

Mestrado em Ciência de Dados

Modelação Bayesiana

Docente: Professor José Gonçalves Dias

Projeto

2024/2025

Alunos:

125712, Bruno Santos
126625, Eduardo Bagulho
125864, Gonçalo Rosa
125767, Miguel Peixoto

19 de maio de 2025

Conteúdo

Lista de Figuras	ii
Lista de Tabelas	iii
1 Introdução	1
2 Descrição do Conjunto de Dados	3
2.1 Imputação de Dados	4
3 Identificação do modelo de medida	6
3.1 Exploração do modelo	6
3.1.1 Distribuições <i>a priori</i> e hiperparâmetros	6
3.1.2 Estimação dos parâmetros do modelo	7
3.2 Discussão dos resultados	8
4 Identificação do modelo estrutural	10
4.1 Discussão dos resultados	11
4.2 Melhoramento do modelo	11
4.3 Aplicação de <i>priors</i> informativas	11
5 Conclusão	13
A Modelo de Medida	15
B Definição do Modelo de Medida	16
B.1 Variáveis Manifestas	16
B.2 Distribuição dos Construtos Latentes	17
C Modelo Estrutural	18
D Modelo Estrutural Refinado	19

Lista de Figuras

1	Modelo de medida	15
1	Modelo Estrutural	18
1	Modelo Estrutural Refinado	19

Lista de Tabelas

2.1	Construtos, variáveis manifestas e respectivas questões. [1]	3
2.2	Média dos erros entre a CFA a partir dos conjuntos de dados pré-processados e os cross divulgados pelos autores.	5
2.3	Comparação do número de parâmetros entre o artigo original e o número obtido por CFA.	5
3.1	Distribuições <i>a priori</i> não informativas utilizadas	7
3.2	CFA de cada construto latente	7
3.3	Parâmetros utilizados na chamada da função <code>bcfa()</code>	8
3.4	CFA com todos os construtos latentes	8
4.1	Distribuições <i>a priori</i> não informativas utilizadas	10
4.2	Parâmetros utilizados na chamada da função <code>bcfa()</code>	10
4.3	<i>Cross-loadings</i> obtidos através da máxima verosimilhança	11
4.4	<i>Cross-loadings</i> obtidos através da máxima verosimilhança com e sem priors informativas	12

Capítulo 1

Introdução

A crescente adoção de sistemas *Enterprise Resource Planning* (ERP) transformou a forma como as organizações gerem os seus processos internos, visando uma maior eficiência, integração e tomada de decisões baseadas em dados. No entanto, embora os sistemas ERP racionalizem as operações, o seu papel na promoção da *Business Model Innovation* (BMI) continua a ser uma área de investigação em curso. A literatura existente sugere que a Complexidade Organizacional (OC) - conceito que engloba as complexidades estruturais, processuais e estratégicas das empresas - pode ter um impacto significativo na forma como a BMI se desenvolve quando as soluções ERP são implementadas. Apesar do potencial do ERP para impulsionar a inovação, a sua eficácia em permitir a BMI depende do grau de CO dentro de uma organização [1].

Este estudo tem como objetivo examinar empiricamente a forma como OC serve como um “facilitador-chave” na relação entre ERP e BMI, abordando uma lacuna na investigação anterior em que o papel mediador de OC não foi completamente explorado. Utilizando medidas validadas de vários itens e quadros teóricos estabelecidos, o conjunto de dados recolhidos para este estudo capta as percepções de 132 profissionais de várias empresas, das quais 104 organizações utilizam ativamente sistemas ERP. A investigação emprega análise quantitativa através de modelação de equações estruturais (SEM) para testar a hipótese de que a OC medeia o efeito do ERP no BMI. As variáveis-chave incluem complexidade tecnológica e utilidade percebida do ERP, bem como custos de inovação e receitas associadas à BMI [1].

Ao aproveitar o conhecimento prévio de estudos sobre implementação de ERP, mecanismos de BMI e teoria de OC, esta investigação contribui para o discurso mais alargado sobre transformação digital e gestão estratégica de empresas. Os resultados têm implicações substanciais para as empresas que procuram aproveitar a inovação impulsionada pelo ERP enquanto navegam nas complexidades das estruturas organizacionais. Além disso, o conjunto de dados fornece uma base para estudos futuros, permitindo comparações inter-industriais e transnacionais da forma como o ERP, CO e BMI interagem em diferentes ambientes empresariais [1].

O presente trabalho tem como principal objetivo replicar o estudo desenvolvido pelos autores do artigo científico, recorrendo à modelação bayesiana. Para além da reprodução do modelo original, pretende-se ainda explorar o potencial desta abordagem no aprofundamento da análise, nomeadamente no que respeita à formulação de novas questões relacionadas com os construtos em estudo e à compreensão das inter-relações existentes entre os mesmos.

Este relatório inicia-se com uma introdução ao problema abordado no artigo original,

incluindo uma breve descrição do processo de recolha dos dados e dos critérios utilizados para a seleção dos questionários incluídos no estudo. Segue-se uma secção dedicada à análise do método adotado pelos autores para o tratamento dos valores em falta no conjunto de dados. No capítulo seguinte, procede-se à exploração da estrutura dos dados, com a identificação das variáveis manifestas que, teoricamente, são explicadas pelos respetivos construtos latentes. Apresenta-se de seguida o modelo de medida (CFA), que descreve as relações entre os construtos latentes e variáveis manifestas, seguido pela formulação do modelo estrutural (SEM), o qual visa compreender de que forma determinadas variáveis latentes influenciam outras. Por fim, discutem-se as conclusões retiradas da análise, com destaque para os principais resultados obtidos e eventuais discrepâncias face ao artigo original.

Capítulo 2

Descrição do Conjunto de Dados

O conjunto de dados utilizado resulta de um questionário aplicado a 132 participantes, dos quais apenas 104 respostas válidas foram consideradas, conforme indicado no capítulo Introdução. Este conjunto de dados está organizado em dois grupos principais: variáveis categóricas, correspondentes às questões Q1 a Q14, e variáveis métricas discretas, correspondentes às questões Q15 a Q19. As variáveis categóricas descrevem as características demográficas dos inquiridos, enquanto as variáveis métricas discretas estão diretamente associadas aos construtos em análise, sendo operacionalizadas através de perguntas cuja resposta segue a escala de Likert com valores entre 1 (discorda totalmente) e 7 (concorda totalmente). Estas questões representam variáveis manifestas que sustentam os construtos definidos no estudo. [1]

A tabela seguinte apresenta a correspondência entre cada pergunta e o construto a que se refere, bem como as perguntas associadas às respetivas variáveis manifestas:

Tabela 2.1: Construtos, variáveis manifestas e respetivas questões. [1]

Construtos	Variáveis manifestas	Questão
ERP Complexity	erpuso2	Using the ERP system in my job is difficult to understand what is going on.
	erpuso3	Using the ERP system in my job involves much time doing mechanical operations.
	erpuso4	Using the ERP system in my job takes too long to learn how to use it.
Perceived Usefulness of Technology	erpresul2	Using the ERP system in my job increases my productivity.
	erpresul3	Using the ERP system in my job enhances my effectiveness.
	erpresul4	Using the ERP system in my job is very useful.

Continua na próxima página

Construtos	Variáveis manifestas	Questão
BMI Complexity	bmiarch1	During the last year, your organization has made changes in its business model that have not been implemented before by competitors.
	bmiarch2	During the last year, your organization has made changes in its business model that transform the way to interact with clients.
	bmiarch3	During the last year, your organization has made changes in its business model that modify the way to organize the relationships with clients.
BMI Cost	bmicost1	During the last year, your organization has made changes in its business model to introduce new ways to reduce fixed costs.
	bmicost2	During the last year, your organization has made changes in its business model to introduce new ways to reduce variable costs.
BMI Revenue	bmirevenue1	During the last year, your organization has made changes in its business model that introduced new ways to be profitable.
	bmirevenue2	During the last year, your organization has made changes in its business model that introduced new pricing mechanisms.

2.1 Imputação de Dados

O conjunto de dados disponibilizado pelos autores do artigo apresenta valores ausentes em algumas das respostas fornecidas pelos participantes às questões formuladas. De modo a replicar a metodologia do artigo original com o máximo de fidelidade possível, foram aplicados quatro métodos de modo a tornar o conjunto de dados o mais próximo do conjunto de dados pré-processado pelos autores.

Os métodos escolhidos para esta investigação foram: imputação por médias, por medianas, por *pair-wise* e pela remoção das linhas que tenham pelo menos um valor nulo. Para avaliar cada conjunto de dados, é realizada a CFA (descrito no Capítulo 3) e, por sua vez, é calculado o erro quadrático médio (MSE) e o erro absoluto médio (MAE). A tabela 2.2 apresenta os erros MSE e MAE para cada tipo de imputação de dados.

Tabela 2.2: Média dos erros entre a CFA a partir dos conjuntos de dados pré-processados e os cross divulgados pelos autores.

Conjunto de dados	MAE	MSE	RMSE
Remoção de nulos	0,0818	0,0199	0,1409
Imputação pela média	0,0801	0,0206	0,1436
Imputação pela mediana	0,0791	0,0203	0,1424
Imputação por <i>pair-wise</i>	0,0855	0,0231	0,1520

Como é possível observar na tabela 2.2, A imputação por mediana dos valores, obteve o valor mais baixo de MAE, quanto a remoção de registos com valores nulos obteve o menor MSE, seguido da imputação por mediana. Remover registos do conjunto de dados pode ser problemático, uma vez que reduz o número de graus de liberdade porque a dimensão dos dados reduz. Comparando os graus de liberdade do artigo com o número de observações (tabela 2.3) é possível observar que existem ligações não explícitas no modelo que foram omitidas pelos autores do artigo.

Tabela 2.3: Comparação do número de parâmetros entre o artigo original e o número obtido por CFA.

Conjunto de dados	Número de observações	Número de parâmetros (CFA)	Graus de liberdade artigo	Diferença de parâmetros
Remoção de nulos	97	36	59	2
Imputação pela média	104	36	59	9
Imputação pela mediana	104	36	59	9
Imputação por <i>pair-wise</i>	104	36	59	9

Como a latente *Perceived Usefulness of Technology* apresenta um dos *cross-loadings* com valor diferente do obtido através da CFA, será utilizado o conjunto de dados com remoção de nulos para replicar o modelo estudado pelos autores.

Capítulo 3

Identificação do modelo de medida

Este capítulo descreve a construção de um modelo de Análise Fatorial Confirmatória (CFA), desenvolvido com base na estrutura do Modelo de Equações Estruturais (SEM) apresentada no artigo de referência [1]. Apesar de os autores não terem proposto explicitamente um modelo CFA, a formulação aqui apresentada procura manter a coerência com a lógica teórica subjacente ao modelo original. A figura incluída no Apêndice A ilustra a estrutura resultante, a qual servirá de base para a análise do problema de investigação definido.

Na secção seguinte, apresentam-se as variáveis manifestas associadas a cada construto, acompanhadas dos respetivos pesos de explicação estimados pelo modelo. Por fim, discutem-se as principais conclusões que podem ser extraídas com base nos resultados obtidos.

3.1 Exploração do modelo

Para a conceção e análise do modelo CFA, é necessário especificar de que forma os construtos latentes explicam as variáveis manifestas identificadas pelos autores do artigo. Neste sentido, o Apêndice B apresenta as expressões matemáticas que formalizam a relação entre os construtos latentes e as respetivas variáveis manifestas.

Após a definição do modelo CFA, procedeu-se à análise individual de cada construto latente, com o objetivo de avaliar a significância das suas relações com as variáveis manifestas associadas. De seguida, foi realizada a análise fatorial confirmatória do modelo na sua totalidade, permitindo validar a estrutura proposta e aferir a sua adequação aos dados.

3.1.1 Distribuições *a priori* e hiperparâmetros

Com o objetivo de replicar os valores obtidos pelos autores do artigo, recorreu-se à biblioteca `blavaan` para a estimação por máxima verosimilhança. Esta abordagem permite reproduzir os resultados do estudo original numa perspetiva bayesiana. Adicionalmente, foram utilizadas distribuições *a priori* não informativas, de forma a garantir que a distribuição *a posteriori* se aproxime da distribuição de máxima verosimilhança, ou seja, assegurando que os dados não são enviesados por conhecimento prévio sobre os construtos em análise. A Tabela 4.1 apresenta as distribuições *a priori* predefinidas na biblioteca `blavaan`.

Tabela 3.1: Distribuições *a priori* não informativas utilizadas

variável	Distribuição <i>a priori</i>
λ	$\mathcal{N}(0, 10)$
μ	$\mathcal{N}(0, 32)$
θ	$Gamma(1, 0.5)$
Variâncias dos construtos (Ψ)	variável fixada
Covariâncias dos construtos	$LKJCorr(1)$

3.1.2 Estimação dos parâmetros do modelo

Para permitir o cálculo dos cross-loadings e da variância dos erros, as variâncias associadas a cada construto latente foram fixadas em 1. Esta parametrização baseia-se no modelo padronizado (*standard*), onde se assume uma média de 0 e um desvio-padrão de 1 para as variáveis latentes. A Tabela 3.2 apresenta os *cross-loadings*, as variâncias dos erros, os indicadores de ajustamento do modelo aos dados, bem como o parâmetro \hat{R} , que permite avaliar a convergência do modelo. Adicionalmente, foram calculadas as correlações residuais para as variáveis `erpresul2`, `erpresul3` e `erpresul4`, com o intuito de avaliar possíveis desvios não explicados pela estrutura latente.

Tabela 3.2: CFA de cada construto latente

Variável Observada/Covariância	λ	θ	\hat{R}	R^2
erpuso2	0,75	0,43	~ 1	0,57
erpuso3	0,52	0,73	~ 1	0,27
erpuso4	0,78	0,39	~ 1	0,62
erpresul2	0,90	0,20	~ 1	0,80
erpresul3	0,94	0,12	~ 1	0,88
erpresul4	0,88	0,22	~ 1	0,78
bmiarch1	0,59	0,65	~ 1	0,35
bmiarch2	0,94	0,11	~ 1	0,89
bmiarch3	0,85	0,28	~ 1	0,72
bmicost1	0,92	0,16	~ 1	0,84
bmicost2	0,92	0,15	~ 1	0,85
bmirevenue1	0,71	0,50	~ 1	0,50
bmirevenue2	0,69	0,53	~ 1	0,47
Convariância entre erpresul3 e erpresul2	-	0,24	~ 1	-
Convariância entre erpresul3 e erpresul4	-	0,17	~ 1	-

Os resultados apresentados nas tabelas anteriores correspondem às estimativas obtidas a partir do modelo gerado com recurso à biblioteca `stan`, utilizando as configurações especificadas na Tabela 3.3.

Tabela 3.3: Parâmetros utilizados na chamada da função `bcfa()`

Parâmetro	Valor
Padronização dos fatores latentes	<code>true</code>
Número de cadeias	3
Período de burn-in	5000 iterações
Amostras pós burn-in	1000
Método de amostragem	<code>stan</code>

Por fim, os parâmetros associados a todos os construtos latentes foram estimados numa perspetiva global, estando os respetivos resultados apresentados na Tabela 3.4.

Tabela 3.4: CFA com todos os construtos latentes

Variável Observada/Covariância	λ	θ	\hat{R}	R^2	<i>Loadings</i> artigo
erpuso2	0,75	0,44	~ 1	0,56	0,76
erpuso3	0,58	0,66	~ 1	0,34	0,54
erpuso4	0,75	0,44	~ 1	0,56	0,76
erpresul2	0,88	0,23	~ 1	0,77	0,90
erpresul3	0,95	0,10	~ 1	0,90	0,58
erpresul4	0,91	0,18	~ 1	0,82	0,88
bmiarch1	0,60	0,64	~ 1	0,36	0,57
bmiarch2	0,94	0,12	~ 1	0,88	0,94
bmiarch3	0,86	0,27	~ 1	0,73	0,86
bmicost1	0,96	0,09	~ 1	0,91	0,87
bmicost2	0,89	0,21	~ 1	0,89	0,94
bmirevenue1	0,82	0,33	~ 1	0,67	0,56
bmirevenue2	0,53	0,72	~ 1	0,28	0,72
Convariância entre erpresul3 e erpresul2	-	0,31	~ 1	-	-
Convariância entre erpresul3 e erpresul4	-	-0,01	~ 1	-	-

3.2 Discussão dos resultados

Ao comparar os resultados divulgados pelos autores com os obtidos na presente análise, é possível identificar algumas incongruências. Os *cross-loadings* dos construtos *BMI Cost* e *BMI Revenue* aparentam estar trocados, quando comparados com os valores obtidos através da máxima verosimilhança. Outra discrepância relevante refere-se ao peso do construto *Perceived Usefulness of Technology* na explicação da variável `erpresul3`, cujo valor estimado pela máxima verosimilhança é significativamente superior ao reportado pelos autores. Importa salientar que os *cross-loadings* foram calculados com base em 38 parâmetros, conforme descrito na Secção 2.1, e também com 36 parâmetros, considerando o modelo SEM original como base para a CFA. Em ambos os cenários, o peso

divergente permanece significativamente distinto dos dados divulgados. Adicionalmente, recorreu-se à biblioteca `lavaan` para calcular os índices de modificação, com o intuito de reduzir os graus de liberdade através da introdução de parâmetros adicionais. No entanto, após uma análise criteriosa, não foram identificados parâmetros plausíveis a incorporar, dado que estes correspondiam, maioritariamente, a correlações residuais entre variáveis manifestas associadas a diferentes construtos. Assim, conclui-se que não foi possível replicar integralmente os *cross-loadings* apresentados no artigo original.

Para além das divergências nos *cross-loadings*, verificou-se que duas variáveis manifestas presentes no modelo: `erpuso3`, que avalia a afirmação “*Using the ERP system in my job involves much time doing mechanical operations*”, e `bmiarch1`, que avalia “*During the last year, your organization has made changes in its business model that have not been implemented before by competitors*”, não foram adequadamente explicadas pelos construtos latentes propostos. Este resultado indica uma fraca associação entre estas variáveis e os fatores teóricos subjacentes, sugerindo que não representam de forma consistente os conceitos que os construtos pretendem medir. A inclusão destas variáveis no modelo contribui apenas para o aumento da sua complexidade, sem benefícios evidentes em termos de validade explicativa.

Capítulo 4

Identificação do modelo estrutural

Com base no diagrama do modelo SEM apresentado pelos autores, procedeu-se à estimação do respetivo modelo com o objetivo de o replicar. Apesar de não ter sido possível reproduzir integralmente os *cross-loadings* observados no modelo CFA analisado anteriormente, foi estimado o modelo SEM conforme a estrutura proposta no artigo original.

Tabela 4.1: Distribuições *a priori* não informativas utilizadas

variável	Distribuição <i>a priori</i>
λ	$\mathcal{N}(0, 10)$
μ	$\mathcal{N}(0, 32)$
θ	$\text{Gamma}(1, 0.5)$
Variâncias dos construtos (Ψ)	variável fixada
Covariâncias dos construtos	$LKJCorr(1)$

À semelhança do que foi feito na exploração do modelo de medida, a estimação do modelo estrutural foi realizada com recurso à função `bcfa`. Para assegurar a convergência do modelo, foi necessário ajustar alguns dos parâmetros de estimação. A Tabela ?? apresenta os valores utilizados nesta configuração.

Tabela 4.2: Parâmetros utilizados na chamada da função `bcfa()`

Parâmetro	Valor
Padronização dos fatores latentes	<code>true</code>
Número de cadeias	4
Período de burn-in	10 000 iterações
Amostras pós burn-in	2000
Método de amostragem	<code>stan</code>

A Tabela 4.3 apresenta uma comparação entre os valores publicados pelos autores e os *loadings* estimados com recurso ao método da máxima verosimilhança. A estrutura completa do modelo SEM divulgado encontra-se ilustrada no Apêndice C.

Tabela 4.3: *Cross-loadings* obtidos através da máxima verosimilhança

Regressão	Estimativa no artigo	Máxima verosimilhança
γ_{put}	-0,31	-0,32
γ_{bmi_comp}	-0,26	-0,26
$\gamma_{put_bmi_comp}$	0,13	0,10
γ_{bmi_cost}	0,24	0,07
$\gamma_{bmi_comp_rev}$	0,39	-0,52
$\gamma_{bmi_revenue}$	0,54	-0,17

4.1 Discussão dos resultados

A partir da análise da Tabela 4.3, observa-se que os *loadings* correspondentes aos parâmetros γ_{put} , γ_{bmi_comp} e $\gamma_{put_bmi_comp}$ refletem os valores estimados através do método da máxima verosimilhança. No entanto, os restantes parâmetros apresentam discrepâncias significativas face aos valores esperados, sugerindo que poderá ter ocorrido algum ajustamento no modelo original com o intuito de obter estimativas mais alinhadas com os objetivos teóricos dos autores.

4.2 Melhoramento do modelo

Com o objetivo de melhorar o ajuste de determinados *loadings* no modelo SEM, foram efetuadas modificações ao modelo original. Especificamente, foram removidas as variáveis manifestas `erpuso3` e `bmiarch1`, eliminada a regressão em que `bmi_complexity` explica `bmi_cost`, e reestruturada a equação em que `bmi_cost` passa a ser explicado por `bmi_revenue` e `bmi_complexity`. A representação gráfica do modelo refinado encontra-se no Apêndice D.

Ao comparar o modelo refinado com o modelo proposto pelos autores, verifica-se que o modelo ajustado obteve um valor de 3535 no *Watanabe-Akaike Information Criterion* (WAIC), em contraste com o valor de 4243 atribuído ao modelo original. Utilizando a biblioteca `blavaan` e a função `blavCompare`, foi possível confirmar que o modelo refinado apresenta menor complexidade.

4.3 Aplicação de *priors* informativas

Com o intuito de explorar mais aprofundadamente as potencialidades da estatística bayesiana, foram aplicadas duas distribuições *a priori*. O modelo utilizado para este estudo corresponde ao modelo refinado descrito na secção anterior. Assumiu-se que o *loading* analisado pelos autores segue uma distribuição normal com média 0 e desvio-padrão de 0.3. Adicionalmente, considerou-se que o *loading* no qual `bmi_cost` explica o construto `bmi_revenue` segue uma distribuição beta com os parâmetros $\alpha = 12$ e $\beta = 2$.

A primeira distribuição *a priori*, de natureza normal, concentra grande parte da sua massa entre -0.1 e 0.1, refletindo a convicção de que o valor do respetivo *loading* se encontra nesse intervalo. Já a segunda distribuição *a priori*, de natureza beta, sugere que

os valores esperados para o *loading* se concentram no intervalo $]0.7, 1[$, refletindo uma forte crença de que o efeito é elevado, mas sem atingir o valor máximo possível.

A Tabela 4.4 apresenta uma comparação entre os dados divulgados pelos autores, os *loadings* estimados através do método da máxima verosimilhança com distribuições *a priori* não informativas, e os *loadings* obtidos utilizando distribuições *a priori* informativas.

Tabela 4.4: *Cross-loadings* obtidos através da máxima verosimilhança com e sem priors informativas

Regressão	Estimativa no artigo	<i>Loadings</i> sem <i>priors</i> informativas	<i>Loadings</i> com <i>priors</i> informativas
γ_{put}	-0,31	-0,32	-0,19
γ_{bmi_comp}	-0,26	-0,26	-0,26
$\gamma_{put_bmi_comp}$	0,13	0,10	0,14
γ_{bmi_cost}	0,24	0,07	-
$\gamma_{bmi_comp_rev}$	0,39	-0,52	0,55
$\gamma_{bmi_revenue}$	0,54	-0,17	-0,18

A distribuição *a priori* aplicada à regressão $\gamma_{put_bmi_comp}$ não produziu alterações significativas nos *loadings*, o que permite concluir que esta *prior* é consistente com os dados e representa adequadamente o parâmetro $\gamma_{put_bmi_comp}$.

Por outro lado, a distribuição *a priori* informativa atribuída ao parâmetro $\gamma_{bmi_revenue}$ revelou-se inadequada, uma vez que os valores obtidos divergem consideravelmente daqueles estimados pela máxima verosimilhança. Tal discrepância sugere que esta *prior* não reflete adequadamente a realidade subjacente aos dados observados.

Capítulo 5

Conclusão

Em síntese, a presente investigação procurou replicar o estudo original recorrendo à modelação bayesiana, não apenas com o intuito de reprodução, mas também para explorar as potencialidades desta abordagem no contexto da modelação de equações estruturais. A replicação revelou-se desafiante, sobretudo devido a discrepâncias nos *cross-loadings* e à fraca representação de algumas variáveis manifestas, o que impediu uma reprodução fiel dos resultados originais.

Ainda assim, o trabalho proporcionou uma análise crítica da estrutura do modelo original, evidenciando possíveis ajustamentos feitos pelos autores para alcançar os resultados desejados. A aplicação de diferentes métodos de imputação foi necessária para tentar replicar com maior fidelidade as decisões pré-analíticas tomadas pelos autores.

A construção de um modelo estrutural refinado, com menor complexidade e melhor ajustamento, demonstrou o potencial da modelação bayesiana para melhorar a parcimónia sem comprometer a explicação dos fenómenos em estudo. A introdução de *priors* informativas revelou-se promissora, embora os resultados mistos alertem para a necessidade de uma criteriosa adequação das *priors* ao contexto empírico.

Em conclusão, embora a replicação integral do modelo original não tenha sido atin- gida, este trabalho contribuiu para a compreensão dos desafios e oportunidades da mo- delação bayesiana em contextos empíricos complexos, sublinhando a importância de uma abordagem crítica e iterativa na construção e validação de modelos SEM.

Bibliografia

- [1] R. Rodríguez, G. Svensson, and F.-J. Molina-Castillo, “Dataset that shows organizational complexity as a key enabler for business model innovation using enterprise resource planning,” *Data in Brief*, vol. 55, p. 110754, 2024. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2352340924007200>

Apêndice A

Modelo de Medida

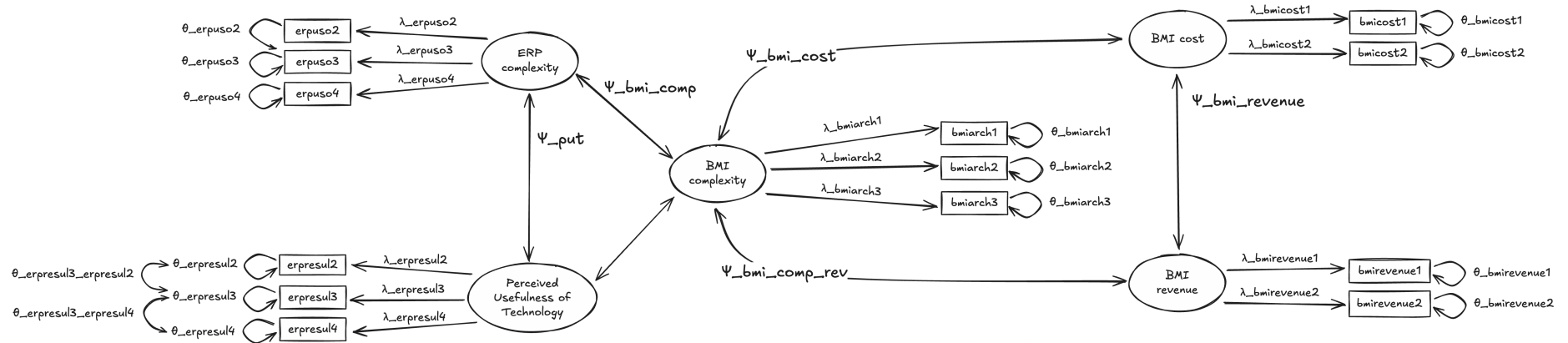


Figura 1: Modelo de medida

Apêndice B

Definição do Modelo de Medida

B.1 Variáveis Manifestas

Construto *ERP Complexity*

$$erpuso2 = \mu_{erpuso2} + \lambda_{erpuso2} \times f(ERP\ Complexity) + \varepsilon_{erpuso2} \quad (B.1)$$

$$erpuso3 = \mu_{erpuso3} + \lambda_{erpuso3} \times f(ERP\ Complexity) + \varepsilon_{erpuso3} \quad (B.2)$$

$$erpuso4 = \mu_{erpuso4} + \lambda_{erpuso4} \times f(ERP\ Complexity) + \varepsilon_{erpuso4} \quad (B.3)$$

Construto *Perceived Usefulness of Technology*

$$erpresul2 = \mu_{erpresul2} + \lambda_{erpresul2} \times f(Perceived\ Usefulness\ of\ Technology) + \varepsilon_{erpresul2} \quad (B.4)$$

$$erpresul3 = \mu_{erpresul3} + \lambda_{erpresul3} \times f(Perceived\ Usefulness\ of\ Technology) + \varepsilon_{erpresul3} \quad (B.5)$$

$$erpresul4 = \mu_{erpresul4} + \lambda_{erpresul4} \times f(Perceived\ Usefulness\ of\ Technology) + \varepsilon_{erpresul4} \quad (B.6)$$

Construto *BMI complexity*

$$bmiarch1 = \mu_{bmiarch1} + \lambda_{bmiarch1} \times f(BMI\ complexity) + \varepsilon_{bmiarch1} \quad (B.7)$$

$$bmiarch2 = \mu_{bmiarch2} + \lambda_{bmiarch2} \times f(BMI\ complexity) + \varepsilon_{bmiarch2} \quad (B.8)$$

$$bmiarch3 = \mu_{bmiarch3} + \lambda_{bmiarch3} \times f(BMI\ complexity) + \varepsilon_{bmiarch3} \quad (B.9)$$

Construto *BMI cost*

$$bmicos1 = \mu_{bmicos1} + \lambda_{bmicos1} \times f(BMI\ cost) + \varepsilon_{bmicos1} \quad (B.10)$$

$$bmicos2 = \mu_{bmicos2} + \lambda_{bmicos2} \times f(BMI\ cost) + \varepsilon_{bmicos2} \quad (B.11)$$

Construto *BMI revenue*

$$bmirevenue1 = \mu_{bmirevenue1} + \lambda_{bmirevenue1} \times f(BMI\ revenue) + \varepsilon_{bmirevenue1} \quad (B.12)$$

$$bmirevenue2 = \mu_{bmirevenue2} + \lambda_{bmirevenue2} \times f(BMI\ revenue) + \varepsilon_{bmirevenue2} \quad (B.13)$$

B.2 Distribuição dos Construtos Latentes

Através do modelo estimado com recurso à biblioteca **blavaan**, e através da máxima verosimilhança, os vários construtos seguem a seguinte distribuição:

$$\begin{bmatrix} \text{ERP Complexity} \\ \text{PUT} \\ \text{BMI Complexity} \\ \text{BMI Cost} \\ \text{BMI Revenue} \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} ; \begin{bmatrix} 1 & -0.285 & -0.279 & -0.184 & -0.084 \\ -0.285 & 1 & 0.192 & -0.076 & 0.064 \\ -0.279 & 0.192 & 1 & 0.016 & 0.486 \\ -0.184 & -0.076 & 0.016 & 1 & -0.161 \\ -0.084 & 0.064 & 0.486 & -0.161 & 1 \end{bmatrix} \right),$$

onde PUT significa *Perceived Usefulness of Technology*.

Apêndice C

Modelo Estrutural

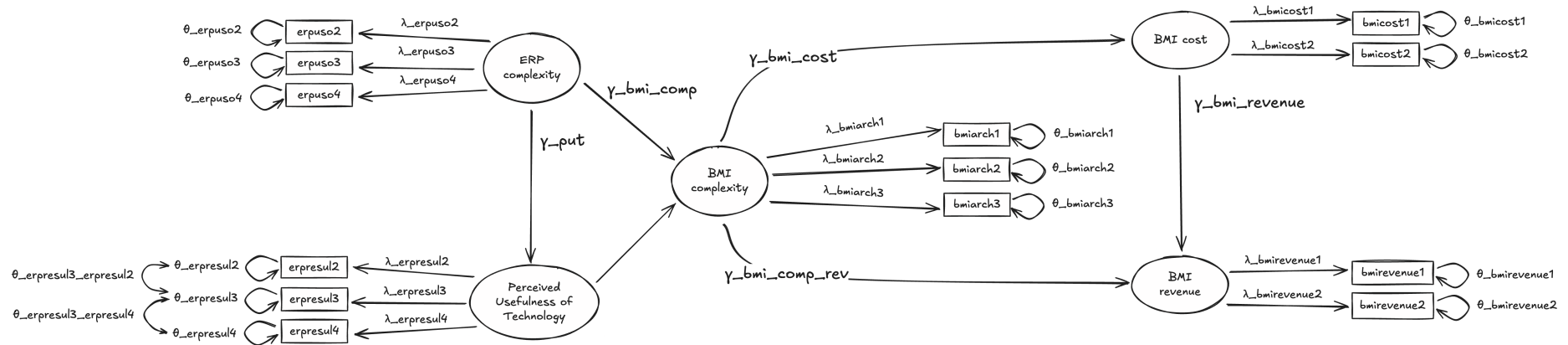


Figura 1: Modelo Estrutural

Apêndice D

Modelo Estrutural Refinado

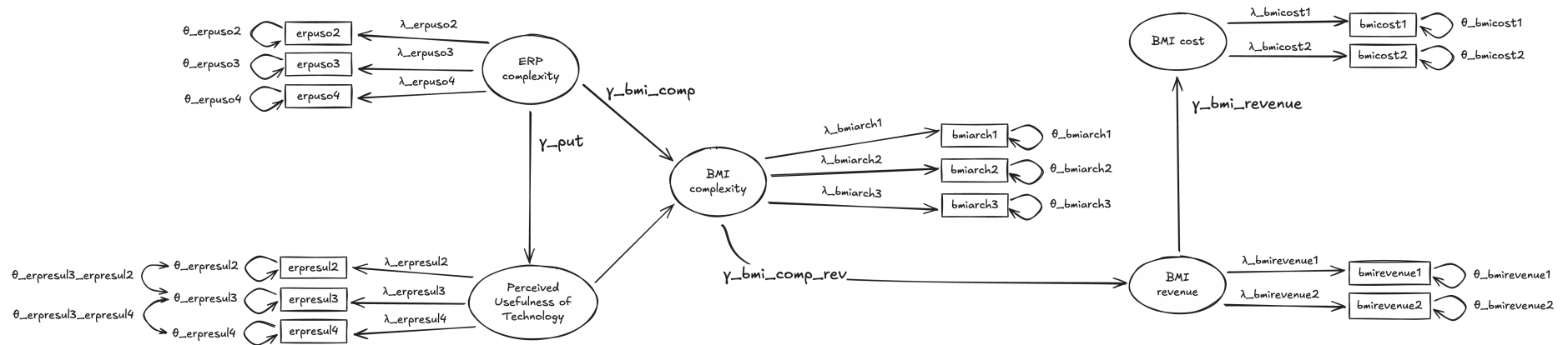


Figura 1: Modelo Estrutural Refinado