resulta da denominação anglo-saxónica de *structural equation models*.

Figura 19 – Modelo de trajectória linear não condicionada da variável y ao longo de 4 momentos



Formalmente, um modelo genérico de *LGCM* não condicionado é dado por

$$y_{it} = \alpha_i + \lambda_t \beta_i + \varepsilon_{it} \quad (1)$$

Onde y_{it} representa o valor da variável y para o i -ésimo caso no momento t ; α_i é o intercepto aleatório para o caso i ; β_i é o declive aleatório para o caso i ; e ε_{it} representa a parcela de erro da trajectória traçada para o elemento i no momento t . λ_t representa a ‘passagem do tempo’. Assim, para um modelo de trajectória linear, assume-se que $\lambda_t = t - 1$, de forma a que o valor esperado do intercepto represente a média da trajectória no momento inicial (em que $\lambda_t = 0$), incorporando-se a trajectória linear a partir da relação entre as variáveis manifestas e o declive. Como se referiu, este é um dos principais pontos de contraste entre os modelos com variáveis latentes e os *LGCM*. Se nos primeiros a matriz dos pesos é estimada, nos modelos de trajectória latente esta é fixada de início.

Para a estimação do modelo definido em (1), alguns pressupostos devem ser enunciados. Como acontece na análise de regressão, pressupõe-se que a parcela do erro tem média nula (*i.e.*, $E(\varepsilon_{it}) = 0, \forall i, t$), e que as variáveis independentes, neste caso os interceptos e declives individuais, não estão correlacionadas com a parcela do erro. Mais formalmente, $\text{cov}(\varepsilon_{it}, \alpha_i) = 0$ e $\text{cov}(\varepsilon_{it}, \beta_i) = 0$, para qualquer i e $t = 1, 2, \dots, T$. Por outro lado, de forma a que os erros não estejam correlacionados ao longo do tempo, pressupõe-se ainda que $\text{cov}(\varepsilon_{it}, \varepsilon_{it+s}) = 0$, para qualquer s diferente de zero.

Relativamente aos interceptos e declives aleatórios, impõe-se que estes não estejam correlacionados entre casos, *i.e.*, que $\text{cov}(\alpha_i, \alpha_j) = 0$, $\text{cov}(\beta_i, \beta_j) = 0$ e $\text{cov}(\alpha_i, \beta_j) = 0$, para $i \neq j$. Por fim, pressupõe-se ainda que a parcela do erro da trajectória de cada indivíduo não está correlacionada com a parcela do erro da trajectória de nenhum dos restantes indivíduos, *i.e.*, $\text{cov}(\varepsilon_{it}, \varepsilon_{jt+s}) = 0$, para $i \neq j$ e qualquer s .

Os parâmetros de interesse num modelo de *LGCM* são as médias dos interceptos, as médias dos declives, e as variâncias e covariância dos efeitos aleatórios (intercepto e declive). Formalmente, definem-se o intercepto e o declive como:

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

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

A equação do intercepto individual (α_i) é função da média dos interceptos de todos os casos (μ_α) e uma perturbação ($\zeta_{\alpha i}$). Da mesma forma, a equação do declive individual (β_i) é função da média dos declives de todos os indivíduos (μ_β) e uma perturbação ($\zeta_{\beta i}$). Isto significa que se introduzem as médias dos interceptos e declives como medidas de tendência central nas trajectórias individuais.

As perturbações $\zeta_{\alpha i}$ e $\zeta_{\beta i}$ têm média zero, variâncias $\psi_{\alpha\alpha}$ e $\psi_{\beta\beta}$ e covariância dos interceptos e declives $\psi_{\alpha\beta}$. Pressupõe-se que estas perturbações não estão correlacionadas com a parcela de erro do modelo genérico apresentado em (1). Porque, no modelo não condicional, as equações do intercepto e do declive individuais apresentadas em (2) e (3) são função apenas do intercepto médio e do declive médio (respectivamente), a variância de α é equivalente à variância de ζ_α e a variância de β é equivalente à variância de ζ_β . O mesmo não acontece no modelo condicional, como se verá na secção 5.2.

Se a maioria dos indivíduos tiver interceptos muito semelhantes e declives também muito semelhantes, espera-se que a variância do intercepto e do declive seja pequena. No caso mais extremo, estas variâncias são zero, e os interceptos e declives de todos os casos são iguais. Inversamente, maior diversidade nos interceptos e declives individuais levam a maiores valores de $\psi_{\alpha\alpha}$ e $\psi_{\beta\beta}$.

Combinando as equações da trajectória, do intercepto e do declive numa única equação, temos

$$y_{it} = (\mu_{\alpha} + \lambda_t \mu_{\beta}) + (\zeta_{\alpha i} + \lambda_t \zeta_{\beta i} + \varepsilon_{it}) \quad (4)$$

Este modelo pode ser denominado de modelo combinado ou forma reduzida da equação do modelo de trajectória latente²⁵ (Bollen e Curran, 2006: 21). A denominação de modelo combinado resulta da junção das equações acima enunciadas em (1), (2) e (3). De facto, a equação enunciada em (4) resulta da reformulação do modelo de trajectória do i -ésimo indivíduo no momento t apresentada em (1), agora definida em função de uma estrutura de médias dos interceptos e dos declives, e de um termo aleatório residual.

A expressão de forma reduzida resulta da substituição dos coeficientes aleatórios α_i e β_i , endógenos, pelos seus determinantes exógenos e perturbações²⁶. Note-se que a parcela de erro compósito é heteroscedástica ao longo do tempo, devido à presença de $\lambda_t \zeta_{\beta i}$, que tem uma variância que depende de λ_t . O primeiro termo entre parêntesis na equação é referido como componente fixa, e o segundo como componente aleatória. A componente fixa representa a estrutura média, e a componente aleatória representa as várias fontes de variabilidade individual.

Os modelos de trajectória latente modelam explicitamente tanto as estruturas de médias como de variâncias e covariâncias. No entanto, impõe-se uma estrutura restritiva a estas médias. Mais concretamente, fixam-se a zero os interceptos das medidas repetidas e estimam-se as médias dos factores da trajectória latente. Assim, a estrutura de médias das medidas repetidas implícitas ao modelo é inteiramente determinada pelas médias dos factores da trajectória latente.

Por fim, a estrutura de covariâncias é determinada da forma usual, as variâncias dos erros são estimadas para as medidas repetidas e são estimadas variâncias e covariâncias do intercepto e do declive da trajectória latente. A variância dos erros das medidas repetidas representa a parcela de variância destas variáveis que não é explicada pela trajectória latente subjacente.

Retomando a questão relativa à estrutura de pesos que reflecte a passagem do tempo, recordamos que esta é inicialmente definida. Na matriz dos Λ , estabelece-se que o peso do

²⁵ Tradução de *combined model* e de *reduced-form equation*.

²⁶ A endogeneidade de α_i e de β_i resultava da sua correlação com a parcela do erro. Definindo intercepto e declive como função das respectivas médias e uma perturbação, tornamos a primeira parcela da equação (4) – componente fixa – exógena.

intercepto em cada uma das medidas repetidas é 1, de forma a que a sua influência sobre as variáveis manifestas se mantenha constante ao longo do tempo.

Paralelamente, é possível introduzir uma variedade de estruturas de pesos para o declive, de forma a evidenciar a passagem do tempo entre as observações. É possível introduzir factores latentes que representem trajectórias não lineares. Por exemplo, a introdução de uma trajectória quadrática obtém-se a partir da relação entre um termo quadrático e as medidas repetidas, onde λ_t tomará valores que espelhem esta trajectória. Neste caso, $\lambda_t = (t-1)^2$, de forma que o termo introduzido represente a trajectória considerada.

Num modelo linear onde se consideram observações recolhidas em momentos temporais igualmente intercalados (*e.g.*, de ano a ano, como acontece relativamente aos dados utilizados neste estudo), estabelece-se que, para o declive, $\lambda_t = 0, 1, 2, \dots, T-1$. Definir o peso no momento inicial a zero permite que o intercepto possa reflectir o valor médio de y para o primeiro período. A média do intercepto representa a média no momento inicial e a média do declive representa a taxa média de mudança nas medidas repetidas para todos os indivíduos na amostra.

Refira-se que a codificação da passagem do tempo deve ter em conta as questões substantivas levantadas pelo investigador. Como referem Preacher *et al.* (2008: 11), “*the zero point of the time scale should be placed at a meaningful occasion of measurement, in part because this choice determines the point in time at which to interpret the mean intercept, intercept variances, and intercept-slope covariance*”.

Para identificação de um modelo não condicional de *LGCM*, aliás como acontece quando utilizamos qualquer outro método de estimação, devemos ter tantas ou mais estimativas conhecidas quanto os parâmetros desconhecidos. Num modelo de trajectória latente não condicionada conhecemos as médias, as variâncias e as covariâncias de y [$E(y_{it})$, $\text{var}(y_{it})$ e $\text{cov}(y_{it}, y_{i,t-s})$], para $s > 0$ e $t-s \geq 1$. Para um modelo de trajectória linear, os parâmetros desconhecidos são a média do intercepto e do declive [μ_α e μ_β , respectivamente], a variância dos erros [$\text{var}(\varepsilon_{it})$], a variância do intercepto e do declive e a sua covariância [$\psi_{\alpha\alpha}$, $\psi_{\beta\beta}$ e $\psi_{\alpha\beta}$, respectivamente]. Como vimos, assume-se que a matriz dos pesos longitudinais é conhecida, tomando valores, no caso de uma tendência linear, $\lambda_t = t-1$.

Assim, num modelo de *LGCM* não condicional, temos uma média amostral para cada período (\bar{y}_t), e variâncias e covariâncias de y_{it} , resultando num total de $\frac{1}{2}T(T+3)$ medidas conhecidas.

Relativamente ao número de parâmetros desconhecidos, temos NT parâmetros para a $\text{var}(\varepsilon_{it})$ e $\frac{1}{2}K(K+3)$ parâmetros para μ_α , μ_β , $\psi_{\alpha\alpha}$, $\psi_{\beta\beta}$ e $\psi_{\alpha\beta}$ (onde K representa o número total de variáveis latentes que descrevem a trajectória²⁷). É fácil perceber que, sem a imposição de alguma restrição, o modelo não poderia estar identificado. Assim, para além da restrição imposta a λ_t , assume-se ainda que os diversos casos têm a mesma variância do erro entre eles, podendo esta variar apenas ao longo do tempo [*i.e.*, $\text{var}(\varepsilon_{it}) = \text{var}(\varepsilon_t)$].

Impostas as restrições, passamos a ter menos parâmetros desconhecidos: se $\text{var}(\varepsilon_{it}) = \text{var}(\varepsilon_t)$, temos apenas T variâncias dos erros a estimar; se considerarmos um modelo linear, $K = 2$, significa que temos mais 5 parâmetros desconhecidos, μ_α , μ_β , $\psi_{\alpha\alpha}$, $\psi_{\beta\beta}$ e $\psi_{\alpha\beta}$. Assim, no caso de um modelo linear de trajectória latente não condicionada, os parâmetros desconhecidos são $T + 5$.

Para que um modelo de *LGCM* esteja identificado, temos de considerar pelo menos três momentos temporais. Facilmente se percebe que, com apenas duas medidas repetidas, não é possível estimar um modelo desta natureza. Teríamos um menor número de valores conhecidos relativamente aos parâmetros a estimar²⁸.

Considerando três períodos de tempo, asseguramos as condições necessárias (mas não suficientes) para a identificação do modelo (passamos a ter 8 parâmetros desconhecidos e 9 valores conhecidos) – neste caso, teríamos apenas um grau de liberdade. Aumentando o número de observações repetidas, aumenta também o número de graus de liberdade.

De forma genérica, o número mínimo de momentos temporais a considerar para que um modelo de trajectória latente esteja identificado deve ser equivalente ao número de factores latentes que definem a trajectória mais um. Assumindo que os parâmetros estimados incluem

²⁷ Por exemplo, $K = 1$ para um intercepto; $K = 2$ se existirem interceptos e declives aleatórios; $K = 3$ quando um termo quadrático aleatório é introduzido, etc. Teríamos mais T parâmetros desconhecidos caso não se impusesse a restrição à matriz λ_t .

²⁸ Num modelo considerando 2 waves teríamos 5 parâmetros conhecidos dados por $\left[\frac{1}{2}T(T+3) = 5\right]$ e 7 parâmetros desconhecidos [2 variâncias dos erros mais 5 parâmetros (μ_α , μ_β , $\psi_{\alpha\alpha}$, $\psi_{\beta\beta}$ e $\psi_{\alpha\beta}$)].

apenas as médias, variâncias e covariâncias dos factores de trajectória latente, e variâncias dos erros, este cálculo é sempre válido (Preacher *et al.*, 2008: 12)²⁹.

Num modelo de *LGCM* com medidas repetidas em escalas métricas, é utilizado o estimador de máxima verosimilhança geralmente utilizado em modelos de equações estruturais, que contém várias características que o tornam especialmente útil e adequado: é consistente, é assintoticamente não enviesado, é assintoticamente normal, e assintoticamente eficiente. Sinteticamente, a ideia subjacente à estimação é a de seleccionar valores para os parâmetros do modelo que reproduzem o mais próximo possível as médias e matrizes de covariância das variáveis observadas. As médias amostrais ($\bar{\mathbf{y}}$) e a matriz de covariâncias amostrais (\mathbf{S}) substituem o vector de médias populacionais ($\boldsymbol{\mu}$) e a matriz de covariâncias populacionais ($\boldsymbol{\Sigma}$), que são desconhecidos. Representando os parâmetros estimados do modelo [*i.e.*, $\hat{\mu}_\alpha$, $\hat{\mu}_\beta$, $VAR(\varepsilon_t)$, $\hat{\psi}_{\alpha\alpha}$, $\hat{\psi}_{\beta\beta}$ e $\hat{\psi}_{\alpha\beta}$] por $\hat{\boldsymbol{\theta}}$, o objectivo é o de escolher valores de $\hat{\boldsymbol{\theta}}$, de forma a que $\boldsymbol{\mu}(\hat{\boldsymbol{\theta}})$ fique próximo de $\bar{\mathbf{y}}$ e $\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})$ próximo de \mathbf{S} . A minimização da função de máxima verosimilhança procura os valores de $\hat{\boldsymbol{\theta}}$ que minimizam as ‘distâncias’ entre os parâmetros de interesse do modelo e os parâmetros amostrais conhecidos³⁰.

5.1.1. Estrutura matricial do modelo de trajectória latente não condicionada

O modelo de trajectória latente pode ser estimado utilizando o enquadramento dos modelos de equações estruturais. Considere-se o vector \mathbf{y} de dimensão ($T \times 1$) que contém o conjunto de T medidas repetidas da variável y . É possível representar \mathbf{y} em termos matriciais da seguinte forma:

$$\mathbf{y} = \boldsymbol{\Lambda}\boldsymbol{\eta} + \boldsymbol{\varepsilon} \quad (5)$$

²⁹ Os autores discutem ainda qual será o número de medidas repetidas a considerar, tendo como valor mínimo o necessário para a identificação do modelo, afirmando que não há nenhuma regra definitiva, apenas sugestões propostas por alguns autores (Preacher *et al.*, 2008: 12).

³⁰ A função de máxima verosimilhança é dada por

$F_{ML} = \ln |\boldsymbol{\Sigma}(\boldsymbol{\theta})| - \ln |\mathbf{S}| + [\boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta})\mathbf{S}]^T - \mathbf{p} - [\bar{\mathbf{y}} - \boldsymbol{\mu}(\boldsymbol{\theta})]' \boldsymbol{\Sigma}^{-1}(\boldsymbol{\theta}) [\bar{\mathbf{y}} - \boldsymbol{\mu}(\boldsymbol{\theta})]$. F_{ML} tem um mínimo em zero e não tem máximo. Aproxima-se do seu mínimo quando $\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}}) \rightarrow \mathbf{S}$ e $\boldsymbol{\mu}(\hat{\boldsymbol{\theta}}) \rightarrow \bar{\mathbf{y}}$, e é zero quando estas relações são idênticas ou quando o *LGCM* está perfeitamente identificado. Para uma leitura mais detalhada da estimação de máxima verosimilhança no contexto do *LGCM* com variáveis manifestas métricas ver, *e.g.*, Bollen e Curran (2006: 39-42).

Onde Λ é uma matriz ($T \times k$) de pesos; η é um vector ($k \times 1$) dos k factores latentes; e ε é um vector ($T \times 1$) dos resíduos. Os elementos da equação (5), para um modelo de trajectória linear ajustado a T medidas repetidas, são:

$$\begin{pmatrix} y_{i1} \\ y_{i2} \\ \dots \\ y_{iT} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ \dots & \dots \\ 1 & T-1 \end{pmatrix} \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} + \begin{pmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \dots \\ \varepsilon_{iT} \end{pmatrix} \quad (6)$$

Assim, cada observação repetida de y para o indivíduo i no momento t é uma combinação de um intercepto aleatório e de um declive linear, mais um resíduo do indivíduo ao longo dos T momentos. η pode ser definido em notação matricial da seguinte forma:

$$\eta = \mu_\eta + \zeta \quad (7)$$

onde μ_η é um vector ($k \times 1$) das médias dos factores de trajectória e ζ é um vector ($k \times 1$) dos resíduos. Para o modelo de trajectória linear definido em (6), os elementos das matrizes presentes na equação (7) são:

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \begin{pmatrix} \mu_\alpha \\ \mu_\beta \end{pmatrix} + \begin{pmatrix} \zeta_{\alpha i} \\ \zeta_{\beta i} \end{pmatrix} \quad (8)$$

A equação (7) pode ser incluída na equação (5), resultando na expressão da forma reduzida de y (equivalente à equação apresentada em (4)):

$$y = \Lambda (\mu_\eta + \zeta) + \varepsilon \quad (9)$$

As variâncias e covariâncias implícitas ao modelo de forma reduzida são dadas por:

$$\Sigma = \Lambda \Psi \Lambda' + \Theta_\varepsilon \quad (10)$$

Onde Σ representa a matriz de covariâncias dos y e Θ_ε representa a estrutura de covariâncias das parcelas de erro para as T medidas repetidas, de forma que:

$$\Theta_{\varepsilon} = \begin{pmatrix} \text{VAR}(\varepsilon_1) & 0 & 0 & \dots & 0 \\ 0 & \text{VAR}(\varepsilon_2) & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \dots & \dots & 0 & \dots & \dots \\ 0 & 0 & 0 & \dots & \text{VAR}(\varepsilon_T) \end{pmatrix} \quad (11)$$

Esta matriz é equivalente à matriz θ_{δ} dos modelos de equações estruturais. Embora se assuma que os erros não estão correlacionados, se o modelo continuar a estar identificado é possível permitir alguma correlação entre erros das medidas repetidas³¹.

A matriz Ψ representa a matriz de covariâncias dos erros ζ entre os factores da trajectória latente, e é equivalente à matriz Φ dos modelos de equações estruturais:

$$\Psi = \begin{pmatrix} \psi_{\alpha\alpha} & \psi_{\alpha\beta} \\ \psi_{\beta\alpha} & \psi_{\beta\beta} \end{pmatrix} \quad (12)$$

Nos modelos com trajectórias não condicionadas, onde não são incluídas variáveis explicativas para a variância dos interceptos e dos declives, a variância de η é igual à variância de ζ .

O valor esperado do modelo reduzido da trajectória é dado por $\mathbf{E}(\mathbf{y}) = \Lambda\boldsymbol{\mu}_{\eta}$.

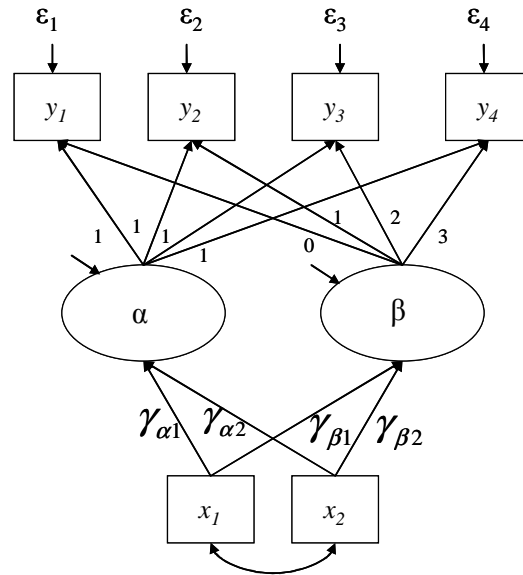
5.2. MODELOS DE TRAJECTÓRIA LATENTE CONDICIONADA

Para os modelos de trajectória latente condicionada, consideraremos apenas variáveis invariantes no tempo como variáveis explicativas, ou seja, variáveis que, podendo assumir valores diferentes entre indivíduos, se mantêm constantes ao longo do tempo (*e.g.* o sexo). Consideram-se ainda como variáveis invariantes aquelas cuja variação é constante ao longo do tempo (*e.g.* a variável idade), ou outras que, por variarem pouco, são modeladas como constantes (*e.g.* a percepção do estado de saúde).

³¹ Nos modelos com variáveis latentes e múltiplos indicadores (apresentados nas secções 5.3. e 5.4.), a matriz que representa a estrutura de covariâncias das medidas repetidas deixará de ser uma matriz diagonal, já que partiremos do pressuposto de que os erros das medidas repetidas estão correlacionados ao longo do tempo.

Um modelo de trajetória linear latente condicionada, considerando quatro momentos temporais e duas variáveis explicativas x_1 e x_2 , pode ser representado da seguinte forma:

Figura 20 – Modelo de trajetória linear latente condicionada da variável y ao longo de 4 períodos consecutivos



Tomemos como ponto de partida um modelo de trajetória latente não condicionada, tal como definido em (1):

$$y_{it} = \alpha_i + \lambda_t \beta_i + \varepsilon_{it} \quad (13)$$

Num *LGCM* com trajetória condicionada, para além das médias dos interceptos e declives, e dos desvios a estas medidas, as trajetórias são agora também função de variáveis explicativas daqueles desvios. Por outras palavras, o modelo condicionado pressupõe a inclusão de variáveis que permitem explorar interceptos e declives, permitindo testar potenciais influências nas trajetórias estimadas.

Partindo das equações dos interceptos e declives definidos em função das suas médias apresentadas em (2) e (3), temos agora que:

$$\alpha_i = \mu_\alpha + \gamma_{\alpha 1} x_{1i} + \gamma_{\alpha 2} x_{2i} + \dots + \gamma_{\alpha M} x_{Mi} + \zeta_{\alpha i} \quad (14)$$

$$\beta_i = \mu_\beta + \gamma_{\beta 1} x_{1i} + \gamma_{\beta 2} x_{2i} + \dots + \gamma_{\beta M} x_{Mi} + \zeta_{\beta i} \quad (15)$$

onde μ_α e μ_β representam, respectivamente, a média dos interceptos e declives de todos os casos quando x_1, x_2, \dots, x_M são zero. Os elementos x_1, x_2, \dots, x_M representam M variáveis explicativas dos interceptos e declives aleatórios, $\gamma_{\alpha 1}, \gamma_{\alpha 2}, \dots, \gamma_{\alpha M}$ são os coeficientes das variáveis x_1, x_2, \dots, x_M na equação dos interceptos aleatórios, e $\gamma_{\beta 1}, \gamma_{\beta 2}, \dots, \gamma_{\beta M}$ são os coeficientes das variáveis x_1, x_2, \dots, x_M na equação dos declives aleatórios. Os coeficientes associados às variáveis explicativas são interpretados como qualquer coeficiente de regressão, *i.e.*, controlando todas as outras variáveis explicativas, exprimem a variação esperada na variável dependente quando a independente aumenta uma unidade. Como se referiu inicialmente, embora seja possível introduzir variáveis com variação irregular ao longo do tempo, as equações apresentadas pressupõem que as variáveis explicativas são constantes em todos os momentos, variando apenas entre indivíduos³².

Agora, ao contrário do que acontecia nos modelos não condicionados, as variâncias das perturbações $\zeta_{\alpha i}$ e $\zeta_{\beta i}$ ($\psi_{\alpha\alpha}$ e $\psi_{\beta\beta}$) não representam as variâncias dos interceptos e declives aleatórios, mas antes variâncias condicionais. Mantêm-se os pressupostos de que as perturbações $\zeta_{\alpha i}$ e $\zeta_{\beta i}$ têm média nula, acrescentando-se agora que, para além de não estarem correlacionadas com ε_{it} , não estão correlacionadas com x_{1i} e x_{2i} .

Combinando a equação das trajectórias explicitada em (13) com as equações (14) e (15), obtemos o modelo combinado ou de forma reduzida para uma trajectória condicionada³³:

$$y_{it} = (\mu_\alpha + \lambda_t \mu_\beta) + (\gamma_{\alpha 1} + \lambda_t \gamma_{\beta 1}) x_{1i} + (\gamma_{\alpha 2} + \lambda_t \gamma_{\beta 2}) x_{2i} + \dots \quad (16)$$

$$\dots + (\gamma_{\alpha M} + \lambda_t \gamma_{\beta M}) x_{Mi} + (\zeta_{\alpha i} + \lambda_t \zeta_{\beta i} + \varepsilon_{it})$$

Vemos agora que a trajectória de y_{it} é função de um primeiro termo composto, coeficientes para $x_{1i}, x_{2i}, \dots, x_{Mi}$ compostos, que mudam com a passagem do tempo (expressa pela presença do termo λ_t), e um termo de erro composto, que também sofre alterações com a passagem do tempo. O conjunto do primeiro termo da equação e dos coeficientes associados às variáveis explicativas $x_{1i}, x_{2i}, \dots, x_{Mi}$ é denominado de componente fixa do modelo de trajectória latente;

³² As variáveis explicativas com variação ao longo do tempo no modelo de LGCM terão impacto quer nas variáveis de trajectória latentes (intercepto e declive), quer nas medidas repetidas. Para uma leitura detalhada dos procedimentos estatísticos envolvidos ver, *e.g.*, Bollen e Curran (2006: 188-228).

³³ Esta equação é o paralelo da equação enunciada em (4) para o modelo não condicional.

o termo de erro composto representa a sua componente aleatória. Tal como nos modelos de trajectória não condicionada, a componente fixa do modelo representa a estrutura de médias, e a componente aleatória as várias fontes de variabilidade individual.

Bollen e Curran (2006: 128-132) mostram que um modelo de *LGCM* estará identificado sempre que o modelo não condicional equivalente seja identificado e as variáveis explicativas sejam observadas (manifestas).

5.2.1. Estrutura matricial do modelo de trajectória latente condicionada

Também o modelo de trajectória condicionada pode ser expresso através de uma abordagem dos modelos de equações estruturais. Nesta secção repetiremos a análise apresentada para os modelos de trajectória não condicionada, agora considerando a presença de variáveis explicativas.

Retomemos a equação das trajectórias individuais na sua forma matricial, apresentada em (5):

$$\mathbf{y} = \Lambda\boldsymbol{\eta} + \boldsymbol{\varepsilon} \quad (17)$$

Considere-se agora a representação matricial das equações dos interceptos e declives enunciada em (14) e (15):

$$\begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix} = \begin{bmatrix} \mu_\alpha \\ \mu_\beta \end{bmatrix} + \begin{bmatrix} \gamma_{\alpha 1} & \gamma_{\alpha 2} & \dots & \gamma_{\alpha M} \\ \gamma_{\beta 1} & \gamma_{\beta 2} & \dots & \gamma_{\beta M} \end{bmatrix} \begin{bmatrix} x_{1i} \\ x_{2i} \\ \dots \\ x_{Mi} \end{bmatrix} + \begin{bmatrix} \zeta_{\alpha i} \\ \zeta_{\beta i} \end{bmatrix} \quad (18)$$

A sua forma matricial é dada por:

$$\boldsymbol{\eta} = \boldsymbol{\mu}_\eta + \boldsymbol{\Gamma}\mathbf{x} + \boldsymbol{\zeta} \quad (19)$$

Se substituirmos o elemento $\boldsymbol{\eta}$ na equação (17) pela sua expressão enunciada na equação (19), obtemos a expressão para o modelo combinado ou de forma reduzida:

$$\mathbf{y} = \Lambda(\boldsymbol{\mu}_\eta + \boldsymbol{\Gamma}\mathbf{x}) + \Lambda\boldsymbol{\zeta} + \boldsymbol{\varepsilon} \quad (20)$$

A expressão formal do modelo combinado em notação matricial considera T medidas repetidas e M variáveis explicativas. A partir da sua forma generalizada, facilmente se obtém a estrutura de médias implícita, bastando substituir os dois lados da equação (20) pelos seus valores esperados:

$$\boldsymbol{\mu}_y(\boldsymbol{\theta}) = \Lambda(\boldsymbol{\mu}_\eta + \Gamma\boldsymbol{\mu}_x) \quad (21)$$

onde $\boldsymbol{\mu}_y(\boldsymbol{\theta})$ representa a estrutura de médias implícita para \mathbf{y} no modelo e $\boldsymbol{\theta}$ contém os parâmetros desconhecidos que se pretendem estimar.

Uma vez que as variáveis agora introduzidas no modelo são exógenas, a sua estrutura de médias implícita é simplesmente a estrutura de médias observadas para as M variáveis. A partir daqui é simples deduzir a expressão para a estrutura de médias implícita a todas as variáveis manifestas, tal que $\boldsymbol{\mu} = \boldsymbol{\mu}(\boldsymbol{\theta})$:

$$\begin{bmatrix} \boldsymbol{\mu}_y \\ \boldsymbol{\mu}_x \end{bmatrix} = \begin{bmatrix} \Lambda(\boldsymbol{\mu}_\eta + \Gamma\boldsymbol{\mu}_x) \\ \boldsymbol{\mu}_x \end{bmatrix} \quad (22)$$

Analogamente, é possível obter a equação das variâncias e covariâncias das variáveis observadas como função dos parâmetros contidos no vector $\boldsymbol{\theta}$, tal que $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\theta})$, onde $\boldsymbol{\Sigma}$ representa a matriz de covariâncias populacionais das variáveis x e y e $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ representa a matriz de covariâncias das mesmas variáveis implícita ao modelo. A partir das equações (20) e (21), da definição de matriz de covariâncias e do vector das variáveis exógenas \mathbf{x} é possível determinar a matriz de covariâncias implícita.

Antes de mais, é necessário subtrair aos vectores \mathbf{x} e \mathbf{y} as suas médias, *i.e.*, obter a expressão matricial para $(\mathbf{y} - \boldsymbol{\mu}_y)$ e $(\mathbf{x} - \boldsymbol{\mu}_x)$. Note-se que Bollen e Curran (2006: 135n) referem a este respeito que “*we are using deviation scores to simplify the analytical expression of the implied matrix (...) not recommending that researchers use [these scores] in their analysis. Rather, the raw scores are more useful in latent curve analysis*”.

Formalmente, obtém-se:

$$\begin{aligned}
\mathbf{y} - \boldsymbol{\mu}_y &= \left[\boldsymbol{\Lambda} (\boldsymbol{\mu}_\eta + \boldsymbol{\Gamma} \mathbf{x}) + \boldsymbol{\Lambda} \zeta + \boldsymbol{\varepsilon} \right] - \left[\boldsymbol{\Lambda} (\boldsymbol{\mu}_\eta + \boldsymbol{\Gamma} \boldsymbol{\mu}_x) \right] \\
&= \boldsymbol{\Lambda} \boldsymbol{\Gamma} (\mathbf{x} - \boldsymbol{\mu}_x) + \boldsymbol{\Lambda} \zeta + \boldsymbol{\varepsilon} \\
&= \boldsymbol{\Lambda} \left[\boldsymbol{\Gamma} (\mathbf{x} - \boldsymbol{\mu}_x) + \zeta \right] + \boldsymbol{\varepsilon}
\end{aligned} \tag{23}$$

Pela definição de matriz de covariâncias, sabemos que $\boldsymbol{\Sigma} = E \left[(\mathbf{X} - E[\mathbf{X}])(\mathbf{X} - E[\mathbf{X}])^T \right]$.

Sabemos ainda que $E[\mathbf{X}] = \boldsymbol{\mu}_x$. Aplicando a definição,

$$\begin{aligned}
\boldsymbol{\Sigma}(\boldsymbol{\theta}) &= E \left(\begin{bmatrix} \mathbf{y} - \boldsymbol{\mu}_y \\ \mathbf{x} - \boldsymbol{\mu}_x \end{bmatrix} \begin{bmatrix} \mathbf{y} - \boldsymbol{\mu}_y \\ \mathbf{x} - \boldsymbol{\mu}_x \end{bmatrix}^T \right) \\
&= \begin{bmatrix} \boldsymbol{\Lambda} (\boldsymbol{\Gamma} \boldsymbol{\Sigma}_{xx} \boldsymbol{\Gamma}^T + \boldsymbol{\Psi}) \boldsymbol{\Lambda}^T + \boldsymbol{\Sigma}_{xx} & \boldsymbol{\Lambda} \boldsymbol{\Gamma} \boldsymbol{\Sigma}_{xx} \\ \boldsymbol{\Sigma}_{xx} \boldsymbol{\Gamma}^T \boldsymbol{\Lambda}^T & \boldsymbol{\Sigma}_{xx} \end{bmatrix}
\end{aligned} \tag{24}$$

A matriz $\boldsymbol{\Sigma}_{xx}$ representa a matriz de covariâncias populacionais do vector \mathbf{x} , $\boldsymbol{\Lambda}$ representa, como já vimos, a matriz de pesos longitudinais, e $\boldsymbol{\Gamma}$ representa a matriz dos coeficientes associados às variáveis explicativas \mathbf{x} . Decompondo a matriz enunciada em (24) verificamos que estamos perante uma matriz usual em modelos de equações estruturais: o primeiro termo da matriz representa a covariância das medidas repetidas (y) definida em termos dos parâmetros do modelo; o último termo, posicionado na última linha e última coluna, representa, como se referiu, a matriz de covariâncias das variáveis x ; os termos fora da diagonal principal são funções dos parâmetros do modelo que igualam a matriz de covariâncias das medidas repetidas à matriz de covariâncias das variáveis explicativas.

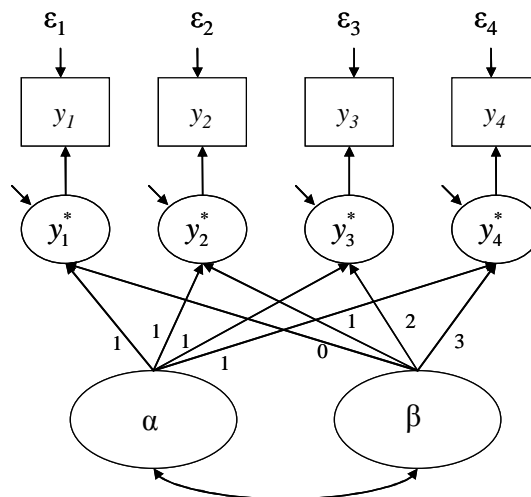
A exposição teórica apresentada considera variáveis manifestas medidas em escalas métricas. No entanto, tal como acontece nos modelos de equações estruturais, os *LGCM* podem ser estimados considerando variáveis ordinais. Uma vez que, utilizando variáveis ordinais, é metodologicamente incorrecto considerar matrizes de variância-covariância e a matriz de correlações de Pearson (*i.e.*, tratar as variáveis como contínuas), é necessário aplicar procedimentos adequados à estimação daquelas matrizes tendo em conta a natureza ordinal das variáveis. Na secção que se segue faremos uma exposição mais detalhada dos procedimentos necessários para a utilização de variáveis ordinais em modelos de trajectória latente.

5.3. MODELOS DE TRAJECTÓRIA LATENTE DE VARIÁVEIS ORDINAIS

Uma leitura atenta da literatura permite perceber que os modelos de trajectória latente considerando variáveis manifestas ordinais estão ainda em fase embrionária, sendo difícil encontrar textos científicos e académicos que utilizem este tipo de modelação. A reduzida dimensão das secções da bibliografia utilizada dedicada a este tema é um indicador do estado da arte nesta matéria.

A figura 21 representa um modelo de trajectória linear latente não condicionada com variáveis ordinais, considerando uma variável medida em quatro momentos consecutivos:

Figura 21 – Modelo de trajectória linear latente não condicionada para uma variável ordinal observada em 4 momentos consecutivos



A modelação de trajectórias latentes de variáveis ordinais é concebida sobre a ideia subjacente aos modelos de equações estruturais que utilizam este tipo de variáveis. A ideia central é a de que, a cada variável ordinal y , está subjacente uma variável contínua latente y^* . A partir destas variáveis, e se nos certificarmos que têm uma distribuição normal bivariada, é possível calcular uma matriz de correlações policóricas (que não é mais do que a matriz de correlações entre as variáveis latentes subjacentes a cada variável ordinal)³⁴.

O debate acerca da utilização de variáveis ordinais em modelos de equações estruturais gira em torno da necessidade de verificação da normalidade bivariada entre as variáveis latentes

³⁴ É também possível calcular a matriz de correlações entre as variáveis latentes subjacentes às variáveis ordinais e variáveis medidas em escalas métricas. Estas correlações, denominadas de *correlações poliserials*, não são relevantes no contexto deste ensaio, uma vez que não consideramos variáveis manifestas contínuas e ordinais simultaneamente no mesmo modelo.

subjacentes às variáveis ordinais. De facto, como refere Salgueiro (2008: 114-115), se este pressuposto não for violado, as estimativas policóricas e poliseriais são as mais adequadas; por outro lado, a violação deste pressuposto leva à obtenção de matrizes de correlações policóricas inflacionadas, e conseqüentemente a erros na estimação.

A exposição metodológica que se apresenta baseia-se na abordagem apresentada por Bollen e Curran (2006:230-240).

5.3.1. Violação de pressupostos

Neste ponto discutiremos as questões levantadas pela violação dos pressupostos de um modelo de trajectória latente pela utilização de medidas repetidas ordinais.

Como se referiu anteriormente, a metodologia aplicada para a modelação de variáveis ordinais no contexto dos modelos de equações estruturais, e também nos modelos de trajectória latente, pressupõe que se considere, para cada variável manifesta ordinal, uma variável latente subjacente contínua, obtida a partir dos *thresholds* das variáveis ordinais³⁵. Assim, ao vector de variáveis ordinais observadas \mathbf{y} , está associado um vector de variáveis contínuas que lhe está subjacente, \mathbf{y}^* . Se até aqui se assumiu que $\mathbf{y} = \mathbf{y}^*$, esta igualdade deixa de se verificar, *i.e.*, o vector de variáveis contínuas subjacentes às medidas repetidas ordinais não é igual ao vector das variáveis de *input*.

O *LGCM* não condicionado é agora formalmente definido como

$$y_{it}^* = \alpha_i + \lambda_t \beta_i + \varepsilon_{it} \quad (25)$$

onde y_{it}^* representa o valor da variável contínua subjacente à ordinal y para o i -ésimo caso no momento t . Com $y_{it} \neq y_{it}^*$, temos ainda que

$$y_{it} \neq \alpha_i + \lambda_t \beta_i + \varepsilon_{it} \quad (26)$$

i.e., a equação de trajectória não é válida para as medidas repetidas ordinais observadas.

³⁵ Discutiremos a estrutura de *thresholds* mais adiante.

Outro pressuposto que se viola remete para as distribuições das variáveis manifestas, que agora dificilmente terão normalidade multivariada. Este problema, que afecta os testes de significância e a eficiência assintótica do estimador de máxima verosimilhança, pode ser ultrapassado pela utilização de métodos de estimação alternativos.

O último pressuposto violado tem consequências bastante mais sérias para a correcta estimação do modelo, e remete para a quebra das equações estruturais. De facto, as equações que estruturam o modelo dificilmente se manterão válidas. Assim, temos que

$$\Sigma^* = \Sigma(\theta) \quad (27)$$

$$\mu^* = \mu(\theta) \quad (28)$$

onde Σ^* e μ^* correspondem, respectivamente, à matriz de covariâncias e ao vector de médias populacionais das variáveis latentes subjacentes às medidas repetidas ordinais, $\Sigma(\theta)$ e $\mu(\theta)$ representam a matriz de covariâncias e o vector de médias implícitas ao modelo, e θ é o vector de parâmetros do modelo. Uma vez que as variáveis observadas são versões truncadas das variáveis contínuas subjacentes, a estrutura de covariâncias não se mantém para as variáveis de *input*. E se $\Sigma \neq \Sigma^*$ e $\mu \neq \mu^*$, então

$$\Sigma \neq \Sigma(\theta) \quad (29)$$

$$\mu \neq \mu(\theta) \quad (30)$$

Como a consistência de um estimador, dada pela sua capacidade de, à medida que aumentamos a amostra, convergir para o verdadeiro valor do parâmetro, é uma das suas mais importantes características, a clara violação deste pressuposto tem consequências graves ao nível da estimação, que não devem ser ignoradas pelo analista.

De forma a superar os problemas de estimação associados à utilização de variáveis ordinais, aplicam-se procedimentos de correcção. Por outro lado, porque o estimador de máxima verosimilhança perde as características de um bom estimador (consistente, assintoticamente não enviesado, assintoticamente normal, e assintoticamente eficiente), é necessário recorrer a métodos de estimação alternativos.

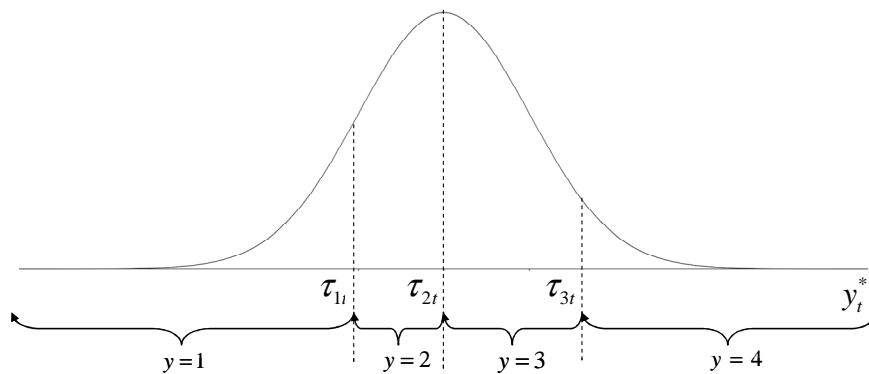
5.3.2. Procedimentos de correcção e estimação por WLS

Como se viu anteriormente, a equação de trajectória definida em (25) não é válida para as variáveis manifestas ordinais. A forma de ultrapassar este problema é impondo uma estrutura auxiliar aos *thresholds*, tal que as variáveis contínuas latentes subjacentes às variáveis ordinais espelhem a distribuição observada nas variáveis de *input*. Assim, impõe-se uma estrutura de *thresholds* às variáveis contínuas, tal que, quando $y_{it} = c_t$

$$\tau_{c_t-1} < y_{it}^* \leq \tau_{c_t} \quad (31)$$

onde $c_t = 1, 2, \dots, C_t$, *i.e.*, as categorias da variável ordinal y_{it} , τ_{c_t-1} e τ_{c_t} são os *thresholds* inferior e superior para a categoria c_t , com $\tau_0 = -\infty$ e $\tau_{C_t} = +\infty$. Teremos $C_t - 1$ *thresholds* ordenados de forma crescente. A figura 22 ilustra a estrutura de *thresholds* imposta a uma variável contínua subjacente a uma variável ordinal com quatro categorias:

Figura 22 – Variável contínua subjacente a uma variável ordinal com quatro categorias



Esta representação permite perceber que todas as observações que caem na primeira categoria da variável ordinal y_{it} terão valores inferiores ao primeiro *threshold* da variável y_{it}^* , as que caem na segunda categorias terão valores em y_{it}^* entre o primeiro e o segundo *threshold*, e assim sucessivamente. A equação para os *thresholds* enunciada em (31) faz a ligação entre as variáveis ordinais observadas e as respectivas variáveis contínuas subjacentes.

No entanto, temos ainda um conjunto de parâmetros desconhecidos, para os quais é necessário encontrar instrumentos que possibilitem a sua identificação. Desconhecemos a média e

variância da variável y_{it}^* , bem como os *thresholds* que a definem. Facilmente se percebe que temos $2 + C_t - 1$ parâmetros desconhecidos para cada medida repetida ordinal que necessitam de ser identificados. Acresce ainda a necessidade de conhecer a distribuição da variável y_{it}^* de forma a estimar os parâmetros do modelo.

Relativamente à distribuição da variável ordinal, conhecemos a proporção de casos que cai em cada uma das suas categorias. Denominemos por $\pi_{1t}, \pi_{2t}, \dots, \pi_{C_t}$ as proporções das respostas dos indivíduos às categorias $1, 2, \dots, C_t$ da variável ordinal manifesta medida no momento t . Obviamente, a soma de todas as proporções é 1, o que significa que temos um número de proporções independentes a considerar equivalente ao número de categorias menos uma.

A única possibilidade de conseguir fazer corresponder a distribuição destas proporções a uma variável contínua subjacente é assumir algum tipo de pressuposto relativamente à distribuição da nova variável. Seguindo a lógica subjacente a um modelo de equações estruturais, assumiremos que a variável y_{it}^* é normalmente distribuída. Impondo este pressuposto, ficamos com $C_t - 1$ valores conhecidos, e ainda dois parâmetros desconhecidos. É necessário impor duas mais restrições. Se, tal como é comum em modelos de SEM, assumirmos ainda que y_{it}^* tem uma distribuição normal estandardizada, *i.e.*, com média zero e desvio-padrão um, reunimos as condições necessárias e suficientes para resolver o sistema de equações subjacente.

Tomemos, por exemplo, a distribuição da variável relativa à perda de sono sentida pelos indivíduos da nossa amostra nos últimos tempos, observada no primeiro momento considerado. As proporções amostrais das respostas em cada uma das categorias cuja ordem, recordamos, corresponde a um acréscimo do sintoma, de *nunca a muito mais do que o habitual*³⁶, é 0,331, 0,488, 0,160 e 0,021. Os *thresholds* para esta variável, que não são mais do que os valores de corte de uma distribuição normal padrão que correspondem à proporção acumulada dos casos que caem em cada categoria, são $\tau_{1_{2003}} = -0,437$, $\tau_{2_{2003}} = 0,913$ e $\tau_{3_{2003}} = 2,033$.

De forma genérica, o valor dos *thresholds* é dado por

³⁶ A representação gráfica da distribuição desta variável é apresentada na figura 3.

$$\tau_{jt} = \Phi^{-1} \left(\sum_{c=1}^j \pi_{ct} \right) \quad (32)$$

com $j = 1, 2, \dots, C - 1$, e Φ^{-1} representando o inverso da distribuição normal padrão acumulada para o valor dentro dos parênteses.

Relativamente à estrutura dos *thresholds* considerando medidas repetidas, encontramos duas sugestões para a sua parametrização. Bollen e Curran (2006: 234) sugerem a utilização da parametrização recomendada por Jöreskog (2001a,b), onde se assume que as médias e variâncias de y_{it}^* são parâmetros livres, e os dois primeiros *thresholds* são fixados a 0 e 1, respectivamente. A sugestão é extensível a todas as medidas repetidas.

Esta parametrização é apresentada como alternativa à parametrização sugerida por vários outros autores, que considera a estandardização das medidas repetidas, fixando as suas médias a zero e as suas variâncias a um. Com esta parametrização, assume-se a igualdade dos *thresholds* em cada medida repetida, *i.e.*, calculam-se os *thresholds* para as variáveis manifestas no primeiro momento, e impõe-se a sua igualdade para as respectivas medidas repetidas.

Embora Bollen e Curran (2006: 234) defendam que “*this alternative assumption [that restricts two thresholds to be equal across time for the repeated measures and estimates the mean and variance for each time period] seems more realistic than the assumption of equal means and variances*”, a parametrização onde se assume a igualdade dos *thresholds* nas medidas repetidas, restringindo as suas médias a zero e variâncias a um, é bastante comum. Nesta situação, as médias e variâncias das variáveis contínuas subjacentes às variáveis ordinais não são estimadas, e estimam-se apenas os *thresholds* no primeiro momento, mantendo-se iguais ao longo do tempo³⁷.

Para a discussão relativa ao método de estimação, recordamos que a hipótese de igualdade da matriz populacional de covariâncias e da matriz de covariâncias implícita ao modelo, habitual em modelos de equações estruturais, não se verifica, como enunciado na equação (29). Nesta situação, como se referiu, o estimador de máxima verosimilhança perde as suas qualidades de bom estimador.

Como referem Flora e Curran (2004: 467), as correlações policóricas são habitualmente calculadas utilizando um procedimento em dois passos, sugerido por Olsson em 1979. O

³⁷ A parametrização de um modelo de trajectória latente utilizando variáveis ordinais será discutida mais detalhadamente na secção seguinte.

primeiro passo consiste no cálculo das variáveis contínuas subjacentes às variáveis observadas ordinais a partir dos *thresholds*, tal como foi explicado anteriormente. No segundo passo, os *thresholds* estimados são utilizados em combinação com a tabela de contingência das variáveis observadas, para estimar a correlação que se obteria caso as variáveis latentes tivessem sido directamente observadas (Flora e Curran, 2004: 468). A função log-verosimilhança para a amostra bivariada é dada por

$$\ell = \ln L + \sum_{c_t=1}^{C_t} \sum_{c_u=1}^{C_u} N_{c_t c_u} \ln(\pi_{c_t c_u}) \quad (33)$$

onde L é uma constante, $N_{c_t c_u}$ representa a frequência das observações na célula (c_t, c_u) , e $\pi_{c_t c_u}$ representa a probabilidade de uma qualquer observação ‘cair’ na célula (c_t, c_u) da tabela de contingência das variáveis ordinais observadas, *i.e.*,

$$\pi_{c_t c_u} = P[y_t = c_t, y_u = c_u] \quad (34)$$

Neste contexto, é fácil perceber porque é que o pressuposto de normalidade bivariada das variáveis latentes subjacentes assume uma relevância central. De facto, só com este pressuposto é possível calcular a probabilidade dada por $\pi_{c_t c_u}$ ³⁸.

Como alternativa ao estimador de máxima verosimilhança, foi desenvolvido o estimador de mínimos quadrados ponderados³⁹ (*WLS*) para a estimação de uma matriz ponderada a partir da matriz de variâncias e covariâncias assintóticas das correlações amostrais. O estimador *WLS* utiliza a função de ajustamento dada por⁴⁰

$$F_{WLS} = [\hat{\boldsymbol{\rho}}^* - \boldsymbol{\rho}^*(\boldsymbol{\theta})]^T \mathbf{W}^{-1} [\hat{\boldsymbol{\rho}}^* - \boldsymbol{\rho}^*(\boldsymbol{\theta})] \quad (35)$$

³⁸ Flora e Curran (2004: 468) apresentam a fórmula de cálculo da probabilidade de uma observação cair numa determinada célula da tabela de contingência das variáveis y_1 e y_2 enunciada por Olsson em 1979:

$\pi_{c_t c_u} = \Phi_2(a_{c_t}, b_{c_u}) - \Phi_2(a_{c_t-1}, b_{c_u}) - \Phi_2(a_{c_t}, b_{c_u-1}) + \Phi_2(a_{c_t-1}, b_{c_u-1})$, onde Φ_2 representa a função normal bivariada de densidade acumulada com correlação ρ .

³⁹ Tradução de *Weighted Least Square* (*WLS*).

⁴⁰ Como referem Bollen e Curran (2006: 239n), Flora e Curran (2004: 469) e Duncan *et al.* (2006: 165-166), esta função é também denominada de *arbitrary distribution function* (*ADF*).

onde $\hat{\boldsymbol{\rho}}^*$ é o vector que contém todas as covariâncias ou correlações policóricas estimadas e as médias das variáveis \mathbf{y}_{it}^* , $\boldsymbol{\rho}^*(\boldsymbol{\theta})$ é o vector dos valores das covariâncias ou correlações policóricas e das médias, implícitas ao modelo, $\boldsymbol{\theta}$ representa um vector de parâmetros do modelo, e \mathbf{W} é uma matriz de ponderação. São seleccionados valores para $\hat{\boldsymbol{\theta}}$ que minimizem a soma ponderada (por \mathbf{W}) dos quadrados dos desvios das correlações policóricas e médias estimadas ($\hat{\boldsymbol{\rho}}^*$), às correspondentes correlações e médias implícitas ao modelo ($\boldsymbol{\rho}^*(\boldsymbol{\theta})$).

O estimador *WLS* encerra as propriedades que o tornam num bom estimador, como a consistência e a eficiência assintótica.

Flora e Curran (2004: 469) apresentam as alterações introduzidas por Muthén ao estimador *WLS*, generalizando a sua aplicação a variáveis categoriais e dicotómicas. Segundo os autores, “*unlike normal-theory estimation, CVM (continuous/categorical variable methodology) provides asymptotically unbiased, consistent, and efficient parameter estimates as well as a correct chi-square test of fit with dichotomous or ordinal observed variables*” (Flora e Curran, 2004: 469).

Os autores evidenciam ainda as limitações do estimador *WLS* quando consideramos variáveis ordinais. Embora se tenha verificado que este estimador é robusto quando há violação de pressupostos de normalidade das variáveis latentes contínuas subjacentes às variáveis ordinais, as ramificações desta violação podem ser graves. Por outro lado, considerando modelos de grande dimensão e amostras pequenas, o estimador *WLS* pode não funcionar, resultando em matrizes \mathbf{W} muitas vezes não positivas definidas, não podendo por isso ser invertidas, impossibilitando a estimação da função de ajustamento.

De forma a lidar com algumas destas limitações, alguns investigadores desenvolveram o estimador *WLS* robusto. Alternativamente ao estimador *WLS*, aqui as estimativas dos parâmetros são calculadas a partir da equação (35), mas agora substituindo a matriz \mathbf{W} pela matriz diagonal \mathbf{V} , cujos elementos são as variâncias assintóticas dos *thresholds* e as correlações policóricas estimadas. Embora os cálculos subjacentes envolvam ainda uma matriz de pesos, deixa de ser necessária e sua inversão, reduzindo assim os problemas encontrados anteriormente.

Em *Mplus* estão implementados dois estimadores *WLS* robustos. O primeiro, *WLSM*, utiliza “*weighted least square parameter estimates using a diagonal weight matrix with standard errors and mean-adjusted chi-square test statistic that use a full weight matrix*”. O estimador *WLSMV* (*Weighted Least Square Means and Variance Adjusted*) utiliza “*weighted least*

square parameter estimates using a diagonal weight matrix with standard errors and mean - and variance adjusted chi-square test statistic that use a full weight matrix” (Muthén e Muthén, 2007: 484). A grande diferença entre o *WLSM* e o *WLSMV* está no ajustamento do teste de χ^2 . Como estes estimadores calculam os graus de liberdade a partir da amostra, e não face ao número de parâmetros a estimar no modelo e ao número de estimativas conhecidas, não é possível fazer testes a partir das diferenças dos χ^2 .

5.3.3. Parametrização de LGCM com variáveis ordinais

Na literatura consultada, os autores sugerem dois tipos de parametrização, incidindo as restrições sobre os factores de escala ou sobre a matriz de variância dos erros associados às variáveis contínuas y_{it}^* .

Os autores do *software* estatístico *Mplus* designam estas duas parametrizações por parametrização Delta e parametrização Theta. Duncan *et al.* (2006: 170) explicam que “*within Mplus, the mean of the slope factor is considered a free parameter while the mean of the intercept factor can be either (a) fixed at zero with free, cross-time-invariant thresholds using the delta parameterization or (b) free with one threshold fixed at all time points using the theta parameterization*”.

Numa modelação de trajectória latente com variáveis contínuas a variância dos erros associados às medidas repetidas pode variar ao longo do tempo. Quando modelamos uma trajectória considerando medidas repetidas ordinais, para que o modelo possa estar identificado⁴¹, é necessário fixar uma de duas medidas no primeiro momento de observação: ou as variâncias dos erros das variáveis y_{it}^* (parametrização Theta), ou os factores de escala⁴² (parametrização Delta). Estes parâmetros podem ser livres para os períodos seguintes, de forma a não distorcer a estrutura do modelo de trajectória.

Os factores de escala tornam possível a articulação entre estrutura de covariâncias e a estrutura de correlações. Na análise tradicional de *LGCM* com medidas repetidas contínuas, a matriz de escala, Λ , não tem qualquer impacto, já que é definida como uma matriz

⁴¹ Recordamos que, num modelo considerando variáveis manifestas ordinais observadas ao longo do tempo, o número de parâmetros desconhecidos aumenta, já que é necessário estimar os *thresholds* das variáveis contínuas associadas às variáveis manifestas.

⁴² Tradução de *scaling factors*.

identidade. Quando todas as variáveis y_{it}^* estão associadas a variáveis y_{it} ordinais⁴³, a matriz Delta (Δ) é utilizada para ajustar a estrutura de covariâncias às covariâncias amostrais. Por outras palavras, para dar escala às variáveis contínuas subjacentes às variáveis ordinais manifestas.

Na parametrização Delta, implementada por defeito no *Mplus*, a diagonal da matriz Δ , que corresponde às estimativas das variâncias das variáveis y_{it}^* , é fixada a um para o primeiro momento, e estimada para os momentos seguintes, de forma a captar as diferenças sofridas nas variáveis y_{it}^* ao longo do tempo. Também a estrutura de *thresholds* tem de ser fixada para o primeiro momento, de forma a possibilitar a estimação.

Na parametrização Theta, as variâncias dos resíduos das variáveis y_{it}^* são fixadas a um para o primeiro período, e estimadas livremente para os períodos seguintes. Os *thresholds* podem ser livres, fixando para todas as variáveis o primeiro e segundo *thresholds*, de forma a ser possível estimar médias e variâncias das variáveis y_{it}^* .

Para uma leitura detalhada da parametrização implementada em *Mplus*, sugere-se a leitura de Muthén, B.O. (2004: 1-11).

A estimação de modelos de trajectória latente de variáveis ordinais em LISREL é mais sinuosa do que em *Mplus*. Uma vez que, em LISREL, nenhuma das parametrizações enunciada está implementada, é necessário parametrizar o modelo em três passos.

No primeiro passo, estimam-se em PRELIS as correlações policóricas, os desvios-padrão e o vector de médias para as variáveis latentes contínuas subjacentes às medidas repetidas ordinais, assumindo igualdade de *thresholds* para cada uma delas ao longo do tempo. No segundo passo, estima-se a matriz de covariâncias assintóticas das correlações amostrais. Por fim, no último passo estima-se o modelo final, com variáveis latentes contínuas subjacentes às medidas repetidas, baseado nas matrizes corrigidas.

Caso se esteja a utilizar a parametrização alternativa, devemos fixar a escala das variáveis y_{it}^* especificando que o primeiro e segundo *thresholds* de todas as variáveis são zero e um, respectivamente. Obtém-se a parametrização alternativa a partir da parametrização descrita anteriormente, aplicando os cálculos sugeridos por Jöreskog (2001: 20)⁴⁴:

⁴³ É possível modelar trajectórias latentes considerando simultaneamente variáveis manifestas ordinais e contínuas, que implica que nem todas as variáveis y_{it}^* correspondem a variáveis manifestas medidas em escalas ordinais (cf. e.g., Muthén, B.O., 2004: 11-12; Duncan *et al.*, 2006: 170-171).

⁴⁴ Citado em Bollen e Curran (2006: 234).

$$\mu_{y_t^*} = \frac{-\tilde{\tau}_{1t}}{\tilde{\tau}_{2t} - \tilde{\tau}_{1t}} \quad (36)$$

$$\sqrt{\sigma_{y_t^* y_t^*}} = \frac{1}{\tilde{\tau}_{2t} - \tilde{\tau}_{1t}} \quad (37)$$

$$\tau_{c_t} = \frac{\tilde{\tau}_{ct} - \tilde{\tau}_{1t}}{\tilde{\tau}_{2t} - \tilde{\tau}_{1t}} \quad (38)$$

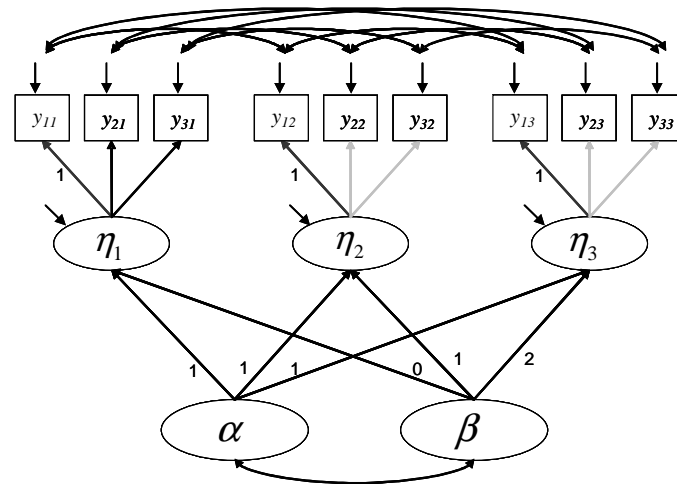
com $c_t = 3, 4, \dots, C_t - 1$, e onde $\tilde{\tau}_{ct} = \tilde{\tau}_{1t}, \tilde{\tau}_{2t}, \dots, \tilde{\tau}_{C_t-1t}$, correspondendo aos *thresholds* estimados a partir do pressuposto distribucional imposto às variáveis contínuas subjacentes às variáveis ordinais manifestas, tal como se explicou anteriormente.

A respeito da estimação dos modelos em LISREL, é ainda de referir que se alterou ligeiramente a proposta apresentada por Duncan *et al.* (2006: 171-172), de forma a facilitar os procedimentos de estimação. Na proposta dos autores, o primeiro passo descrito não incluía a gravação da matriz de correlações policóricas, sendo necessário copiá-la do *output* resultante deste passo para a mancha de comandos necessária para a estimação do modelo de trajectória. Naturalmente, em modelos considerando 12 medidas repetidas em quatro momentos, como são alguns dos modelos estimados neste trabalho, o número de variáveis aumenta para 48, resultando numa matriz de dimensão considerável. A simples cópia destes valores introduz uma margem considerável para erro. Assim, recomenda-se a gravação da matriz de correlações policóricas estimada no primeiro passo num ficheiro externo, e posteriormente a introdução do ficheiro gravado na estimação do modelo final, evitando o trabalho moroso de cópia da matriz, e tornando o procedimento menos permeável a erros.

5.4. MODELOS DE TRAJECTÓRIA LATENTE CONSIDERANDO MÚLTIPLOS INDICADORES

Na figura 23 representa-se um modelo de trajectória linear latente de um factor medido por três variáveis manifestas, considerando três períodos consecutivos. A trajectória a modelar é agora não a de um único indicador observado, mas sim a de um factor latente medido por vários indicadores, que podem ser contínuos ou ordinais.

Figura 23 – Modelo de trajetória linear latente de um factor latente medido por três indicadores, em três momentos consecutivos



Como no contexto dos *LGCM* “*due to the use of the multivariate approach, Mplus does not consider a growth model to be a two-level model as in multilevel modeling but a single-level model*” (Muthén e Muthén: 2007, 87), numa análise longitudinal o número de níveis corresponde ao nível habitualmente atribuído num modelo de equações estruturais menos um. Embora este modelo seja por vezes conhecido como um *LGCM* de segunda ordem, optou-se por adoptar a denominação de modelo com múltiplos indicadores sugerida por Bollen e Curran (2006: 247).

Em modelos desta natureza estima-se a trajetória da variável latente medida por diversas variáveis manifestas. Aqui, modelo de trajetória latente é formalmente dado por:

$$\eta_{it} = \alpha_i + \lambda_t \beta_i + \varepsilon_{it} \quad (39)$$

onde η_{it} representa a variável latente para o indivíduo i no momento t ; ε_{it} representa a perturbação para o caso i no momento t que, como vimos, tem média zero e não está correlacionada com α_i , β_i e λ_t . Os restantes termos são os mesmos já apresentados para os modelos de trajetória latente não condicionada.

As equações para o intercepto e para o declive da trajetória latente são equivalentes às apresentadas em (2) e (3):

$$\alpha_i = \mu_\alpha + \zeta_{\alpha i} \quad (40)$$

$$\beta_i = \mu_\beta + \zeta_{\beta i} \quad (41)$$

Como no modelo não condicional apresentado anteriormente, as perturbações $\zeta_{\alpha i}$ e $\zeta_{\beta i}$ têm média zero e não estão correlacionadas com ε_{it} e λ_t . Até aqui, o modelo apresentado é muito semelhante a um modelo de trajectória latente não condicional.

Quando acrescentamos um modelo de medida ao modelo de trajectória, de forma a considerar os múltiplos indicadores, o modelo altera-se. Assim, é necessário introduzir mais um índice, de forma a introduzir os vários indicadores:

$$y_{jit} = \nu_{jt} + \Lambda_{jt} \eta_{it} + v_{jit} \quad (42)$$

onde $j = 1, 2, \dots, J$, sendo J o número total de indicadores considerados, ν_{jt} corresponde ao intercepto do indicador j no momento t , Λ_{jt} é a matriz de pesos do indicador j no momento t , e v_{jit} é o erro associado à trajectória do i -ésimo caso no momento t para o indicador j . Os pressupostos são equivalentes aos enunciados na secção 5.1., relativamente aos modelos de trajectória latente não condicionada. Note-se que λ_t e Λ_{jt} referem-se a parâmetros diferentes: λ_t é a matriz que incorpora a trajectória, dependendo directamente do tempo (recordamos que, num modelo de trajectória linear, $\lambda_t = t - 1$); Λ_{jt} representa a matriz de pesos da relação entre os indicadores manifestos e a variável latente η_{it} .

O modelo apresentado parte do pressuposto de que temos os mesmos indicadores em cada momento de observação. Seria possível incorporar indicadores diferentes ao longo do tempo, desde que tivéssemos pelo menos um em comum (Bollen e Curran, 2006: 248). No entanto, nos modelos que estimamos tal não acontece.

Antes de analisar as medidas de bondade do ajustamento, refira-se ainda que, porque “*models for panel data must accommodate the fact that observations for the same unit over time are unlikely to be independent of one another*” (Berrington e Smith, 2006: 4), introduziremos nos modelos estimados correlações dos erros de medida das medidas repetidas ao longo do tempo. Isto significa que libertamos as covariâncias entre os erros de medida das medidas repetidas. A identificação de um modelo desta natureza complica-se. Bollen e Curran (2006: 251)

sugerem a utilização do que denominam de *Two-Step Rule*: no primeiro passo, consideram os parâmetros a estimar para o modelo de medida, sem ter ainda em conta o modelo de trajetória latente, assumindo correlações entre os factores ao longo do tempo, e utilizam as regras de identificação para um modelo de análise factorial confirmatória; no segundo passo, consideram os factores latentes como variáveis manifestas, e determinam se um modelo de trajetória latente não condicionada, considerando um número de indicadores equivalente, estaria identificado. Os autores concluem que, com base na análise da identificação de um modelo não condicional, um modelo considerando múltiplos indicadores estará identificado considerando pelo menos três períodos de tempo.

5.5. MEDIDAS DE BONDADE DO AJUSTAMENTO

Para averiguar a bondade do ajustamento dos modelos testados utilizaremos as medidas sugeridas por Muthén, B.O. (2004: 21-24). São propostos testes ao modelo baseados no χ^2 , critérios de informação como o *Akaike Information Criterion (AIC)* ou o *Bayesian Information Criterion (BIC)*, medidas como *Root Mean Square Error of Approximation (RMSEA)* e medidas para os resíduos, como o índice de ajustamento *Standardized Root Mean Square Residual (SRMR)* ou o *Weighted Root Mean Square Residual (WRMR)*.

O teste mais popular ao ajustamento global de um modelo estimado por máxima verosimilhança tem a seguinte estatística:

$$T_{ML} = (N - 1)F_{ML} \quad (43)$$

A estatística T_{ML} é dada pelo valor da função de ajustamento utilizada pelo estimador de máxima verosimilhança, ponderada por $(N - 1)$. A hipótese em teste é a de que

$$H_0 : \boldsymbol{\mu} = \boldsymbol{\mu}(\boldsymbol{\theta}); \boldsymbol{\Sigma} = \boldsymbol{\Sigma}(\boldsymbol{\theta}) \quad (44)$$

Se a hipótese nula for verdadeira e os pressupostos distribucionais impostos para que o estimador de máxima verosimilhança se verificarem (fazendo com que se mantenham as suas

propriedades de bom estimador), a estatística de teste T_{ML} tem uma distribuição assintótica de χ^2 com $\frac{1}{2}T(T+3)-u$ graus de liberdade, onde u representa o número de parâmetros livres estimados pelo modelo. Num modelo perfeitamente identificado, onde o número de parâmetros a estimar é igual ao número de valores conhecidos, a função de ajustamento, e portanto também a estatística de teste, assumem o valor zero, perdendo a sua utilidade. A estatística T_{ML} é apenas aplicável a modelos com variáveis manifestas contínuas. Para modelos considerando variáveis explicativas a estatística de teste é idêntica, mas agora os graus de liberdade são dados por $df = \frac{1}{2}(T+M)(T+M+3)-u$. Recordamos que M corresponde ao número de variáveis explicativas consideradas, T ao número de medidas repetidas, e u ao número de parâmetros a estimar.

Outras medidas de ajustamento utilizadas são denominadas *baseline fit indices*, que comparam o modelo nulo ao modelo estimado, *i.e.*, comparam um modelo mais restritivo, geralmente tomando como parâmetros livres as variâncias das variáveis observadas e fixando a zero as suas covariâncias, com o modelo estimado. A propósito da parametrização utilizada para o modelo nulo, Bollen e Curran (2006: 45) referem que a parametrização da sua estrutura de médias é menos consensual. No entanto, consideram que um modelo com covariâncias entre as variáveis manifestas a zero e libertando as suas variâncias e estrutura de médias para estimação é um modelo nulo aceitável.

A lógica inerente aos *baseline fit indices* é a de considerar o ajustamento do modelo em estudo num *continuum* que vai do pior ao melhor ajustamento possível, *i.e.*, coloca o ajustamento do modelo em análise no intervalo que separa o ajustamento do modelo nulo do ajustamento de um modelo saturado, que tem estatística de teste nula.

Foram desenvolvidas várias medidas desta natureza. Uma das mais utilizadas é o índice de *Tucker-Lewis (TLI)*, que é dado por

$$TLI = \frac{\frac{\chi_N^2}{df_N} - \frac{\chi_H^2}{df_H}}{\frac{\chi_N^2}{df_N} - 1} \quad (45)$$

com χ_N^2 a representar a estatística de teste ao ajustamento do modelo nulo e df_N os respectivos graus de liberdade, e χ_H^2 a estatística de teste ao ajustamento do modelo em análise, com df_H graus de liberdade. Esta medida varia entre 0 – ausência total de ajustamento – e 1 – ajustamento ideal. Bollen e Curran (2006: 46) consideram que, com valores inferiores a 0,9, o investigador deve analisar a adequação do modelo.

O *comparative fit index* (*CFI*) é também uma medida de ajustamento calculada por defeito no *software Mplus*, e é dada por

$$CFI = \frac{1 - \max(\chi_H^2 - df_H, 0)}{\max(\chi_H^2 - df_H, \chi_N^2 - df_N, 0)} \quad (46)$$

Esta medida varia também entre 0 e 1 e a sua interpretação é idêntica à do índice *TLI*.

A medida de ajustamento *RMSEA* é uma medida baseada no erro de aproximação à população, que considera o erro da aproximação à população e a precisão de medida (cf. Salgueiro, 2008: 49). Uma estimativa da função de discrepância da população (*population discrepancy function*) é dada por:

$$\widehat{PDF} = \max\left\{F - \frac{df}{N-1}, 0\right\} \quad (47)$$

onde F representa o mínimo da função de ajustamento. Como \widehat{PDF} tende a decrescer com o aumento do número de parâmetros do modelo, é recomendada a utilização da medida *RMSEA*, calculada a partir da função \widehat{PDF} :

$$RMSEA = \sqrt{\frac{\widehat{PDF}}{df}} \quad (48)$$

Alguns autores sugerem que valores de *RMSEA* inferiores a 0,05 apontam para um bom ajustamento, embora valores até 0,08 sejam ainda aceitáveis.

Os critérios de informação *AIC* e *BIC*, por serem calculados a partir da função de máxima verosimilhança, são disponibilizados apenas para os modelos com variáveis contínuas. Estas medidas permitem a comparação de modelos, e favorecem modelos parcimoniosos. O critério

de informação AIC é dado pela equação seguinte, onde u representa o número de parâmetros estimados pelo modelo.

$$AIC = \chi_H^2 + 2u \quad (49)$$

Quanto menor o valor de AIC , melhor o ajustamento⁴⁵. O critério de informação BIC é interpretado de forma idêntica, mas considera, para além do número de parâmetros do modelo, a dimensão da amostra⁴⁶:

$$BIC = \chi_H^2 + u \ln n \quad (50)$$

O *Mplus* calcula ainda medidas de ajustamento baseadas nos resíduos. O índice de ajustamento $SRMR$ é definido como

$$SRMR = \sqrt{\sum_j \sum_{k \leq j} \frac{r_{jk}^2}{e}}, \quad (51)$$

com $e = p(p+1)/2$, p correspondendo ao número de variáveis contínuas manifestas, e r_{jk} dado por:

$$r_{jk} = \frac{s_{jk}}{\sqrt{s_{jj}} \sqrt{s_{kk}}} - \frac{\hat{\sigma}_{jk}}{\sqrt{\hat{\sigma}_{jj}} \sqrt{\hat{\sigma}_{kk}}}, \quad (52)$$

onde s_{jk} e $\hat{\sigma}_{jk}$ representam as covariâncias amostrais e estimadas pelo modelo entre variáveis contínuas manifestas y_j e y_k , respectivamente. Quando as variáveis manifestas são ordinais, esta medida é calculada apenas no caso de não se considerarem variáveis explicativas no modelo (cf. Muthén, B.O., 2004: 23).

Por fim, o índice $WRMR$ é dado pela seguinte expressão:

⁴⁵ Bollen e Curran (2006: 47n) referem que “these information-based measures are sometimes defined differently in different sources. The different formulas do not alter the ordering of the fit of the models”.

⁴⁶ Como referem Muthén, B.O. (2004: 22), Sclove (1987) sugere um ajustamento do critério BIC considerando $n^* = (n+2)/24$.

$$WRMR = \sqrt{\sum_r^e \frac{(s_r - \hat{\sigma}_r)}{v_r}} / e \quad (53)$$

onde s_r é um elemento do vector de estatísticas amostrais, $\hat{\sigma}_r$ é a correspondente estimativa calculada pelo modelo, v_r é uma estimativa da variância assintótica de s_r , e e é o número de estatísticas amostrais. O índice é adequado para casos em que as estatísticas amostrais têm variâncias muito dispersas, e quando estas estatísticas estão em diferentes escalas, como acontece nos modelos com estruturas de médias e/ou *thresholds*. Para além destas situações, o índice *WRMR* é também indicado para modelos em que as medidas repetidas não têm uma distribuição normal. Concretamente no caso de medidas repetidas em escalas ordinais, e recorrendo o estimador *WLS* robusto, $WRMR = \sqrt{2nF_{\min}} / e$, onde F_{\min} corresponde ao mínimo da função de ajustamento utilizada, n representa a dimensão da amostra e e o número de estatísticas amostrais (Muthén, B.O., 2004: 24).

Muthén, B.O. (2004: 24) apresentam as sugestões de Hu e Bentler (1999) para valores de corte das medidas de bondade do ajustamento a considerar para um bom modelo: $TLI > 0,95$; $CFI > 0,95$; $RMSEA < 0,06$; $SRMR < 0,08$; $WRMR < 0,90$. Os autores afirmam ainda que estes valores são aceitáveis para modelos com variáveis manifestas ordinais.

6. ANÁLISE DOS RESULTADOS

Neste capítulo apresentam-se os resultados dos vários modelos estimados. De forma a facilitar a leitura, optou-se por dividi-lo em duas secções, analisando primeiro os resultados da modelação das trajectórias sem introdução de variáveis explicativas. Assim, na primeira secção serão interpretadas estimativas e comparados os modelos com variáveis manifestas ordinais e contínuas, ainda sem introdução de variáveis explicativas.

Na secção seguinte serão interpretadas e comparadas as soluções estimadas para os mesmos modelos, considerando variáveis explicativas das trajectórias.

Os diagramas representativos dos modelos testados são apresentados em anexo. À medida que se analisar cada trajectória, será feita uma referência ao respectivo diagrama.

6.1. MODELOS DE TRAJECTÓRIA LATENTE NÃO CONDICIONADA DO GHQ-12

Recordamos que o primeiro passo na modelação adoptada do GHQ-12 é a comparação de um modelo de trajectória do indicador único de bem-estar subjectivo (SWB) ao longo dos períodos considerados, com dois modelos de LGCM com os *scores* das três dimensões de bem-estar: um que modela as trajectórias das três novas variáveis de *scores*, correlacionando os interceptos e os declives das trajectórias de cada uma; e outro que considera um factor latente subjacente àquelas três variáveis, que se denominou de debilidade psicológica (PM)⁴⁷. A tabela 6 apresenta os resultados da estimação dos três modelos com medidas repetidas tratadas como contínuas⁴⁸.

Para a leitura da significância dos parâmetros do modelo, recordamos que “*an informal test of significance of a parameter is conducted by dividing the point estimate by its standard error; if the ratio exceeds about 2,00 (1,96 in very large samples), the parameter estimate is said to be significantly different from zero at the 0,05 level*” (Preacher *et al.*, 2008: 18).

Como vimos no capítulo relativo à parametrização de modelos com medidas repetidas contínuas, para as trajectórias considerando os *scores* do GHQ-12, quer num indicador único, quer em três indicadores, permite-se que os indivíduos tenham valores iniciais de trajectória diferentes (*i.e.*, estimamos livremente a média do intercepto). Estes valores serão,

⁴⁷ Os diagramas que representam os três modelos referidos são apresentados nos pontos 10.4.1, 10.4.5 e 10.4.6.

⁴⁸ A exposição metodológica relativa a estes modelos foi apresentada na secção 5.1.

naturalmente, muito próximos das médias das variáveis, apresentadas na secção relativa à caracterização da amostra.

Tabela 6 – Estimativas para os modelos 1, 2 e 3, com trajectórias não condicionadas

	Modelo 1		Modelo 2						Modelo 3	
	SWB		AD		SD		LC		PM	
	intSWB	decSWB	intAD	decAD	intSD	decSD	intLC	decLC	intPM	decPM
MD	11,00	0,107	3,652	0,035	6,266	0,044	1,083	0,027	0,000	0,045
intSWB	14,779	-0,605	---	---	---	---	---	---	---	---
decSWB	---	0,611	---	---	---	---	---	---	---	---
intAD	---	---	3,437	-0,135	3,090	-0,649	1,883	-0,211	---	---
decAD	---	---	---	0,136	-0,673	0,454	-0,265	0,177	---	---
intSD	---	---	---	---	2,231	-0,119	1,487	-0,243	---	---
decSD	---	---	---	---	---	0,109	-0,299	0,203	---	---
intLC	---	---	---	---	---	---	0,852	0,006	---	---
decLC	---	---	---	---	---	---	---	0,018	---	---
intPM	---	---	---	---	---	---	---	---	2,306	-0,073
decPM	---	---	---	---	---	---	---	---	---	0,086

NOTA: MD – médias; os restantes valores são variâncias (na diagonal) e covariâncias entre interceptos e declives

Relativamente à trajectória do indicador único de bem-estar subjectivo (*SWB*, modelo 1)⁴⁹, a média do intercepto mostra que os indivíduos consideram ter, em média, níveis de bem-estar psicológico subjectivo positivos⁵⁰. A média do declive é positiva e significativa, o que quer dizer que os níveis de debilidade psicológica percebidos pelos indivíduos aumentam, em média, entre os períodos considerados, cerca de 0,1.

Por outro lado, as variâncias do intercepto e do declive são significativas, sugerindo que se verificam diferenças nas trajectórias individuais do *SWB* relativamente à trajectória média. Por outras palavras, há evidência de variabilidade significativa das trajectórias individuais em torno do intercepto e declive médios. Note-se ainda que a covariância entre declive e intercepto é significativa e negativa, sugerindo que indivíduos com médias mais elevadas no *score* de bem-estar subjectivo (que significa serem indivíduos mais debilitados psicologicamente), aumentam o nível de debilidade psicológica média mais lentamente do que indivíduos com melhores níveis de bem-estar psicológico médio inicial.

⁴⁹ O diagrama representativo da modelação do indicador *SWB* é semelhante ao diagrama apresentado na figura 19 da secção 5.1., onde a medida repetida é agora o indicador de bem-estar subjectivo. Diagrama específico apresentado em anexo, ponto 10.4.1.

⁵⁰ Recordamos que o índice *SWB* é constituído pela soma das respostas dos indivíduos aos itens do GHQ-12, variando entre 0 (menos perturbado(a)) e 36 (mais perturbado(a)).

Antes de se proceder à estimação do modelo 2, que modela conjuntamente as três trajectórias de ansiedade e depressão, disfunção social e perda de confiança, estudaram-se previamente, e em modelos separados, as três trajectórias de cada uma destas dimensões. Os resultados da estimação dos três modelos de trajectória latente para as variáveis *ansiedade e depressão* (*AD*), *disfunção social* (*SD*) e *perda de confiança* (*LC*), quando considerados separadamente, estão sintetizados na tabela 7⁵¹.

Tabela 7 – Estimativas para os modelos de trajectória dos *scores* ansiedade e depressão (*AD*), *disfunção social* (*SD*) e *perda de confiança* (*LC*)

	<i>AD</i>		<i>SD</i>		<i>LC</i>	
	intAD	decAD	intSD	decSD	intLC	decLC
MD	3,652	0,035	6,265	0,045	1,083	0,027
intAD	3,477	-0,153	---	---	---	---
decAD	---	0,142	---	---	---	---
intSD	---	---	2,300	-0,154	---	---
decSD	---	---	---	0,127	---	---
intLC	---	---	---	---	0,877	-0,003
decLC	---	---	---	---	---	0,020

Tabela 8 – Medidas de bondade do ajustamento dos modelos de trajectória dos *scores* ansiedade e depressão (*AD*), *disfunção social* (*SD*) e *perda de confiança* (*LC*)

	<i>AD</i>	<i>SD</i>	<i>LC</i>
CFI	0,091	0,987	0,994
TLI	0,989	0,984	0,993
RMSEA	0,048	0,041	0,039
SRMR	0,017	0,020	0,012
AIC	79482,550	80618,099	55228,065
BIC	79540,379	80675,928	55285,894

As três trajectórias são semelhantes, apresentando médias dos interceptos e declives significativas e positivas, sugerindo que, em média, em 2003, os indivíduos têm níveis de ansiedade e depressão, disfunção social e perda de confiança relativamente baixos⁵², aumentando entre períodos consecutivos a uma taxa média de 0,035, 0,045 e 0,027, respectivamente. Por outro lado, a significância das variâncias de interceptos e declives aponta para diferenças significativas entre as trajectórias individuais e a trajectória média em cada uma das variáveis, quer no momento inicial, quer ao nível dos acréscimos médios sofridos entre 2003 e 2006. As covariâncias entre intercepto e declive são significativas apenas para os modelos relativos às trajectórias de ansiedade e depressão e de disfunção social em 2003, sugerindo trajectórias destes factores semelhantes à trajectória identificada para o indicador único de bem-estar subjectivo (*SWB*): indivíduos que, em média, manifestam

⁵¹ Os diagramas dos modelos em análise neste ponto são apresentados nos pontos 10.4.2 a 10.4.4 dos anexos.

⁵² Recordamos que as dimensões *AD*, *SD* e *LC* resultam da soma das respostas dos indivíduos aos itens que correspondem a cada um dos factores de bem-estar, codificados em escalas de 0 a 3. Assim, o factor *AD* varia entre 0 e 12 (esta dimensão corresponde à soma de 4 itens do GHQ-12), o factor *SD* varia entre 0 e 18 (corresponde à soma de 6 itens do GHQ-12), e o factor *LC* varia entre 0 e 6 (resultante da soma de 2 itens do GHQ-12). Para todos eles, o valor mais baixo corresponde a menores graus de ansiedade e depressão, disfunção social e perda de confiança, respectivamente.

maiores graus de ansiedade e depressão e de disfunção social, têm nestas dimensões taxas de mudança mais baixas, *i.e.*, aumentam mais lentamente o nível médio nestas dimensões. Relativamente ao indicador de perda de confiança, não há evidência de que diferentes médias iniciais tenham impacto ao nível das trajectórias (a covariância entre intercepto e declive não é significativa).

Na tabela 8 apresentam-se medidas de bondade do ajustamento modelo-dados para as trajectórias das três dimensões. Todas as medidas consideradas apontam para um bom ajustamento nos três modelos.

Retomando a análise dos resultados apresentados na tabela 6, agora considerando o modelo 2⁵³, que modela simultaneamente as trajectórias dos três indicadores apresentados relacionando os seus interceptos e declives, verificamos que os resultados obtidos para as três trajectórias não contrastam com que se disse anteriormente acerca da modelação em separado de cada uma delas. Importa agora perceber como se relacionam os três processos entre si.

Considerando a articulação entre os três processos, as covariâncias entre os diferentes interceptos permitem perceber que as médias iniciais em cada uma das dimensões estão relacionadas com as médias iniciais estimadas para as restantes. Estas covariâncias são todas positivas e significativas, sugerindo que, de facto, níveis médios iniciais de ansiedade e depressão altos (ou baixos) estão associados a níveis médios iniciais de disfunção social e de perda de confiança altos (ou baixos). Analogamente, níveis médios iniciais de disfunção social altos (ou baixos) estão associados aos níveis médios iniciais de perda de confiança altos (ou baixos), ou vice-versa.

A significância observada nas covariâncias entre os declives das três trajectórias aponta também para associações múltiplas entre uns e outros processos. De forma semelhante ao que se verificou relativamente às covariâncias entre as médias iniciais de umas e outras dimensões, também os ritmos de mudanças em cada uma das dimensões se reflectem mutuamente de forma directa: o aumento (ou decréscimo) nos níveis de ansiedade e depressão médios ao longo dos quatro anos é acompanhado de um aumento (ou decréscimo) nos níveis de disfunção social e de perda de confiança médios (ou vice-versa); também o aumento (ou decréscimo) no nível médio de disfunção social é acompanhado de um aumento (ou decréscimo) no nível médio de perda de confiança (ou vice-versa).

Percebemos ainda que diferenças iniciais nas médias das três variáveis estão associadas às trajectórias das restantes. Mais concretamente, indivíduos com médias iniciais de ansiedade e

⁵³ Diagrama do modelo em anexo, ponto 10.4.5.

depressão mais elevadas em 2003 tenderão a aumentar a média nas dimensões de disfunção social e de perda de confiança ao longo destes anos de forma mais lenta. Analogamente, médias iniciais mais elevadas de disfunção social estão associadas a ritmos de mudança na trajectória das percepções de ansiedade e depressão e perda de confiança mais lentos.

De forma sucinta, embora as trajectórias globais dos três factores tenham níveis médios iniciais diferentes (que resulta directamente dos diferentes intervalos de variação possíveis em cada um deles), em todas elas encontramos diferenças das trajectórias individuais relativamente à trajectória média. Por outro lado, a articulação entre os seus níveis iniciais e as suas taxas de mudança é análoga relativamente aos três processos. Com esta ideia em mente, e recordando que as dimensões de ansiedade e depressão, disfunção social e perda de confiança são concebidas como constructos de um mesmo factor de debilidade psicológica, faz também sentido que as trajectórias dos três índices estejam fortemente relacionadas entre elas. A evolução estimada para cada uma destas dimensões está relacionada com a evolução das restantes.

A partir desta ideia, e tendo em conta o modelo conceptual desenhado, estimou-se outro modelo a partir dos *scores* nas três dimensões, considerando agora um constructo medido pelas três variáveis que, como se discutiu anteriormente, se denominou de debilidade psicológica (*PM*) (modelo 3⁵⁴). As estimativas obtidas para o modelo apresentam-se na tabela 6.

O comportamento da trajectória do factor de debilidade psicológica é semelhante ao comportamento da trajectória do indicador único de bem-estar subjectivo, interpretado anteriormente. Por um lado, verifica-se existirem desvios significativos das médias individuais de debilidade psicológica no momento inicial relativamente à média global no mesmo momento. De forma semelhante, conclui-se existir variabilidade significativa das trajectórias individuais em torno da trajectória média, representada pela variância do declive estimado. Por fim, a covariância entre intercepto e declive da trajectória de *PM*, negativa e significativa, sugere que valores médios de debilidade psicológica mais elevados em 2003 estão associados a trajectórias médias menos acentuadas ou de crescimento menos pronunciado.

As medidas de bondade do ajustamento obtidas para os modelos 1, 2 e 3 analisados, apresentadas na tabela 9, mostram que o modelo 1 e o modelo 3 têm um melhor ajustamento. Comparativamente, o modelo 2, com uma estrutura de variâncias e covariâncias mais

⁵⁴ Diagrama do modelo em anexo, ponto 10.4.6.

complexa, aumentando consideravelmente o número de parâmetros a estimar, tem medidas de ajustamento que ultrapassam os (ou ficam aquém dos) valores de corte sugeridos. Por outro lado, os critérios de informação, que permitem a comparação entre os modelos, sugerem que a modelação da trajectória do indicador único apresenta um melhor ajustamento modelo-dados.

Tabela 9 – Medidas de bondade do ajustamento dos modelos de trajectória 1, 2 e 3

	Modelo 1	Modelo 2	Modelo 3
CFI	0,987	0,765	0,974
TLI	0,985	0,696	0,963
RMSEA	0,054	0,183	0,064
SRMR	0,021	0,072	0,063
AIC	107993,145	204858,800	197946,309
BIC	108050,975	205109,395	198222,606

Esta conclusão, algo desanimadora do ponto de vista da conceptualização inicialmente estabelecida, não invalida a análise dos modelos considerando os itens que constituem a escala do *GHQ-12* tal como foram medidos. Recordamos a estrutura do modelo conceptual, que considera três constructos distintos, relativos à ansiedade e depressão, à disfunção social e à perda de confiança, medidos por diferentes conjuntos de variáveis, tal como se apresenta no anexo 10.2.

Na tabela 10 apresentam-se os resultados da estimação de dois modelos, considerando a escala ordinal das medidas repetidas.

Tabela 10 – Estimativas dos modelos 4 e 5, com trajectórias não condicionadas

	Modelo 4						Modelo 5	
	<i>AD</i>		<i>SD</i>		<i>LC</i>		<i>PM</i>	
	intAD	decAD	intSD	decSD	intLC	decLC	intPM	decPM
MD	---	0,015	---	0,015	---	0,033	---	0,015
intPM	---	---	---	---	---	---	0,326	-0,018
decPM	---	---	---	---	---	---	---	0,013
intAD	0,330	-0,013	0,376	-0,090	0,537	-0,077	---	---
decAD	---	0,009	-0,091	0,059	-0,082	0,05	---	---
intSD	---	---	0,251	-0,016	0,436	-0,085	---	---
decSD	---	---	---	0,012	-0,096	0,062	---	---
intLC	---	---	---	---	0,546	-0,006	---	---
decLC	---	---	---	---	---	0,003	---	---

Até agora foram modeladas trajectórias a partir de *scores* obtidos como soma de itens observados. Nos modelos que analisaremos de seguida consideram-se factores latentes medidos por múltiplos indicadores ordinais. Recordamos que, na parametrização imposta a modelos que consideram variáveis observadas ordinais, se restringe a matriz dos factores de escala no primeiro momento considerado, e as médias dos interceptos não são estimadas, mas forçadas a zero⁵⁵.

Para a análise dos modelos considerando medidas repetidas ordinais, procurou-se verificar o pressuposto de normalidade bivariada das variáveis contínuas subjacentes às variáveis ordinais⁵⁶. Segundo refere Salgueiro (2008: 118), Jöreskog, tendo verificado que o teste à hipótese nula de normalidade bivariada, cuja estatística de teste tem distribuição de χ^2 , era sensível a desvios de normalidade, desenvolveu uma medida tipo *RMSEA* que, caso assumia valores inferiores a 0,1, permite que se assumia normalidade bivariada. A verificação deste pressuposto foi feita em LISREL, e permitiu concluir, através da significância ao teste associado à medida *RMSEA* referida⁵⁷, que 17 pares de variáveis poderão não satisfazer este pressuposto. A este respeito, Flora e Curran (2004: 482) mostram que, utilizando o estimador *WLS* robusto, uma “*modest violation of normality for latent response variables of a degree that might be expected in applied research leads to only slightly biased estimates of polychoric correlations*”. Sublinhe-se ainda que a verificação deste pressuposto foi feita em LISREL, sendo os resultados apresentados estimados em *Mplus*. Seria conveniente testar o pressuposto no *software* utilizado para a estimação, sob pena de ligeiras diferenças nos procedimentos levarem a conclusões não completamente idênticas. No entanto, não deixa de parecer importante referir este aspecto.

Analisando agora as soluções estimadas, refira-se como primeira nota que o modelo 4⁵⁸, equivalente na estrutura das relações entre trajectórias ao modelo 2, sugere um conjunto de conclusões muito semelhantes àquele. Antes de mais, a significância das médias dos declives de cada um dos processos sugere que, em média, os graus de ansiedade e depressão média, de disfunção social média e de perda de confiança média crescem ligeiramente entre os períodos considerados (0,015, 0,015 e 0,033, respectivamente).

⁵⁵ A parametrização dos modelos em análise, que corresponde à parametrização Delta, é discutida na secção 5.3.3.

⁵⁶ A discussão dos pressupostos subjacentes aos modelos de LGCM com variáveis ordinais foi apresentada nos pontos 5.3.1. e 5.3.2.

⁵⁷ O teste tem como hipótese nula que a medida *RMSEA* é inferior a 0,1 na população (*cf.* Salgueiro, 2008: 118).

⁵⁸ Diagrama do modelo em anexo, ponto 10.4.6.

A significância das variâncias dos interceptos e declives das trajectórias dos factores que medem a ansiedade e depressão e a disfunção social permitem concluir que existe variabilidade significativa das trajectórias individuais em torno da trajectória média, quer ao nível do seu valor médio inicial, quer ao nível da tendência (*i.e.*, da taxa de mudança, ou do declive da recta de trajectória média). Relativamente ao factor referente à perda de confiança, apenas a variância do intercepto é significativa, apontando para diferenças individuais em torno da média inicial, mas não na sua tendência. Isto significa que, em termos médios, a trajectória global do factor referente à perda de confiança reflecte a tendência das trajectórias individuais, mas não os seus pontos médios iniciais⁵⁹.

Neste ponto, a comparação deste modelo com o modelo 2 estimado anteriormente revela uma diferença interessante. Anteriormente concluímos existir variabilidade individual (ainda que reduzida) não só em torno do nível médio de perda de confiança em 2003, como também relativamente à tendência média. Não podemos deixar de colocar a hipótese de que a perda de significância da variância do declive se relaciona com a natureza das variáveis agora introduzidas.

Quanto à covariância entre os interceptos e declives respeitantes a cada um dos processos, verificamos que, para as trajectórias dos constructos relativos à ansiedade e depressão e à disfunção social, os seus valores médios em 2003 associam-se significativamente às respectivas tendências estimadas ao longo dos 4 anos em estudo. Mais concretamente, médias mais altas (ou mais baixas) em 2003 estão associadas a tendências menos acentuadas (ou mais acentuadas), *i.e.*, taxas de crescimento dos níveis de ansiedade e depressão e de disfunção social mais baixas (ou mais altas). O mesmo não acontece relativamente à trajectória do factor de perda de confiança. Verifica-se que, neste caso, os valores médios de perda de confiança iniciais não têm qualquer relação com a tendência da sua evolução nos períodos considerados. A articulação das trajectórias dos três constructos conduz à conclusão que, globalmente, elas estão associadas. Verificamos também aqui que os níveis médios em 2003 em cada uma das dimensões têm relação com os níveis médios iniciais das restantes, *i.e.*, aumentando (ou diminuindo) o nível médio de uma das dimensões em 2003, aumenta (ou diminui) também o nível médio nas restantes dimensões no mesmo momento. Analogamente, a significância das covariâncias entre os vários declives, sempre positivas, sugere que a taxa média de crescimento dos níveis de ansiedade e depressão, disfunção social e perda de confiança entre

⁵⁹ Isto significa que, se traçássemos uma recta para a trajectória média estimada e rectas para as trajectórias individuais, estas últimas teriam pontos iniciais diferentes da trajectória média, mas rectas que lhe seriam paralelas.

2003 e 2006 se reflectem mutuamente de forma directa, *i.e.*, aumentando (ou diminuindo) a taxa média de crescimento em cada um dos sintomas entre os períodos considerados, aumenta (ou diminui) a taxa média de crescimento nos restantes.

Por fim, as covariâncias entre interceptos e declives são todas negativas e significativas, o que significa que as médias iniciais de ansiedade e depressão, disfunção social e perda de confiança têm uma associação inversa ao nível das tendências de evolução entre si. Por outras palavras, níveis médios iniciais mais altos (ou mais baixos) de cada um dos sintomas em 2003 reflectem-se em ritmos de aumento dos restantes sintomas ao longo dos anos observados mais lentos (ou mais acelerados).

O modelo 5⁶⁰ estima a trajectória do modelo conceptual desenhado a partir do debate em torno da escala do *GHQ-12* e da sua multidimensionalidade, ou seja, modela a trajectória do factor de segunda ordem de debilidade psicológica, subjacente aos constructos relativos à ansiedade e depressão, disfunção social e perda de confiança. As estimativas obtidas para este modelo não contrastam com o que se tem dito até aqui. Verificamos que, entre cada período em análise, o nível médio de debilidade psicológica aumenta ligeiramente (0,015 em média). Por outro lado, verificamos que existe variabilidade das trajectórias individuais em torno da trajectória média, quer nos níveis médios iniciais, quer na taxa de mudança. Por fim, o nível médio de debilidade psicológica em 2003 reflecte-se de forma inversa na sua trajectória média, o que significa que a um nível médio inicial de debilidade psicológica mais alto (ou mais baixo) está associada uma taxa de crescimento mais baixa (ou mais alta). Por outras palavras, maiores graus de debilidade psicológica inicial estão associados a um crescimento menos acentuado ao longo dos 4 anos. Inversamente, se os níveis médios de debilidade psicológica forem mais baixos, a tendência de crescimento é mais acentuada.

A tabela 11 apresenta as medidas de bondade do ajustamento dos modelos 4 e 5. Embora não sejam fornecidas medidas para a comparação de modelos⁶¹, analisando separadamente cada um deles, vemos que os seus ajustamentos estão longe de ser perfeitos. Relativamente ao modelo 5, a medida *RMSEA* aponta para um fraco ajustamento. Quanto aos valores da medida *WRMR*, relativa à bondade de ajustamento aferida a partir da distribuição dos resíduos

⁶⁰ Diagrama do modelo em anexo, ponto 10.4.7.

⁶¹ Estimando os mesmos modelos em LISREL, é possível obter valores para o critério AIC, que permitem a comparação dos modelos. Os valores deste critério para os modelos 4 e 5 são $AIC_4 = 48988,411$ e $AIC_5 = 39322,064$. A estimação em LISREL aponta o modelo 5 como o melhor dos dois. Embora, em rigor, não seja possível estender esta conclusão para os modelos estimados em *Mplus*, não deixamos de colocar a hipótese de que o mesmo acontece nos modelos apresentados. É possível consultar os comandos utilizados para a estimação em LISREL nos anexos 10.7. e 10.8.

estimados⁶², elas assumem valores bastante elevados. No entanto, como refere Linda Muthén num dos fóruns de debate do *Mplus*, pode acontecer que as diferentes medidas de bondade do ajustamento apresentem valores contraditórios⁶³, o que não significa que o modelo tenha um mau ajustamento.

Tabela 11 – Medidas de bondade do ajustamento dos modelos de trajectória 4 e 5

	Modelo 4	Modelo 5
CFI	0,854	0,804
TLI	0,973	0,952
RMSEA	0,075	0,100
WRMR	3,772	5,067

Com o objectivo de tentar explicar a variabilidade individual em torno das trajectórias médias estimadas, introduziram-se algumas variáveis explicativas. Analisaremos o impacto destas variáveis nas trajectórias estimadas do indicador único de bem-estar subjectivo (modelo 1), da dimensão de debilidade psicológica medida pelos *scores* de ansiedade e depressão, disfunção social e perda de confiança (modelo 3), e do factor de segunda ordem das três dimensões referidas, medidas pelas variáveis ordinais que constituem a escala do *GHQ-12* (modelo 5). Porque os modelos 1, 3 e 5 propõem diferentes estruturas para medir uma única trajectória linear do fenómeno em análise (um indicador único no modelo 1, três indicadores construídos a partir dos *scores* no modelo 3, e um factor de segunda ordem construído a partir dos factores de ansiedade e depressão, disfunção social e perda de confiança, medidos pelas respectivas variáveis ordinais observadas), optou-se por introduzir variáveis explicativas apenas neste modelos, procurando assim perceber se alguma das três propostas de modelação acrescenta capacidade explicativa. Futuramente, poderá ser interessante introduzir variáveis explicativas nos modelos que consideram três trajectórias distintas correlacionadas entre si (modelos 2 e 4), sob a hipótese de algumas variáveis explicativas terem impacto numas ou noutras trajectórias.

⁶² As medidas de ajustamento apresentadas foram explicadas na secção 5.5.

⁶³ Muthén, L.K. (2004) [<http://www.statmodel.com/discussion/messages/23/416.html>], consultada em 28 de Junho de 2009].

6.2. MODELOS DE TRAJECTÓRIA LATENTE CONDICIONADA DO GHQ-12

Neste ponto apresentaremos os resultados da estimação dos modelos 1, 3 e 5 considerando três variáveis explicativas invariantes: o sexo, a idade e o estado de saúde percebido⁶⁴. Os resultados do impacto destas variáveis nos três modelos são apresentados na tabela 12.

Tabela 12 – Impacto das variáveis explicativas nos modelos 1, 3 e 5 (modelos 1c, 3c e 5c) de trajectória latente condicionada

	Modelo 1c		Modelo 3c		Modelo 5c	
	SWB		PM		PM ordinais	
	Intcpt	Slope	Intcpt	Slope	Intcpt	Slope
Sexo:						
mulher	1,314		0,601		0,213	
Escalão etário:						
30-44 anos	0,647		0,215		0,112	
45-64 anos						
65 + anos			-0,301		-0,088	
Estado de saúde:						
boa	1,383	-0,149	0,586		0,243	-0,026
fraca/muito fraca	4,237	-0,329	1,780	-0,126	0,664	-0,055

As variáveis explicativas, todas categoriais, foram transformadas em variáveis *dummy*, de forma a possibilitar a sua introdução nos modelos de trajectória latente⁶⁵. Para a variável sexo, a categoria de referência é o sexo masculino. Na variável de escalão etário, o escalão de referência corresponde a idades entre os 16 e os 29 anos. Quanto ao estado de saúde percebido, a categoria correspondente a estados de saúde excelentes é a referência.

Relativamente à trajectória do indicador único de bem-estar subjectivo (modelo 1c), verificamos que as mulheres têm *scores* médios de *SWB* em 2003 significativamente mais altos do que os homens (média do intercepto positiva), sugerindo portanto que as mulheres em 2003 se sentem, em média, mais debilitadas psicologicamente do que os homens. Por outro lado, indivíduos com idades entre os 30 e os 44 anos apresentam, em 2003, graus médios de debilidade psicológica significativamente mais elevados do que os indivíduos mais novos. Por fim, indivíduos que consideram ter um estado de saúde mais fraco têm valores

⁶⁴ Como se explorou no capítulo relativo à caracterização da amostra em estudo, a variável relativa ao estado de saúde percebido praticamente não varia entre 2003 e 2006, considerando-se, para efeitos de estimação, como constante ao longo dos quatro anos. Ainda, a variável idade, por ter uma variação constante no tempo, é também introduzida como variável invariante.

⁶⁵ A apresentação dos modelos de trajectória condicionada foi feita na secção 5.2.

médios de debilidade psicológica mais elevados em 2003 do que aqueles que consideram estar num estado excelente de saúde, acentuando-se esta assimetria nos indivíduos que consideram ter um estado de saúde fraco a muito fraco.

Relativamente ao impacto destas variáveis na tendência da trajectória, e ainda para o modelo 1c, verificamos que os indivíduos que afirmam ter piores estados de saúde em 2003 têm decréscimos significativos no ritmo médio de evolução da sua debilidade psicológica relativamente ao ritmo de evolução do grau de debilidade psicológica sentida pelos indivíduos com estados de saúde excelentes em 2003 (médias dos interceptos negativas).

O modelo 3c, que considera agora uma estrutura tripartida do indicador de bem-estar subjectivo, medido pelos *scores* das três dimensões que medem a debilidade psicológica, se por um lado parece atenuar o impacto das variáveis explicativas na trajectória estimada, evidencia o impacto de categorias que antes não eram significativas.

Começando pela análise do efeito das variáveis explicativas introduzidas sobre o intercepto aleatório, verificamos que, também aqui, as mulheres apresentam valores médios iniciais de debilidade psicológica significativamente superiores aos dos homens (média do intercepto positiva).

Quanto ao estado de saúde percebido, a conclusão a retirar é semelhante: indivíduos com estados de saúde percebidos piores em 2003 tendem a manifestar níveis médios mais altos de debilidade psicológica no primeiro ano considerado, quando comparados com indivíduos que afirmam estar em excelente estado de saúde.

Relativamente ao impacto da idade na trajectória estimada, verificamos que, ao passo que os indivíduos com idades entre os 30 e os 44 anos têm graus médios de debilidade psicológica em 2003 significativamente mais elevados do que os indivíduos mais novos, acontece o inverso quando consideramos os indivíduos mais velhos, com 65 ou mais anos. De facto, indivíduos com 65 ou mais anos têm valores médios de debilidade psicológica significativamente mais baixos em 2003 quando comparados com os mais novos. Embora este resultado pareça contradizer as nossas noções mais intuitivas relativamente à evolução do bem-estar psicológico, também Clark (2002) chega às mesmas conclusões. Na sua análise longitudinal do impacto da idade no bem-estar subjectivo dos indivíduos, a partir da noção de que a relação entre a idade e o bem-estar subjectivo em dados seccionais é *U-shaped* (*i.e.*, não é linear), o autor utiliza os dados do painel BHPS com o objectivo de distinguir a componente do envelhecimento nesta relação, do efeito do *coorte*, *i.e.*, o efeito de, por exemplo, indivíduos nascidos entre 1940 e 1950, independentemente da idade que têm no momento de análise.

Clark (2002: 7) conclui não só que “*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*”, como ainda que “*those born earlier report higher levels of well-being on the GHQ scale, independently of their current age*”.

Analisando agora o impacto das variáveis explicativas sobre o declive da trajectória, percebemos que apenas a percepção do estado de saúde tem um efeito significativo. Concretizando, é possível concluir que indivíduos com estados de saúde fracos ou muito fracos têm um decréscimo significativo na taxa de mudança da sua debilidade psicológica, quando comparados com indivíduos em excelente estado de saúde (média do declive negativa).

O último modelo em análise introduz variáveis explicativas na trajectória do factor de segunda ordem relativo à debilidade psicológica (modelo 5c). Um aspecto a sublinhar é que, neste modelo, todas as categorias com impacto ao nível da trajectória dos modelos anteriores são significativas (quer aquelas que eram significativas nos dois, quer as que eram num ou noutro).

Relativamente às conclusões a retirar, elas são, do ponto de vista substantivo, em tudo idênticas. As mulheres têm níveis iniciais médios de debilidade psicológica mais altos do que os homens. Tal como vimos acontecer nos outros modelos, indivíduos com idades entre os 30 e os 44 anos têm graus médios de debilidade psicológica no momento inicial significativamente mais elevados do que os indivíduos com idades entre os 16 e os 29 anos. Também neste modelo indivíduos mais velhos, com idades superiores aos 65 anos, têm valores médios de debilidade psicológica significativamente mais baixos em 2003 quando comparados com os mais novos. Por fim, o impacto da introdução do estado de saúde percebido é idêntico ao avaliado para os modelos anteriores: indivíduos com estados de saúde mais fracos têm valores médios de debilidade psicológica em 2003 significativamente superiores aos níveis médios percebidos em indivíduos que consideram ter um estado de saúde excelente.

Quanto ao impacto das variáveis explicativas no declive da trajectória média de debilidade psicológica, apenas as categorias relativas a estados de saúde mais débeis têm impacto, quando comparadas com a categoria relativa a estados de saúde excelentes: conclui-se que os primeiros têm um decréscimo significativo na taxa de mudança da sua debilidade psicológica, quando comparados com indivíduos em excelente estado de saúde.

A tabela 13 apresenta as medidas de bondade do ajustamento para os modelos de trajectória latente condicionada. O modelo 5c revela um melhor ajustamento do que o modelo 5, que não considera variáveis explicativas. Note-se que, uma vez que as medidas apresentadas não permitem a comparação directa dos modelos, não é possível, em rigor, dizer que um é melhor do que outro. Mas, uma vez que os modelos consideram exactamente a mesma trajectória, com e sem variáveis que expliquem a variabilidade individual em torno da trajectória média, não parece desadequada a conclusão de que a introdução de variáveis melhora o ajustamento da trajectória do factor de segunda ordem de debilidade psicológica.

Tabela 13 – Medidas de bondade do ajustamento dos modelos de trajectória latente condicionada (modelos 1c, 3c e 5c)

	Modelo 1c	Modelo 3c	Modelo 5c
CFI	0,986	0,955	0,888
TLI	0,974	0,942	0,968
RMSEA	0,033	0,056	0,066
SRMR	0,011	0,05	---
AIC	136475,679	226415,559	---
BIC	136610,615	226768,962	---
WRMR	---	---	3,463

Comparando agora o ajustamento modelo-dados obtido para os pares de modelos de trajectória do indicador único de bem-estar subjectivo (modelo 1 e modelo 1c) e de trajectória do factor latente de debilidade psicológica medido pelas variáveis contínuas *AD*, *SD* e *LC* (modelo 3 e modelo 3c), os critérios de informação *AIC* e *BIC* sugerem que a introdução de variáveis explicativas piora o ajustamento dos modelos. No entanto, é relevante notar que todas as medidas de bondade do ajustamento decrescem nos modelos considerando variáveis explicativas. Sublinhando mais uma vez que estamos a modelar exactamente a mesma trajectória, incluindo agora variáveis com vista à explicação da variabilidade dos indivíduos em torno da trajectória média, julgamos não ser desadequada a comparação destas medidas. Em suma, mesmo concluindo que o ajustamento não melhora com a introdução de variáveis explicativas, são incontestáveis os ganhos em capacidade explicativa relativamente aos modelos de trajectória não condicionada.

7. DISCUSSÃO DOS RESULTADOS

A análise dos resultados da estimação dos modelos de trajectória não condicionada de bem-estar psicológico permitiu concluir que, tal como sugeriam Shevlin e Adamson (2005), é pouca a capacidade explicativa adicional de um modelo mais complexo relativamente a um modelo mais simples que considera a trajectória do indicador único. Apesar de se terem evidenciado relações estatisticamente significativas entre as trajectórias das três dimensões de bem-estar psicológico, quer considerando indicadores construídos directamente pela soma das respostas dos indivíduos, quer a partir dos indicadores medidos em escalas de tipo Likert, as conclusões retiradas não vão muito além da constatação, já evidenciada, de que estas dimensões estão altamente correlacionadas. Mais ainda, os resultados da modelação das trajectórias dos três processos correlacionados (modelos 2 e 4), que consideram dimensões distintas de debilidade psicológica, pouco diferem entre eles relativamente às conclusões a retirar. Sugere-se que a única diferença encontrada entre os modelos 2 e 4 é a diluição das relações estatisticamente significativas evidenciadas no modelo com trajectórias das três variáveis, que poderá resultar da introdução de complexidade no modelo, ‘mediando’ as relações entre as trajectórias das dimensões por mais um nível de variáveis, *i.e.*, os factores latentes.

Poderia ser interessante averiguar se a introdução de variáveis explicativas teria impactos diferenciados nos três processos. Se tal acontecesse, poderíamos concluir que de facto uma estrutura tripartida permite evidenciar relações entre as trajectórias latentes e variáveis externas ao modelo, que não transparecem numa estrutura unifactorial. Apesar de se ter evidenciado que, tal como sugerem Graetz (1991), Cheung (2002), Gao *et al.* (2004), Campbell e Knowles (2007) e Ye (2008), a escala do *GHQ-12* pode encerrar várias dimensões de debilidade psicológica, a análise dos modelos não condicionados parece levar a uma conclusão, do ponto de vista conceptual, que contradiz a hipótese de que a utilização desta estrutura acrescenta capacidade explicativa ao modelo.

Por outro lado, embora o modelo que modela a trajectória de um factor de segunda ordem não acrescente muito a um modelo de trajectória do factor único de bem-estar psicológico, quando consideramos modelos condicionados somos levados a repensar esta ideia. De facto, quando introduzimos variáveis explicativas das trajectórias estimadas, quer para o factor único, quer

para o modelo que considera um factor de ordem superior às três dimensões de debilidade psicológica, no modelo mais complexo reúnem-se as conclusões retiradas relativamente aos dois modelos considerando as medidas repetidas contínuas (para o indicador único e para os três indicadores calculados). Esta conclusão sugere que afinal é possível ganhar especificidade estimando a trajectória de um modelo conceptual mais sofisticado, que considera as dimensões de ansiedade e depressão, disfunção social e perda de confiança como medindo um constructo comum, de debilidade psicológica.

Acrescente-se ainda que a estrutura factorial de segunda ordem utilizada trouxe para primeiro plano as conclusões retiradas por Clark (2002) e Clark e Oswald (2006) relativamente à relação que se estabelece entre o bem-estar subjectivo e a idade, evidenciando o impacto não linear da idade sobre a debilidade psicológica.

8. CONCLUSÃO

Como foi sendo referido ao longo do trabalho apresentado, os modelos de trajectória latente encerram um conjunto de características que lhes confere relevo no âmbito de estudos de carácter longitudinal, permitindo não só estimar trajectórias médias dos fenómenos em estudo, como ainda ter em linha de conta a heterogeneidade individual. Permite também uma fácil articulação entre o desenho de modelo teórico consistente e a sua estimação.

Em Portugal, os estudos que recorrem a esta metodologia contam-se pelos dedos de uma mão. Encontramos aplicações de *LGCM* apenas em Salgueiro (2009) e Salgueiro *et al.* (2008).

Duncan *et al.* (2006: 176-178) referem que “*latent growth modeling with categorical outcomes is an emerging methodology [and] the accommodation of these variables within LGM framework is still evolving [, and] applications of these methods are not widespread*”.

Um indicador de que o estado da arte neste contexto está ainda em evolução é sem dúvida a reduzida extensão dos capítulos dedicados a esta matéria nas obras de referência, onde os modelos de trajectória latente considerando variáveis ordinais estão geralmente entre outras extensões dos modelos de *LGCM*.

Se isto é verdade relativamente à utilização de variáveis ordinais, a modelação de um factor de segunda ordem confere a esta pesquisa um carácter duplamente inovador. De facto, nos vários artigos e obras que aplicam ou se dedicam inteiramente a modelos desta natureza, não se encontrou nenhum exemplo de aplicação a um modelo de trajectória remotamente semelhante ao que aqui se utiliza. No entanto, os autores do *software Mplus*, reconhecidamente figuras de destaque no contexto dos *LGCM*, asseguram que esta modelação é conceptualmente viável⁶⁶.

Naturalmente que o trabalho aqui desenvolvido fica longe de esgotar o debate quer em torno da modelação longitudinal do *GHQ-12*, quer relativamente à modelação de trajectórias latentes com medidas repetidas ordinais e, mais ainda, a modelos considerando um factor latente de segunda ordem. Relativamente à modelação longitudinal do *GHQ-12*, fica em aberto se os resultados obtidos serão válidos considerando um maior intervalo de tempo. Como refere Pavot (2008: 134-135), uma análise seccional do bem-estar psicológico permite

⁶⁶ Num dos fóruns do *Mplus*, Muthén, L.K. (2003) afirma: “*you can do three-level if one level is time. Otherwise, in a cross-sectional model, you can do two-level*”.

(<http://www.statmodel.com/discussion/messages/14/263.html?1046724068>, consultado a 28 de Junho de 2009)

avaliar como alguns factores interferem no bem-estar dos indivíduos, mas falham em compreender como é que essa interferência se processa ao longo do tempo. Estudos longitudinais, particularmente estudos que consideram intervalos temporais mais alargados, permitem evidenciar dinâmicas de causa-efeito entre o bem-estar dos indivíduos e um conjunto de factores externos. O autor refere como exemplo estudos longitudinais que mostram que indivíduos com maiores níveis iniciais de bem-estar subjectivo têm mais probabilidade de casar num momento temporal mais distante. Esta conclusão põe em causa a ideia, demonstrada em pesquisas com dados seccionais ou análises considerando curtos intervalos temporais, de que o casamento tem uma causa directa no bem-estar percebido, *i.e.*, que as pessoas casadas tendem a manifestar níveis de bem-estar mais elevados.

Diener (2008), referência incontestável no estudo da felicidade, refere a multiplicidade de factores relevantes no estudo do bem-estar subjectivo. Também Clark (2002) e Clark e Oswald (2002, 2006) sugerem, para além do sexo, da idade e do estado de saúde percebido, um conjunto mais alargado de variáveis que interferem no bem-estar subjectivo dos indivíduos, que não são aqui consideradas. O autor mostra, por exemplo, que o rendimento e a condição perante o trabalho são variáveis de especial interesse neste contexto. A introdução de variáveis como estas, cuja variação no tempo pode permitir evidenciar alterações significativas no bem-estar subjectivo dos indivíduos, sem dúvida acrescentaria capacidade explicativa, reduzindo certamente a margem do que se deixa por explicar neste ensaio. A introdução de variáveis explicativas com variações ao longo do tempo é uma sugestão que se deixa em aberto para trabalhos futuros.

Outro ponto, apenas superficialmente abordado neste estudo, e que nos parece ser da maior importância, remete para a verificação dos pressupostos subjacentes aos modelos de *LGCM*, necessária para a validação e boa estimação destes modelos. Aliás, sublinha-se esta falha como transversal a uma multiplicidade de aplicações do método consultadas. Por exemplo, apesar de se ter referido a questão da violação do pressuposto de normalidade bivariada das variáveis contínuas subjacentes às medidas repetidas ordinais, e de nos protegermos sob a capacidade dos estimadores utilizados produzirem estimativas consistentes quando tal acontece, os efeitos da violação deste pressuposto deveriam ser analisados em maior detalhe.

Por fim, a aplicação de modelos de trajectória latente a um factor de segunda ordem não foi encontrada em nenhum outro estudo. A replicação de uma estrutura factorial equivalente à que se apresenta traria certamente novas pistas para o estudo e a aplicação de modelos de *LGCM* a desenhos conceptuais complexos. Por um lado, seria importante validar a abordagem

ao *GHQ-12* adoptada neste trabalho recorrendo a outras amostras. Também a modelação de um factor de segunda ordem a partir de outras estruturas factoriais, aplicadas a outros fenómenos sociais ou económicos, poderia ajudar a resolver e reequacionar algumas das opções tomadas, contribuindo para alargar o debate em torno da aplicação dos *LGCM*.

Espera-se que este trabalho possa contribuir para futuras aplicações no âmbito dos modelos de trajectória latente. É nesse sentido que se introduzem em anexo os comandos para a estimação dos modelos nos dois *softwares* utilizados (*LISREL* e *Mplus*), com comentários que pretendem ajudar outros analistas que venham a utilizar a metodologia dos *LGCM* noutros contextos, facilitando a sua implementação ao nível mais técnico.

9. BIBLIOGRAFIA

- Berrington, A. e P.W.F. Smith (2006), *An Overview of Methods for the Analysis of Panel Data*. ESRC National Centre for Research Methods, NCRM Methods Review Paper (NCRM/007).
- Bollen, K.A. e P.J. Curran (2006), *Latent Curve Models. A Structural Equation Perspective*. New Jersey: John Wiley & Sons.
- Campbell, A. e S. Knowles (2007), A Confirmatory Factor Analysis of the GHQ-12 Using a Large Australian Sample, *European Journal of Psychological Assessment*, 23(1), 2-8.
- Cheung, Y. B. (2002), A confirmatory factor analysis of the 12-item General Health Questionnaire among older people, *International Journal of Geriatric Psychiatry*, 17, 739-744.
- Clark, A. (2002), Born to be mild: cohort effects in subjective well-being, DELTA mimeo.
- Clark, A. e A. Oswald (2002), Well-Being in Panels. University of Orléans: mimeo.
- Clark, A. e A. Oswald (2006), The curved relationship between subjective well-being and age, *Paris-Jourdan Sciences Economiques*, working paper n°. 2006-9, Centre National de la Recherche Scientifique.
- Diener, E. (2008), Myths in the Science of Happiness, and Directions for Future Research, em Eid, Michael e Randy J. Larsen (ed.), *The Science of Subjective Well-Being*, Nova Iorque: The Guilford Press, 493-514.
- Duncan, T., S. Duncan e L. Strycker (2006), *An Introduction to Latent Variable Growth Curve Modelling. Concepts, Issues, and Applications*. New Jersey: Lawrence Erlbaum Associates.
- Flora, D.B. e P.J. Curran (2004), An Empirical Evaluation of Alternative Methods of Estimation for Confirmatory Factor Analysis with Ordinal Data, *Psychological Methods*, American Psychological Association, 9(4), 466-491.
- Gao, F., N. Luo, J. Thumboo, C. Fones, S.C. Li e Y.B. Cheung (2004), Does the 12-item General Health Questionnaire contain multiple factors and do we need them?, *Health and Quality of Life Outcomes*, 2(63), 2-63.
- Graetz, B. (1991), Multidimensional properties of the General Health Questionnaire, *Social Psychiatry and Psychiatric Epidemiology*, 26,132-138.
- Jones, A. M. e J. Widman (2008), Health, income and relative deprivation: Evidence from BHPS, *Journal of Health Economics*, 27, 308-324.
- Hankins, M. (2008a), The factor structure of the twelve item General Health Questionnaire (GHQ-12): the result of negative phasing?, *Clinical Practice and Epidemiology in Mental Health*, 4:10.
- Hankins, M. (2008b), The reliability of the twelve-item General Health Questionnaire (GHQ-12) under realistic assumptions, *BMC Public Health*, 8:355.
- Muthén, B.O. (2004), *Mplus Technical Appendices*. Los Angeles, CA : Muthén & Muthén.
<http://www.statmodel.com/download/techappen.pdf>

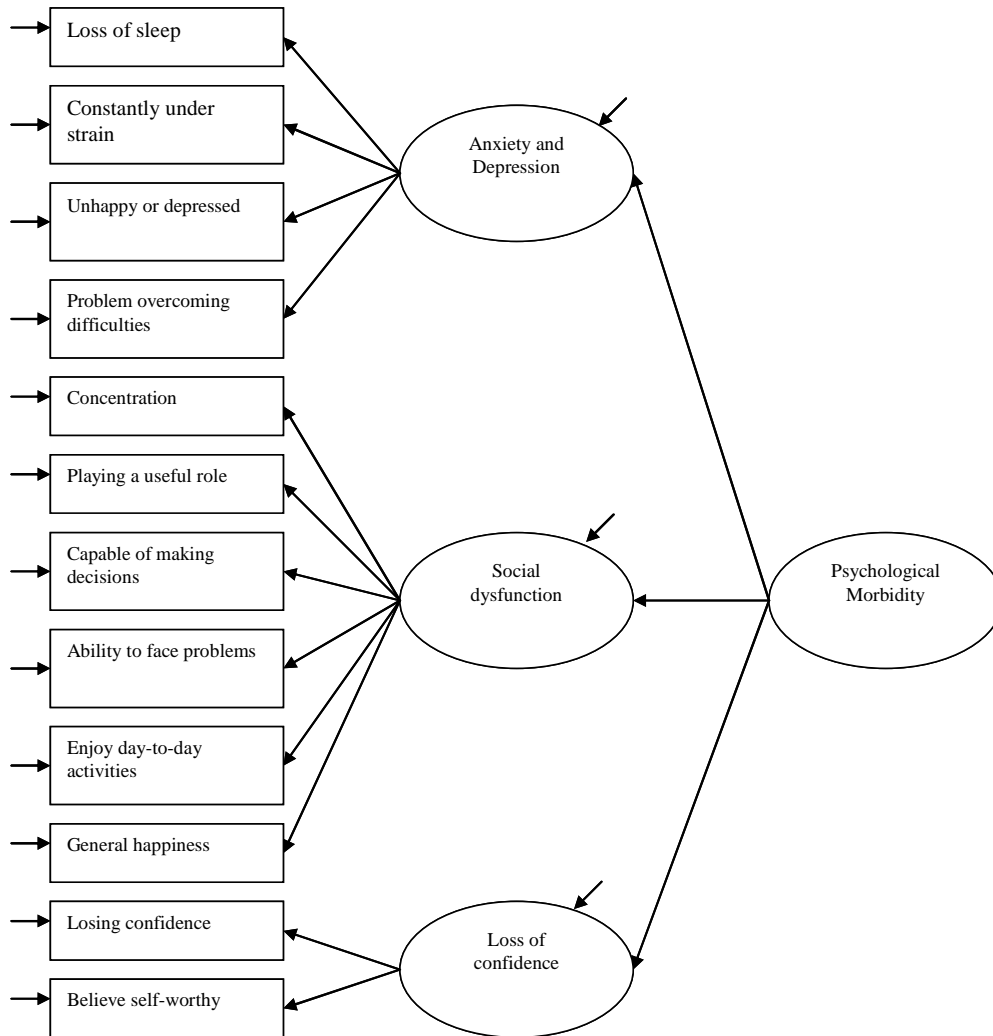
- Muthén, L.K. (2003), Three level LGM with latent variables, under the topic *Growth modeling of longitudinal data*, Mplus Discussion, <http://www.statmodel.com/discussion/messages/23/416.html>.
- Muthén, L.K. (2004), Fit indices. How is it possible!!!, under the topic *Categorical Data Modeling*, Mplus Discussion, <http://www.statmodel.com/discussion/messages/23/416.html>.
- Muthén, L.K. e B.O. Muthén (2007), *Mplus User's Guide*, 5ª edição. Los Angeles, CA : Muthén & Muthén. <http://www.statmodel.com/download/usersguide/Mplus%20Users%20Guide%20v5.pdf>
- Pavot, W. (2008), The Assessment of Subjective Well-Being, em Eid, Michael e Randy J. Larsen (ed.), *The Science of Subjective Well-Being*, Nova Iorque: The Guilford Press, 124-140.
- Preacher, K.J., A.L. Wichman, R.C. MacCallum e N.E. Briggs (2008), *Latent Growth Curve Modeling*. Thousand Oaks: Sage publications.
- Salgueiro, M.F. (2008), *Modelos de Equações Estruturais – Aplicações com LISREL*, manuscrito não publicado.
- Salgueiro, M.F. (2009), Modelling Job Satisfaction Trajectories in the British Household Panel Survey using Latent Growth Curve Models, em Pires Manso e J.D. Monteiro (dir.), *Anais de Economia Aplicada*, Nº XXIII, 581-593.
- Salgueiro, M.F., P.W.F. Smith e M.D.T. Vieira (2008), Modelação de trajectórias de satisfação com dados do British Household Panel Survey, *Programas e Resumos – XVI Congresso Anual da SPE*, Sociedade Portuguesa de Estatística, 140.
- Shevlin, M. e G. Adamson (2005), Alternative Factor Models and Factorial Invariance of the GHQ-12: A Large Sample Analysis Using Confirmatory Factor Analysis, *American Psychological Association*, 17(2), 231-236.
- Taylor, M.F. (ed.), with John Brice, Nick Buck and Elaine Prentice-Lane (2008), *British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices*. Colchester: University of Essex.
- Ye, S. (2008), Factor Structure of the General Health Questionnaire (GHQ-12): the role of wording effects, *Personality and Individual Differences*, 46, 197-201.

10. ANEXOS

10.1. LISTA DE VARIÁVEIS DA ESCALA GHQ-12 E INDICADOR ÚNICO DE SWB

Nome	Etiqueta	Questão	Escala de medida
wGHQB	GHQ: loss of sleep	Have you recently....lost much sleep over worry?	4 pontos: Not at all; No more than usual; Rather more thn usual; Much more than usual
wGHQE	GHQ: constantly under strain	Have you recently....felt constantly under strain?	4 pontos: Not at all; No more than usual; Rather more thn usual; Much more than usual
wGHQF	GHQ: problem overcoming difficulties	Have you recently....felt you couldn't overcome your difficulties?	4 pontos: Not at all; No more than usual; Rather more thn usual; Much more than usual
wGHQI	GHQ: unhappy or depressed	Have you recently....been feeling unhappy or depressed?	4 pontos: Not at all; No more than usual; Rather more thn usual; Much more than usual
wGHQA	GHQ: concentration	Have you recently....been able to concentrate on whatever you're doing?	4 pontos: More than usual; Same as usual; Less so; Much less
wGHQC	GHQ: playing a useful role	Have you recently....felt that you were playing a useful part in things?	4 pontos: More than usual; Same as usual; Less so; Much less
wGHQD	GHQ: capable of making decisions	Have you recently....felt capable of making decisions about things?	4 pontos: More than usual; Same as usual; Less so; Much less
wGHQG	GHQ: enjoy day-to-day activities	Have you recently....been able to enjoy your normal day-to- day activities?	4 pontos: More than usual; Same as usual; Less so; Much less
wGHQH	GHQ: ability to face problems	Have you recently....been able to face up to problems?	4 pontos: More than usual; Same as usual; Less so; Much less
wGHQL	GHQ: general happiness	Have you recently....been feeling reasonably happy, all things considered?	4 pontos: More so than usual; Same as usual; Less so than usual; Much less than usual
wGHQJ	GHQ: losing confidence	Have you recently....been losing confidence in yourself?	4 pontos: Not at all; No more than usual; Rather more thn usual; Much more than usual
wGHQK	GHQ: believe in self-worth	Have you recently....been thinking of yourself as a worthless person?	4 pontos: Not at all; No more than usual; Rather more thn usual; Much more than usual
wHLGHQ1	Subjective wellbeing (GHQ) 1: Likert	Uses OGHQA to OGHQL. This measure converts valid answers to questions wGHQA to wGHQL to a single scale by recoding so that the scale for individual variables runs from 0 to 3 instead of 1 to 4, and then summing, giving a scale running from 0 (the least distressed) to 36 (the most distressed). See Cox, B.D et al, The Health and Lifestyle Survey. (London: Health Promotion Research Trust, 1987).	

10.2. MODELO CONCEPTUAL EM ANÁLISE



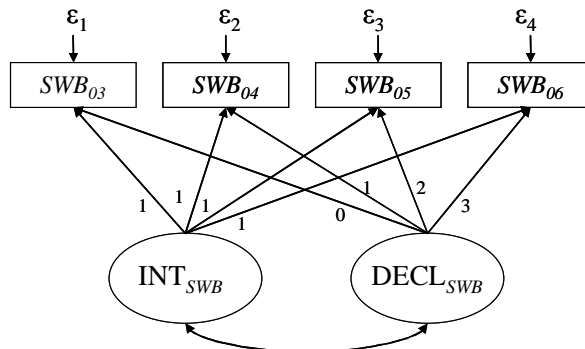
10.3. EQUAÇÕES, DEFINIÇÕES E PRESSUPOSTOS DE UM MODELO DE TRAJECTÓRIA LATENTE NÃO CONDICIONADA

Modelo não condicional		
$y_{it} = \alpha_i + \lambda_t \beta_i + \varepsilon_{it}$	Equação da trajetória	$i = 1, 2, \dots, N$ (N número total de casos); $t = 1, 2, \dots, T$ (T número total de waves)
$\alpha_i = \mu_\alpha + \zeta_{\alpha_i}$	Equação do intercepto individual	
$\beta_i = \mu_\beta + \zeta_{\beta_i}$	Equação do declive individual	
$y_{it} = (\mu_\alpha + \lambda_t \mu_\beta) + (\zeta_{\alpha_i} + \lambda_t \zeta_{\beta_i} + \varepsilon_{it})$	Equação do modelo combinado, ou forma reduzida	
Definições		
y_{it}	valor da variável de trajetória para o caso i , no momento t	
α_i	intercepto da trajetória do caso i	
β_i	declive da trajetória para o caso i	

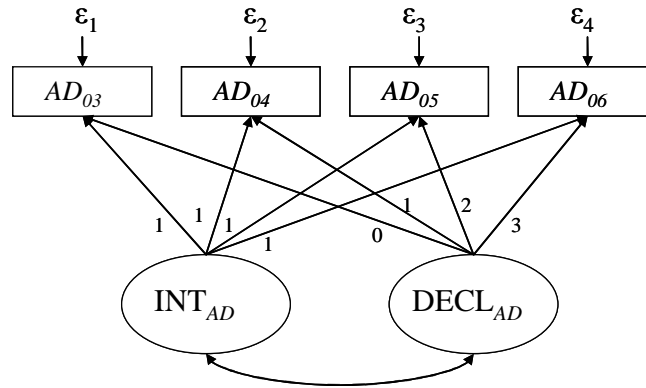
λ_i	valor da variável de tendência para o momento i
ε_{it}	parcela de erro de y_{it}
$\mu_\alpha = E(\alpha_i)$	média dos interceptos
$\mu_\beta = E(\beta_i)$	média dos declives
ζ_{α_i}	desvio do caso i a μ_α
ζ_{β_i}	desvio do caso i a μ_β
$\psi_{\alpha\alpha} = \text{var}(\zeta_{\alpha_i})$	variância dos interceptos
$\psi_{\beta\beta} = \text{var}(\zeta_{\beta_i})$	variância dos declives
$\psi_{\alpha\beta} = \text{cov}(\zeta_{\alpha_i}, \zeta_{\beta_i})$	covariância entre interceptos e declives
$\theta_{\varepsilon_{it}} = \text{var}(\varepsilon_{it})$	variância dos erros no momento t para o caso i
Pressupostos	
$E(\varepsilon_{it}) = 0, i = 1, 2, \dots, N; t = 1, 2, \dots, T$	erro tem média zero
$E(\zeta_{\alpha_i}) = 0, \text{ para } i = 1, 2, \dots, N$	erro tem média zero
$E(\zeta_{\beta_i}) = 0, \text{ para } i = 1, 2, \dots, N$	erro tem média zero
$\text{cov}(\varepsilon_{it}, \zeta_{\alpha_i}) = 0$	as duas parcelas de erro não estão correlacionadas
$\text{cov}(\varepsilon_{it}, \zeta_{\beta_i}) = 0$	as duas parcelas de erro não estão correlacionadas
$\text{cov}(\zeta_{\alpha_i}, \zeta_{\alpha_j}) = 0, \text{ para } i \neq j$	os desvios individuais relativamente à média dos interceptos não estão correlacionados
$\text{cov}(\zeta_{\beta_i}, \zeta_{\beta_j}) = 0, \text{ para } i \neq j$	os desvios individuais relativamente à média dos declives não estão correlacionados
$\text{cov}(\zeta_{\beta_i}, \zeta_{\alpha_j}) = 0, \text{ para } i \neq j$	os desvios individuais relativamente à media dos interceptos não estão correlacionados com os desvios individuais relativamente aos declives
$\text{cov}(\varepsilon_{it}, \varepsilon_{jt}) = 0, \text{ para } i \neq j$	o erro associado ao indivíduo i não está correlacionado com o erro associado ao indivíduo j

10.4. DIAGRAMAS DOS MODELOS ESTIMADOS

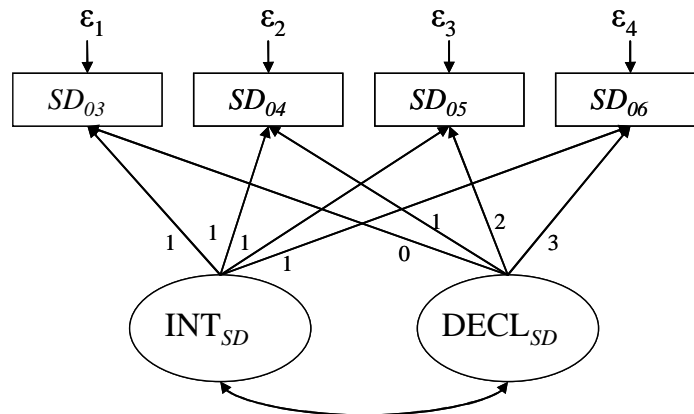
10.4.1. Diagrama do modelo 1: trajectória latente não condicionada do índice agregado de bem-estar subjectivo (SWB)



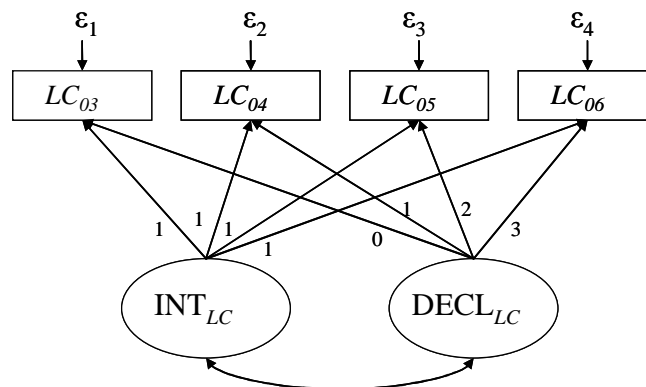
10.4.2. Diagrama do modelo AD: trajetória latente não condicionada do índice de ansiedade e depressão (AD)



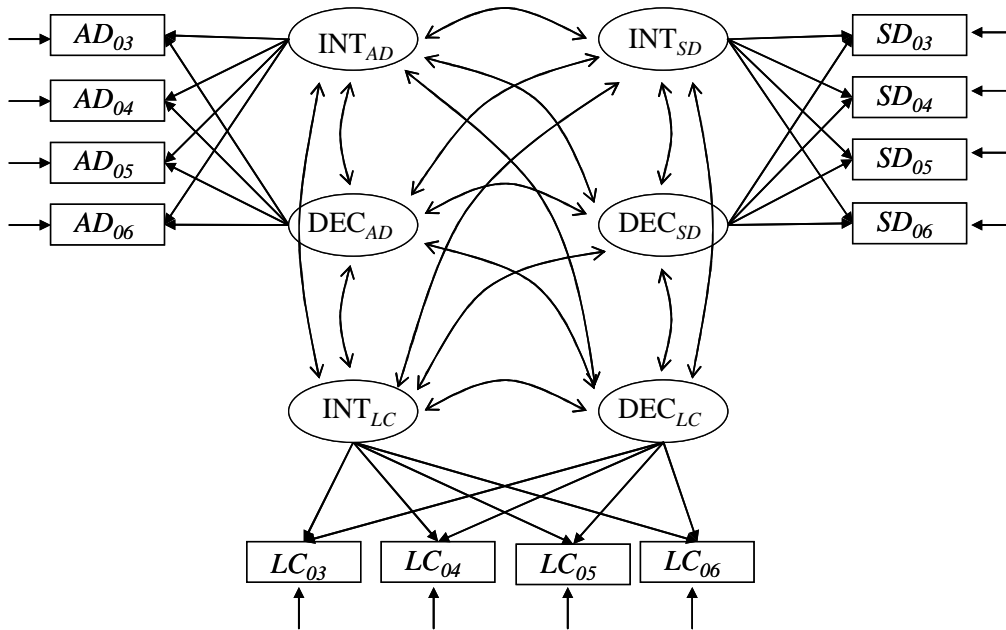
10.4.3. Diagrama do modelo SD: trajetória latente não condicionada do índice de disfunção social (SD)



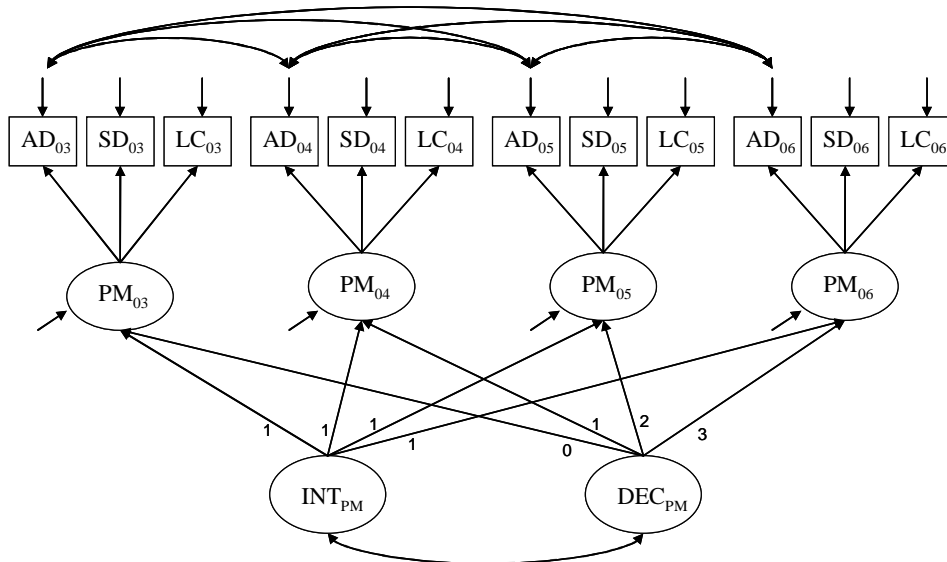
10.4.4. Diagrama do modelo LC: trajetória latente não condicionada do índice de perda de confiança (LC)



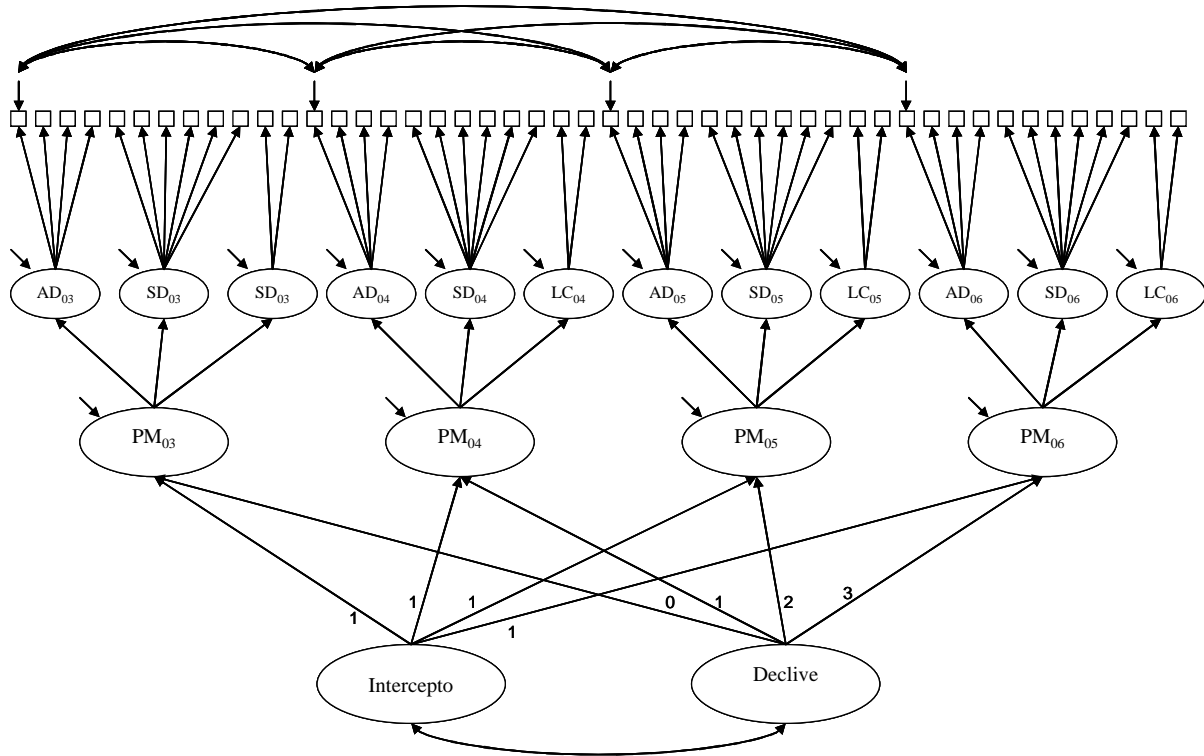
10.4.5. Diagrama do modelo 2: relacionando as trajectórias latentes dos três processos (AD, SD e LC)



10.4.6. Diagrama do modelo 3: trajectória latente não condicionada do factor de debilidade psicológica (PM), modelado a partir dos índices AD, SD e LC



10.4.8. Diagrama do modelo 5: trajectória latente do factor de segunda ordem de PM



NOTA: Por simplificação omitiu-se a representação dos erros das medidas repetidas e a correlação apresentada entre indicadores deve ler-se como a correlação entre os erros das medidas repetidas.

10.5. COMANDOS DE MPLUS UTILIZADOS PARA A ESTIMAÇÃO DO MODELO 4

Data:

File is ghqm_p.dat ;

Variable:

Names are

mghqb mghqe mghqf mghqi mghqa mghqc mghqd mghqg mghqh mghql mghqj
 mghqk pid nghqb nghqe nghqf nghqi nghqa nghqc nghqd nghqg nghqh nghql
 nghqj nghqk oghqb oghqe oghqf oghqi oghqa oghqc oghqd oghqg oghqh
 oghql oghqj oghqk PGHQB PGHQE PGHQF PGHQI PGHQA PGHQC PGHQD PGHQG
 PGHQH PGHQL PGHQJ PGHQK ;

Missing are all (-9999) ;

USEVARIABLES ARE mghqb mghqe mghqf mghqi mghqa mghqc mghqd mghqg
 mghqh mghql mghqj

mghqk nghqb nghqe nghqf nghqi nghqa nghqc nghqd nghqg nghqh nghql
 nghqj nghqk oghqb oghqe oghqf oghqi oghqa oghqc oghqd oghqg oghqh
 oghql oghqj oghqk PGHQB PGHQE PGHQF PGHQI PGHQA PGHQC PGHQD PGHQG
 PGHQH PGHQL PGHQJ PGHQK ;

CATEGORICAL ARE mghqb mghqe mghqf mghqi mghqa mghqc mghqd mghqq
 mghqh mghql mghqj
 mghqk nghqb nghqe nghqf nghqi nghqa nghqc nghqd nghqq nghqh nghql
 nghqj nghqk oghqb oghqe oghqf oghqi oghqa oghqc oghqd oghqq oghqh
 oghql oghqj oghqk PGHQB PGHQE PGHQF PGHQI PGHQA PGHQC PGHQD PGHQQ
 PGHQH PGHQL PGHQJ PGHQK;

MODEL: AnxDep03 BY mghqb
 mghqe-mghqi (1-3);
 AnxDep04 BY nghqb
 nghqe-nghqi (1-3);
 AnxDep05 BY oghqb
 oghqe-oghqi (1-3);
 AnxDep06 BY pghqb
 pghqe-pghqi (1-3);

[mghqb\$1 nghqb\$1 oghqb\$1 pghqb\$1] (4);
 [mghqb\$2 nghqb\$2 oghqb\$2 pghqb\$2] (5);
 [mghqb\$3 nghqb\$3 oghqb\$3 pghqb\$3] (6);
 [mghqe\$1 nghqe\$1 oghqe\$1 pghqe\$1] (7);
 [mghqe\$2 nghqe\$2 oghqe\$2 pghqe\$2] (8);
 [mghqe\$3 nghqe\$3 oghqe\$3 pghqe\$3] (9);
 [mghqf\$1 nghqf\$1 oghqf\$1 pghqf\$1] (10);
 [mghqf\$2 nghqf\$2 oghqf\$2 pghqf\$2] (11);
 [mghqf\$3 nghqf\$3 oghqf\$3 pghqf\$3] (12);
 [mghqi\$1 nghqi\$1 oghqi\$1 pghqi\$1] (13);
 [mghqi\$2 nghqi\$2 oghqi\$2 pghqi\$2] (14);
 [mghqi\$3 nghqi\$3 oghqi\$3 pghqi\$3] (15);

{mghqb-mghqi@1 nghqb nghqe nghqf nghqi oghqb oghqe oghqf
 oghqi PGHQB PGHQE PGHQF PGHQI};

IntAD SlopAD | AnxDep03@0 AnxDep04@1 AnxDep05@2 AnxDep06@3 ;

mghqb-mghqi PWITH nghqb-nghqi;
 mghqb-mghqi PWITH oghqb-oghqi;
 mghqb-mghqi PWITH pghqb-pghqi;
 nghqb-nghqi PWITH oghqb-oghqi;
 nghqb-nghqi PWITH pghqb-pghqi;
 oghqb-oghqi PWITH pghqb-pghqi;

SD03 BY mghqa
 mghqc-mghql (16-20);
 SD04 BY nghqa
 nghqc-nghql (16-20);
 SD05 BY oghqa
 oghqc-oghql (16-20);
 SD06 BY pghqa
 pghqc-pghql (16-20);

[mghqa\$1 nghqa\$1 oghqa\$1 pghqa\$1] (21);
 [mghqa\$2 nghqa\$2 oghqa\$2 pghqa\$2] (22);
 [mghqa\$3 nghqa\$3 oghqa\$3 pghqa\$3] (23);

[mghqc\$1 nghqc\$1 oghqc\$1 pghqc\$1] (24);
 [mghqc\$2 nghqc\$2 oghqc\$2 pghqc\$2] (25);
 [mghqc\$3 nghqc\$3 oghqc\$3 pghqc\$3] (26);

[mghqd\$1 nghqd\$1 oghqd\$1 pghqd\$1] (27);

```

[mghqd$2 nghqd$2 oghqd$2 pghqd$2] (28);
[mghqd$3 nghqd$3 oghqd$3 pghqd$3] (29);

[mghqg$1 nghqg$1 oghqg$1 pghqg$1] (30);
[mghqg$2 nghqg$2 oghqg$2 pghqg$2] (31);
[mghqg$3 nghqg$3 oghqg$3 pghqg$3] (32);

[mghqh$1 nghqh$1 oghqh$1 pghqh$1] (33);
[mghqh$2 nghqh$2 oghqh$2 pghqh$2] (34);
[mghqh$3 nghqh$3 oghqh$3 pghqh$3] (35);

[mghql$1 nghql$1 oghql$1 pghql$1] (36);
[mghql$2 nghql$2 oghql$2 pghql$2] (37);
[mghql$3 nghql$3 oghql$3 pghql$3] (38);

{mghqa-mghql@1 nghqa nghqc nghqd nghqg nghqh nghql
  oghqa oghqc oghqd oghqg oghqh oghql PGHQA PGHQC PGHQD PGHQG
  PGHQH PGHQL };

IntSD SlopSD | SD03@0 SD04@1 SD05@2 SD06@3 ;

mghqa-mghql PWITH nghqa-nghql ;
mghqa-mghql PWITH oghqa-oghql ;
mghqa-mghql PWITH pghqa-pghql ;
nghqa-nghql PWITH oghqa-oghql ;
nghqa-nghql PWITH pghqa-pghql ;
oghqa-oghql PWITH pghqa-pghql ;

LC03 BY mghqj
      mghqk (39);
LC04 BY nghqj
      nghqk (39);
LC05 BY oghqj
      oghqk (39);
LC06 BY pghqj
      pghqk (39);

[mghqj$1 nghqj$1 oghqj$1 pghqj$1] (40);
[mghqj$2 nghqj$2 oghqj$2 pghqj$2] (41);
[mghqj$3 nghqj$3 oghqj$3 pghqj$3] (42);
[mghqk$1 nghqk$1 oghqk$1 pghqk$1] (43);
[mghqk$2 nghqk$2 oghqk$2 pghqk$2] (44);
[mghqk$3 nghqk$3 oghqk$3 pghqk$3] (45);

{mghqj-mghqk@1 nghqj nghqk oghqj oghqk PGHQJ PGHQK};

IntLC SlopLC | LC03@0 LC04@1 LC05@2 LC06@3 ;

mghqj-mghqk PWITH nghqj-nghqk;
mghqj-mghqk PWITH oghqj-oghqk;
mghqj-mghqk PWITH pghqj-pghqk;
nghqj-nghqk PWITH oghqj-oghqk;
nghqj-nghqk PWITH pghqj-pghqk;
oghqj-oghqk PWITH pghqj-pghqk;

SlopAD SlopSD SlopLC PWITH IntAD IntSD IntLC ;

OUTPUT: STANDARDIZED tech4 ;

```

10.6. COMANDOS DE MPLUS UTILIZADOS PARA A ESTIMAÇÃO DO MODELO 5 E 5C (COM VARIÁVEIS EXPLICATIVAS) COMENTADOS

```

Data:
  File is GHQm_p_deter_invariantes.dat ;
[introdução do ficheiro dos dados]
Variable:
  Names are
    mghqb mghqe mghqf mghqi mghqa mghqc mghqd mghqg mghqh mghql mghqj
    mghqk nghqb nghqe nghqf nghqi nghqa nghqc nghqd nghqg nghqh nghql
    nghqj nghqk oghqb oghqe oghqf oghqi oghqa oghqc oghqd oghqg oghqh
    oghql oghqj oghqk PGHQB PGHQE PGHQF PGHQI PGHQA PGHQC PGHQD PGHQG
    PGHQH PGHQL PGHQJ PGHQK msex mhlstat1 mhlstat2 mhlstat3 mage12_1z
    mage12_2 mage12_3 mage12_4 pid;
[nome de todas as variáveis que constam na base de dados]
  Missing are all (-9999) ;
[definição do código para os valores omissos]
  USEVARIABLES ARE
    mghqb mghqe mghqf mghqi mghqa mghqc mghqd mghqg mghqh mghql mghqj
    mghqk nghqb nghqe nghqf nghqi nghqa nghqc nghqd nghqg nghqh nghql
    nghqj nghqk oghqb oghqe oghqf oghqi oghqa oghqc oghqd oghqg oghqh
    oghql oghqj oghqk PGHQB PGHQE PGHQF PGHQI PGHQA PGHQC PGHQD PGHQG
    PGHQH PGHQL PGHQJ PGHQK msex mhlstat1 mhlstat2 mhlstat3 mage12_2
    mage12_3
    mage12_4 ;
[variáveis a utilizar na estimação; as variáveis msex mhlstat2 mhlstat3 mage12_2
mage12_3 mage12_4 correspondem às variáveis explicativas]
  CATEGORICAL ARE mghqb mghqe mghqf mghqi mghqa mghqc mghqd mghqg mghqh
    mghql mghqj mghqk nghqb nghqe nghqf nghqi nghqa nghqc nghqd nghqg
    nghqh nghql nghqj nghqk oghqb oghqe oghqf oghqi oghqa oghqc oghqd
    oghqg oghqh oghql oghqj oghqk PGHQB PGHQE PGHQF PGHQI PGHQA PGHQC
    PGHQD PGHQG PGHQH PGHQL PGHQJ PGHQK;
[nome das variáveis que são categóricas]

MODEL:
  AnxDep03 BY mghqb
    mghqe-mghqi (1-3);
[variável latente 'ansiedade e depressão' (AnxDep) em 2003 é medida pelas variáveis
manifestas mghqb e mghqe até mghqi; os comandos seguintes são idênticos, mas
relativos aos restantes períodos]
  AnxDep04 BY nghqb
    nghqe-nghqi (1-3);
  AnxDep05 BY oghqb
    oghqe-oghqi (1-3);
  AnxDep06 BY pghqb
    pghqe-pghqi (1-3);

    [mghqb$1 nghqb$1 oghqb$1 pghqb$1] (4);
    [mghqb$2 nghqb$2 oghqb$2 pghqb$2] (5);
    [mghqb$3 nghqb$3 oghqb$3 pghqb$3] (6);
    [mghqe$1 nghqe$1 oghqe$1 pghqe$1] (7);
    [mghqe$2 nghqe$2 oghqe$2 pghqe$2] (8);
    [mghqe$3 nghqe$3 oghqe$3 pghqe$3] (9);
    [mghqf$1 nghqf$1 oghqf$1 pghqf$1] (10);
    [mghqf$2 nghqf$2 oghqf$2 pghqf$2] (11);
    [mghqf$3 nghqf$3 oghqf$3 pghqf$3] (12);
    [mghqi$1 nghqi$1 oghqi$1 pghqi$1] (13);
    [mghqi$2 nghqi$2 oghqi$2 pghqi$2] (14);
    [mghqi$3 nghqi$3 oghqi$3 pghqi$3] (15);

```

[restrição da igualdade de thresholds para cada variável ao longo do tempo]

```
{mghqb-mghqi@1  nghqb nghqe nghqf nghqi oghqb oghqe oghqf
  oghqi PGHQB PGHQE PGHQF PGHQI};
```

[factores de escala (*scaling factors*) restringidos a 1 no primeiro momento e livres nos momentos seguintes]

```
mghqb-mghqi PWITH nghqb-nghqi;
mghqb-mghqi PWITH oghqb-oghqi;
mghqb-mghqi PWITH pghqb-pghqi;
nghqb-nghqi PWITH oghqb-oghqi;
nghqb-nghqi PWITH pghqb-pghqi;
oghqb-oghqi PWITH pghqb-pghqi;
```

[estimação livre das variâncias e covariâncias dos erros das medidas repetidas ao longo dos períodos considerados]

```
SD03 BY mghqa
      mghqc-mghql (16-20);
```

[variável latente 'disfunção social' (SD) em 2003 é medida pelas variáveis manifestas mghqa e mghqc até mghql; os comandos seguintes são idênticos, mas relativos aos restantes períodos]

```
SD04 BY  nghqa
      nghqc-nghql (16-20);
SD05 BY  oghqa
      oghqc-oghql (16-20);
SD06 BY  pghqa
      pghqc-pghql (16-20);
```

```
[mghqa$1 nghqa$1 oghqa$1 pghqa$1] (21);
[mghqa$2 nghqa$2 oghqa$2 pghqa$2] (22);
[mghqa$3 nghqa$3 oghqa$3 pghqa$3] (23);
```

```
[mghqc$1 nghqc$1 oghqc$1 pghqc$1] (24);
[mghqc$2 nghqc$2 oghqc$2 pghqc$2] (25);
[mghqc$3 nghqc$3 oghqc$3 pghqc$3] (26);
```

```
[mghqd$1 nghqd$1 oghqd$1 pghqd$1] (27);
[mghqd$2 nghqd$2 oghqd$2 pghqd$2] (28);
[mghqd$3 nghqd$3 oghqd$3 pghqd$3] (29);
```

```
[mghqg$1 nghqg$1 oghqg$1 pghqg$1] (30);
[mghqg$2 nghqg$2 oghqg$2 pghqg$2] (31);
[mghqg$3 nghqg$3 oghqg$3 pghqg$3] (32);
```

```
[mghqh$1 nghqh$1 oghqh$1 pghqh$1] (33);
[mghqh$2 nghqh$2 oghqh$2 pghqh$2] (34);
[mghqh$3 nghqh$3 oghqh$3 pghqh$3] (35);
```

```
[mghql$1 nghql$1 oghql$1 pghql$1] (36);
[mghql$2 nghql$2 oghql$2 pghql$2] (37);
[mghql$3 nghql$3 oghql$3 pghql$3] (38);
```

[restrição da igualdade de thresholds para cada variável ao longo do tempo]

```
{mghqa-mghql@1  nghqa nghqc nghqd nghqg nghqh nghql oghqa oghqc oghqd
  oghqg oghqh oghql PGHQA PGHQC PGHQD PGHQG PGHQH PGHQL };
```

[factores de escala (*scaling factors*) restringidos a 1 no primeiro momento e livres nos momentos seguintes]

```
mghqa-mghql PWITH nghqa-nghql ;
mghqa-mghql PWITH oghqa-oghql ;
```

```

mghqa-mghql PWITH pghqa-pghql ;
nghqa-nghql PWITH oghqa-oghql ;
nghqa-nghql PWITH pghqa-pghql ;
oghqa-oghql PWITH pghqa-pghql ;
[estimação livre das variâncias e covariâncias dos erros das medidas
repetidas ao longo dos períodos considerados]

LC03 BY mghqj
      mghqk (39);
[variável latente 'perda de confiança' (LC) em 2003 é medida pelas variáveis
manifestas mghqj e mghqk; os comandos seguintes são idênticos, mas relativos aos
restantes períodos]
LC04 BY nghqj
      nghqk (39);
LC05 BY oghqj
      oghqk (39);
LC06 BY pghqj
      pghqk (39);

[mghqj$1 nghqj$1 oghqj$1 pghqj$1] (40);
[mghqj$2 nghqj$2 oghqj$2 pghqj$2] (41);
[mghqj$3 nghqj$3 oghqj$3 pghqj$3] (42);
[mghqk$1 nghqk$1 oghqk$1 pghqk$1] (43);
[mghqk$2 nghqk$2 oghqk$2 pghqk$2] (44);
[mghqk$3 nghqk$3 oghqk$3 pghqk$3] (45);
[restrição da igualdade de thresholds para cada variável ao longo do tempo]

{mghqj-mghqk@1 nghqj nghqk oghqj oghqk PGHQJ PGHQK};
[factores de escala (scaling factores) restringidos a 1 no primeiro momento e
livres nos momentos seguintes]

mghqj-mghqk PWITH nghqj-nghqk;
mghqj-mghqk PWITH oghqj-oghqk;
mghqj-mghqk PWITH pghqj-pghqk;
nghqj-nghqk PWITH oghqj-oghqk;
nghqj-nghqk PWITH pghqj-pghqk;
oghqj-oghqk PWITH pghqj-pghqk;
[estimação livre das variâncias e covariâncias dos erros das medidas
repetidas ao longo dos períodos considerados]

PM03 BY AnxDep03 SD03 LC03;
PM04 BY AnxDep04 SD04 LC04;
PM05 BY AnxDep05 SD05 LC05;
PM06 BY AnxDep06 SD06 LC06;
[relação estrutural entre factores de primeira e de segunda ordem ao longo
dos anos em análise]

IntPM SlopPM | PM03@0 PM04@1 PM05@2 PM06@3 ;
[Definição do modelo de trajectória linear do factor de segunda ordem]

IntPM SlopPM ON msex mhlstat2 mhlstat3 mage12_2
mage12_3 mage12_4 ;
[Introdução de variáveis explicativas na trajectória do factor de segunda
ordem]

OUTPUT: STANDARDIZED tech1 tech4 ;
[apresentar estimativas estandardizadas dos parâmetros e respectivos erros-
padrão; apresentar médias, covariâncias e correlações estimadas para as
variáveis latentes]

```


NOTA: a opção tech1 é utilizada para que sejam impressas as matrizes e vectores contendo as especificações e os valores iniciais de todos os parâmetros livres no modelo; a opção tech4 é utilizada para pedir estimativas das médias, covariâncias e correlações para as variáveis latentes do modelo.

10.7. COMANDOS DE LISREL UTILIZADOS PARA A ESTIMAÇÃO DO MODELO 4

10.7.1. Passo 1: comandos para a estimação dos thresholds e da matriz de correlações policóricas comentados

```

ESPECIFICAÇÕES EM PRELIS PARA A ESTIMAÇÃO DA MATRIZ DE CORRELAÇÕES
POLICÓRICAS,
DESVIOS-PADRÃO E VECTOR DE MÉDIAS PARA UM MODELO DE LGCM COM VARIÁVEIS
ORDINAIS
DA NI=48 NO=4562
LA
mghqb mghqe mghqf mghqi mghqa mghqc mghqd mghqg mghqh mghql mghqj mghqk
nghqb nghqe nghqf nghqi nghqa nghqc nghqd nghqg nghqh nghql nghqj nghqk
oghqb oghqe oghqf oghqi oghqa oghqc oghqd oghqg
oghqh oghql oghqj oghqk PGHQB PGHQE PGHQF PGHQI PGHQA PGHQC PGHQD PGHQG
PGHQH PGHQL PGHQJ PGHQK
RA=C:\LGCM_LISREL\LGCM_corrigido\GHQ12_LGunc.dat
[utilização da base de dados em formato livre]
OR ALL
[todas as variáveis ordinais]
ET mghqb nghqb oghqb PGHQB
ET mghqe nghqe oghqe PGHQE
ET mghqf nghqf oghqf PGHQF
ET mghqi nghqi oghqi PGHQI
ET mghqa nghqa oghqa PGHQA
ET mghqc nghqc oghqc PGHQC
ET mghqd nghqd oghqd PGHQD
ET mghqg nghqg oghqg PGHQG
ET mghqh nghqh oghqh PGHQH
ET mghql nghql oghql PGHQL
ET mghqj nghqj oghqj PGHQJ
ET mghqk nghqk oghqk PGHQK
[restrição de igualdade de thresholds ao longo do tempo para as variável
subjacentes a cada variável ordinal manifesta]
OU MA=PM PM=C:\LGCM_LISREL\LGCM_corrigido\GHQ12_LGunc.PM
[cálculo, e gravação em ficheiro autónomo, da matriz de correlações
policóricas]

```

10.7.2. Passo 2: comandos para a estimação da matriz de covariâncias assintóticas comentados

```

ESPECIFICAÇÕES EM PRELIS PARA A ESTIMAÇÃO DA MATRIZ DE COVARIÂNCIAS
ASSIMPTÓTICAS DAS CORRELAÇÕES AMOSTRAIS
SOB O PRESSUPOSTO DE IGUALDADE DE THRESHOLDS ENTRE AS MEDIDAS REPETIDAS
DA NI=48
LA

```

```

mghqb mghqe mghqf mghqi mghqa mghqc mghqd mghqg mghqh mghql mghqj mghqk
nghqb nghqe nghqf nghqi nghqa nghqc nghqd nghqg nghqh nghql nghqj nghqk
oghqb oghqe oghqf oghqi oghqa oghqc oghqd oghqg
oghqh oghql oghqj oghqk PGHQB PGHQE PGHQF PGHQI PGHQA PGHQC PGHQD PGHQG
PGHQH PGHQL PGHQJ PGHQK
RA=C:\LGCM_LISREL\LGCM_corrigido\GHQ12_LGunc.dat
OR ALL
ET mghqb nghqb oghqb PGHQB
ET mghqe nghqe oghqe PGHQE
ET mghqf nghqf oghqf PGHQF
ET mghqi nghqi oghqi PGHQI
ET mghqa nghqa oghqa PGHQA
ET mghqc nghqc oghqc PGHQC
ET mghqd nghqd oghqd PGHQD
ET mghqg nghqg oghqg PGHQG
ET mghqh nghqh oghqh PGHQH
ET mghql nghql oghql PGHQL
ET mghqj nghqj oghqj PGHQJ
ET mghqk nghqk oghqk PGHQK
OU MA=CM AC=C:\LGCM_LISREL\LGCM_corrigido\GHQ12_unc.ACC
[cálculo, e gravação em ficheiro autónomo, da matriz de covariâncias
assimptóticas]

```

10.7.3. Passo 3: comandos para a estimação do modelo 4

```

COMANDOS PARA A ESTIMAÇÃO DE UM LGCM COM VARIÁVEIS ORDINAIS (MODELO 4)
Trajectória latente com utilização da matriz de correlações policóricas
vector de médias e de desvios-padrão e matriz de covariâncias assimptóticas
obtidas em Prelis, nos passos 1 e 2
DA NI=48 NO=4562 MA=CM
LA
mghqb mghqe mghqf mghqi mghqa mghqc mghqd mghqg mghqh mghql mghqj mghqk
nghqb nghqe nghqf nghqi nghqa nghqc nghqd nghqg nghqh nghql nghqj nghqk
oghqb oghqe oghqf oghqi oghqa oghqc oghqd oghqg
oghqh oghql oghqj oghqk PGHQB PGHQE PGHQF PGHQI PGHQA PGHQC PGHQD PGHQG
PGHQH PGHQL PGHQJ PGHQK
PM FI=C:\LGCM_LISREL\LGCM_corrigido\GHQ12_LGunc.PM
AC FI=C:\LGCM_LISREL\LGCM_corrigido\GHQ12_unc.ACC
MODEL NY=48 NE=18 AL=FI PS=SY,FI LY=FU,FI BE=SD,FI TE=SY,FI
LE
INTAnx SLpAnx INTSoc SLPSoc INTLoss SLPLoss ANXDEPR1 SOCIDISF1 LOSSCONF1
ANXDEPR2 SOCIDISF2 LOSSCONF2 ANXDEPR3 SOCIDISF3 LOSSCONF3 ANXDEPR4
SOCIDISF4 LOSSCONF4
VA 1 LY 1 7 LY 5 8 LY 11 9 LY 13 10 LY 17 11 LY 23 12 LY 25 13 LY 29 14 LY
35 15 LY 37 16 LY 41 17 LY 47 18
FR LY 2 7 LY 3 7 LY 4 7 LY 6 8 LY 7 8 LY 8 8 LY 9 8 LY 10 8 LY 12 9
EQ LY 2 7 LY 14 10 LY 26 13 LY 38 16
EQ LY 3 7 LY 15 10 LY 27 13 LY 39 16
EQ LY 4 7 LY 16 10 LY 28 13 LY 40 16
EQ LY 6 8 LY 18 11 LY 30 14 LY 42 17
EQ LY 7 8 LY 19 11 LY 31 14 LY 43 17
EQ LY 8 8 LY 20 11 LY 32 14 LY 44 17
EQ LY 9 8 LY 21 11 LY 33 14 LY 45 17
EQ LY 10 8 LY 22 11 LY 34 14 LY 46 17
EQ LY 12 9 LY 24 12 LY 36 15 LY 48 18
VA 0 BE 7 2 BE 8 4 BE 9 6

```

```

VA 1 BE 7 1 BE 10 1 BE 13 1 BE 16 1 BE 10 2 BE 8 3 BE 11 3 BE 14 3 BE 17 3
BE 11 4 BE 9 5 BE 12 5 BE 15 5 BE 18 5 BE 12 6
VA 2 BE 13 2 BE 14 4 BE 15 6
VA 3 BE 16 2 BE 17 4 BE 18 6
FR PS 1 1 PS 2 2 PS 3 3 PS 4 4 PS 5 5 PS 6 6 PS 7 7 PS 8 8 PS 9 9 PS 10 10
PS 11 11 PS 12 12 PS 13 13 PS 14 14 PS 15 15 PS 16 16 PS 17 17 PS 18 18
FR PS 1 3 PS 2 4 PS 5 6 PS 1 4 PS 1 6 PS 3 2 PS 3 5 PS 2 6 PS 4 5 PS 3 4 PS
1 2 PS 2 5 PS 4 6
FR TE 1 1 TE 2 2 TE 3 3 TE 4 4 TE 5 5 TE 6 6 TE 7 7 TE 8 8 TE 9 9 TE 10 10
TE 11 11 TE 12 12 TE 13 13 TE 14 14 TE 15 15 TE 16 16 TE 17 17 TE 18 18 TE
19 19 TE 20 20 TE 21 21 TE 22 22 TE 23 23 TE 24 24 TE 25 25 TE 26 26 TE 27
27
FR TE 28 28 TE 29 29 TE 30 30 TE 31 31 TE 32 32 TE 33 33 TE 34 34 TE 35 35
TE 36 36 TE 37 37 TE 38 38 TE 39 39 TE 40 40 TE 41 41 TE 42 42 TE 43 43 TE
44 44 TE 45 45 TE 46 46 TE 47 47 TE 48 48
FR TE 1 13 TE 13 25 TE 25 37 TE 2 14 TE 14 26 TE 26 38 TE 3 15 TE 15 27 TE
27 39 TE 4 16 TE 16 28 TE 28 40 TE 5 17 TE 17 29 TE 29 41 TE 6 18 TE 18 30
TE 30 42 TE 7 19 TE 19 31 TE 31 43 TE 8 20 TE 20 32 TE 32 44 TE 9 21 TE 21
33 TE 33 45
FR TE 10 22 TE 22 34 TE 34 46 TE 11 23 TE 23 35 TE 35 47 TE 12 24 TE 24 36
TE 36 48
FR TE 1 25 TE 1 37 TE 2 26 TE 2 38 TE 3 27 TE 3 39 TE 4 28 TE 4 40 TE 5 29
TE 5 41 TE 6 30 TE 6 42 TE 7 31 TE 7 43 TE 8 32 TE 8 44 TE 9 33 TE 9 45 TE
10 34 TE 10 46 TE 11 35 TE 11 47 TE 12 36 TE 12 48
FR TE 13 37 TE 14 38 TE 15 39 TE 16 40 TE 17 41 TE 18 42 TE 19 43 TE 20 44
TE 21 45 TE 22 46 TE 23 47 TE 24 48
FR AL 1 AL 2 AL 3 AL 4 AL 5 AL 6
ME
-0.039      -0.020      -0.015      0.000      -0.050      -0.026      -0.044
-0.056      -0.024      -0.056      -0.035      -0.049      0.017      -0.025
-0.029      -0.007      0.026      -0.005      0.017      0.039      0.022
0.023      -0.007      0.017      0.007      0.018      -0.022      0.032
0.023      -0.010      -0.005      0.010      -0.015      0.002      0.024
0.019      0.015      0.026      0.011      0.028      0.001      0.040
0.032      0.006      0.017      0.030      0.018      0.012
SD
0.986      0.999      0.949      1.025      0.998      0.989      0.989
1.011      1.011      1.018      0.991      1.014      0.997      1.012
0.980      1.052      0.983      1.000      1.010      0.982      1.005
0.977      0.994      0.986      1.026      0.998      0.895      1.100
1.004      1.020      1.006      0.998      0.977      1.001      0.985
0.974      0.990      0.990      0.965      1.020      1.014      0.990
0.995      1.008      1.006      1.003      1.030      1.025
PD
OU SC SE TV ND=5 AD=OFF

```

10.8. COMANDOS DE LISREL UTILIZADOS PARA A ESTIMAÇÃO DO MODELO 5 COMENTADOS

Os passos 1 e 2 são idênticos aos apresentados anteriormente. Apresentam-se apenas os comandos para a estimação do modelo de trajectória do factor de segunda ordem de debilidade psicológica.

```

COMANDOS PARA A ESTIMAÇÃO DE UM LGCM DE UM FACTOR DE SEGUNDA ORDEM COM
VARIÁVEIS ORDINAIS (MODELO 5)
Trajectória latente com utilização da matriz de correlações policóricas
vector de médias e de desvios-padrão e matriz de covariâncias assimpóticas
obtidas em Prelis, nos passos 1 e 2
DA NI=48 NO=4562 MA=CM
LA
mghqb mghqe mghqf mghqi mghqa mghqc mghqd mghqg mghqh mghql mghqj mghqk
nghqb nghqe nghqf nghqi nghqa nghqc nghqd nghqg nghqh nghql nghqj nghqk
oghqb oghqe oghqf oghqi oghqa oghqc oghqd oghqg
oghqh oghql oghqj oghqk PGHQB PGHQE PGHQF PGHQI PGHQA PGHQC PGHQD PGHQG
PGHQH PGHQL PGHQJ PGHQK
PM FI=C:\LGCM_LISREL\LGCM_corrigido\GHQ12_LGunc.PM
[introdução da matriz de correlações policóricas gravada no primeiro passo]
AC FI=C:\LGCM_LISREL\LGCM_corrigido\GHQ12_unc.ACC
[introdução da matriz de covariâncias assimpóticas gravada no segundo
passo]
MODEL NY=48 NE=18 AL=FI PS=SY,FI LY=FU,FI BE=SD,FI TE=SY,FI
LE
INTER SLOPE PHMORB1 PHMORB2 PHMORB3 PHMORB4 ANXDEPR1 SOCIDISF1 LOSSCONF1
ANXDEPR2 SOCIDISF2 LOSSCONF2 ANXDEPR3 SOCIDISF3 LOSSCONF3 ANXDEPR4
SOCIDISF4 LOSSCONF4
VA 1 LY 1 7 LY 5 8 LY 11 9 LY 13 10 LY 17 11 LY 23 12 LY 25 13 LY 29 14 LY
35 15 LY 37 16 LY 41 17 LY 47 18
[atribuição de escala a cada variável latente]
FR LY 2 7 LY 3 7 LY 4 7 LY 6 8 LY 7 8 LY 8 8 LY 9 8 LY 10 8 LY 12 9
[estimação livre dos restantes indicadores de cada variável latente]
EQ LY 2 7 LY 14 10 LY 26 13 LY 38 16
EQ LY 3 7 LY 15 10 LY 27 13 LY 39 16
EQ LY 4 7 LY 16 10 LY 28 13 LY 40 16
EQ LY 6 8 LY 18 11 LY 30 14 LY 42 17
EQ LY 7 8 LY 19 11 LY 31 14 LY 43 17
EQ LY 8 8 LY 20 11 LY 32 14 LY 44 17
EQ LY 9 8 LY 21 11 LY 33 14 LY 45 17
EQ LY 10 8 LY 22 11 LY 34 14 LY 46 17
EQ LY 12 9 LY 24 12 LY 36 15 LY 48 18
[restrição de invariância das medidas repetidas ao longo do tempo]
VA 0 BE 3 2
VA 1 BE 3 1 BE 4 1 BE 5 1 BE 6 1 BE 4 2
VA 2 BE 5 2
VA 3 BE 6 2
[definição dos pesos que definem a trajectória linear]
FR BE 7 3 BE 8 3 BE 9 3 BE 10 4 BE 11 4 BE 12 4 BE 13 5 BE 14 5 BE 15 5 BE
16 6 BE 17 6 BE 18 6
[estimação livre da relação entre variáveis latentes de primeira e segunda
ordem]
FR PS 1 1 PS 2 2 PS 3 3 PS 4 4 PS 5 5 PS 1 2 PS 6 6 PS 7 7 PS 8 8 PS 9 9 PS
10 10 PS 11 11 PS 12 12 PS 13 13 PS 14 14 PS 15 15 PS 16 16 PS 17 17 PS 18
18
[estimação livre das variâncias dos erros das relações estruturais]
FR TE 1 1 TE 2 2 TE 3 3 TE 4 4 TE 5 5 TE 6 6 TE 7 7 TE 8 8 TE 9 9 TE 10 10
TE 11 11 TE 12 12 TE 13 13 TE 14 14 TE 15 15 TE 16 16 TE 17 17 TE 18 18 TE
19 19 TE 20 20 TE 21 21 TE 22 22 TE 23 23 TE 24 24 TE 25 25 TE 26 26 TE 27
27
FR TE 28 28 TE 29 29 TE 30 30 TE 31 31 TE 32 32 TE 33 33 TE 34 34 TE 35 35
TE 36 36 TE 37 37 TE 38 38 TE 39 39 TE 40 40 TE 41 41 TE 42 42 TE 43 43 TE
44 44 TE 45 45 TE 46 46 TE 47 47 TE 48 48
[estimação livre das variâncias dos erros de medida das variáveis
manifestas]

```

FR TE 1 13 TE 13 25 TE 25 37 TE 2 14 TE 14 26 TE 26 38 TE 3 15 TE 15 27 TE
 27 39 TE 4 16 TE 16 28 TE 28 40 TE 5 17 TE 17 29 TE 29 41 TE 6 18 TE 18 30
 TE 30 42 TE 7 19 TE 19 31 TE 31 43 TE 8 20 TE 20 32 TE 32 44 TE 9 21 TE 21
 33 TE 33 45
 FR TE 10 22 TE 22 34 TE 34 46 TE 11 23 TE 23 35 TE 35 47 TE 12 24 TE 24 36
 TE 36 48
 FR TE 1 25 TE 1 37 TE 2 26 TE 2 38 TE 3 27 TE 3 39 TE 4 28 TE 4 40 TE 5 29
 TE 5 41 TE 6 30 TE 6 42 TE 7 31 TE 7 43 TE 8 32 TE 8 44 TE 9 33 TE 9 45 TE
 10 34 TE 10 46 TE 11 35 TE 11 47 TE 12 36 TE 12 48
 FR TE 13 37 TE 14 38 TE 15 39 TE 16 40 TE 17 41 TE 18 42 TE 19 43 TE 20 44
 TE 21 45 TE 22 46 TE 23 47 TE 24 48

[estimação livre das covariâncias dos erros das medidas repetidas ao longo dos períodos considerados]

FR AL 1 AL 2

[estimação livre das médias do intercepto e do declive]

ME

-0.039	-0.020	-0.015	0.000	-0.050	-0.026	-0.044
-0.056	-0.024	-0.056	-0.035	-0.049	0.017	-0.025
-0.029	-0.007	0.026	-0.005	0.017	0.039	0.022
0.023	-0.007	0.017	0.007	0.018	-0.022	0.032
0.023	-0.010	-0.005	0.010	-0.015	0.002	0.024
0.019	0.015	0.026	0.011	0.028	0.001	0.040
0.032	0.006	0.017	0.030	0.018	0.012	

[médias calculadas no primeiro passo - das latentes subjacentes à medidas repetidas ordinais]

SD

0.986	0.999	0.949	1.025	0.998	0.989	0.989
1.011	1.011	1.018	0.991	1.014	0.997	1.012
0.980	1.052	0.983	1.000	1.010	0.982	1.005
0.977	0.994	0.986	1.026	0.998	0.895	1.100
1.004	1.020	1.006	0.998	0.977	1.001	0.985
0.974	0.990	0.990	0.965	1.020	1.014	0.990
0.995	1.008	1.006	1.003	1.030	1.025	

[desvios-padrão calculados no primeiro passo - das latentes subjacentes à medidas repetidas ordinais]

PD

OU SC SE TV ND=5 AD=OFF