UNIVERSIDADE FEDERAL DE MINAS GERAIS - UFMG
INSTITUTO DE CIÊNCIAS EXATAS - ICEx
DEPARTAMENTO DE ESTATÍSTICA

MODELOS DE REGRESSÃO t-TOBIT COM ERROS NAS COVARIÁVEIS

Gustavo Henrique Mitraud Assis Rocha
Orientadora: Rosangela Helena Loschi
Co-orientador: Reinaldo Boris Arellano-Valle

Tese de doutorado

MODELOS DE REGRESSÃO t-TOBIT COM ERROS NAS COVARIÁVEIS

Gustavo Henrique Mitraud Assis Rocha

2014

Dedico este trabalho
aos meus pais, Ozeres (+) e Mary,
e à minha irmã Rita e sobrinha Paloma.

Agradecimentos

Agradeço a Deus por tudo que tem me proporcionado e por mais uma conquista.

Aos meus pais, Ozeres (*in memorian*) e Mary, à minha irmã Rita e à minha sobrinha Paloma pelo apoio em relação aos estudos.

À minha família por torcerem por mim e desejarem muito sucesso em minha vida acadêmica e profissional e, em especial, à minha tia Miriam por toda a ajuda prestada ao longo da vida.

À minha orientadora, Professora Rosangela Helena Loschi, pelo apoio, paciência, auxílio e até mesmo amizade durante os anos de UFMG. Ao meu co-orientador Professor Reinaldo Boris Arellano-Valle (PUC-Chile), pela confiança e também pela oportunidade que me foi dada de conhecer um país tão rico quanto o Chile.

Aos professores do Departamento de Estatística da UFMG, pelos ensinamentos concedidos.

Aos membros da banca examinadora, Prof. Marcos Oliveira Prates (UFMG), Profª. Lourdes Coral Contreras Montenegro (UFMG), Prof. Manuel Jesus Galea Rojas (PUC - Chile) e Prof. Filidor Edilfonso Vilca Labra (UNICAMP) pela leitura, correções e sugestões para a tese.

Aos colegas de trabalho da ENCE - Escola Nacional de Ciências Estatísticas - por toda a ajuda e oportunidade para o desenvolvimento do trabalho. Em especial à Aline, Daniela, Larissa, Maria Luiza e Renata, pelos bons momentos de descontração e adaptação.

Agradeço a todos os alunos que já tive pela troca de conhecimento.

À Fundação João Pinheiro e à Faculdade Pitágoras pela experiência.

À CAPES pela bolsa de doutorado, à FAPEMIG por diversos apoios financeiros prestados para participação em eventos e ao PIBIC-CNPq pela bolsa de iniciação científica, importante para o surgimento do interesse em pesquisas.

A todos os amigos e amigas cujas amizades foram feitas na UFMG - graduação, mestrado e doutorado - e que se mantem até os dias de hoje.

Aos amigos que tornam a vida bem mais divertida.

Em especial aos amigos Bruno Freitas, Fabrício Teixeira, Lucas Rodrigues, Watson Hermann, Rômulo Nadler, Rafael Valentim e Tiago Santos, que tanto me apoiaram e me entreteram ao longo dos anos.

Ao Lucas Cavalcante, que tanto me auxilia e se faz presente.

A todos, o meu muito obrigado!

Gustavo Henrique Mitraud Assis Rocha

Resumo

Neste trabalho é desenvolvida uma análise de regressão linear considerando que a variável dependente é censurada e também que algumas das variáveis explicativas são medidas com erros aditivos. Esse modelo de regressão censurado com erros de medidas é especificado assumindo distribuições com cauda pesada para o processo probabilístico. Especificamente, assume-se uma distribuição t-Student multivariada para modelar o comportamento conjunto dos erros e das verdadeiras covariáveis não observadas. Nesse sentido, o modelo será robusto o suficiente para proteger as inferências de observações atípicas e influentes. Para a estimação do modelo considera-se a metodologia de máxima verossimilhança, em que inclui-se a estimação da variância assintótica dos estimadores de máxima verossimilhança e também desenvolve-se um algoritmo do tipo EM para obter as estimativas, e também o paradigma bayesiano, onde considera-se o procedimento de aumento de dados e desenvolve-se um algoritmo MCMC para amostrar das distribuições *a posteriori*. A metodologia proposta é flexível o bastante para ser adaptada para distribuições com caudas pesadas vindas da classe de misturas de escala da distribuição normal. A performance da nova metodologia desenvolvida é avaliada através de um estudo Monte Carlo e também de uma análise de um estudo de caso.

Abstract

In this work, we develop a non-standard linear regression analysis by considering that the dependent variable is censored and also that some of the explanatory variables are measured with additive errors. In addition, our censored measurement error regression model is specified by assuming heavy-tailed distributions for the underlying probabilistic process. Specifically, our analysis focuses on assuming a multivariate Student-t joint distribution for the error terms and the unobserved true covariates. In this sense, the proposed model will be robust enough to protect our inferences of atypical or influential observations. For the model estimation, we consider the maximum likelihood methodology, in which we include the estimation of the asymptotic variance of the maximum likelihood estimators and we also develop an EM type algorithm to obtain the estimates, and also the Bayesian paradigm, in which we use a data augmentation approach and develop a MCMC algorithm to sample from the posterior distributions. The proposed methodology is flexible enough to be adapted for heavy-tailed distributions coming from the class of scale mixture of the normal distribution. The performance of the newly developed methodology is evaluated throughout a Monte Carlo study as well as a case sudy analysis.

Sumário

Capítulo 1

Introdução

Os modelos de regressão censurados ou modelos Tobit são bastante usados na análise estatística para modelar variáveis respostas que são parcialmente observadas ou que possuem uma quantidade de valores agrupados em um valor limite. O campo de aplicação desses modelos cobre várias áreas tais como econometria, biometria, ensaios clínicos dentre outros. No trabalho pioneiro de Tobin (Tobin, 1958), o modelo Tobit assume que o valor limite é zero então apenas valores positivos da variável dependente são efetivamente observados. Inferências nos modelos Tobit, no entanto, levam em consideração todos os dados, tanto as respostas efetivamente observadas, quanto as censuradas. Como em Tobin (1958), a maior parte da teoria desenvolvida ao redor do modelo Tobit é baseada na suposição de normalidade para a distribuição dos erros do modelo. A teoria de máxima verossimilhança e outros métodos de estimação para esses modelos podem ser encontrados em Amemiya (1984, 1985), Maddala (1983), Greene (1990), Olsen (1978) e Barros *et al.* (2010), dentre outros. Métodos bayesianos são também considerados por Carriquiry *et al.* (1987), Sweeting (1987), Cowles *et al.* (1996), Hamilton (1999) e, mais recentemente, por Austin (2002). Chib (1992) estabeleceu condições para a existência dos momentos *a posteriori* no modelo Tobit e introduziu uma estratégia de aumento de dados para amostras da distribuição *a posteriori* na presença de dados censurados. Tal estratégia facilita as aproximações numéricas para tais momentos *a posteriori*.

Em tais trabalhos, também é assumido que as variáveis explicativas estão livres de erros de medida. Tal suposição não é realística em vários problemas práticos e podem levar à estimativas inconsistentes e imprecisas para os parâmetros (Fuller, 1987). Dellaportas e Stephens (1995) analisam modelos com erros nas covariáveis sob o ponto de

vista bayesiano utilizando MCMC e Bolfarine e Arellano-Valle (2005) discutem a inferência bayesiana em modelos elípiticos dependentes e independentes com erros de medida. Algumas extensões para o modelo Tobit tem sido propostas nos últimos anos. Polasek e Krause (1993) e Wang (1998), por exemplo, consideram um modelo Tobit com erro nas covariáveis com disribuições normais independentes para os erros e as covariáveis latentes. Esse modelo é chamado aqui de modelo N-Tobit com erro de medida ou com erro nas covariáveis. Seguindo Chib (1992), Polasek e Krause (1993) introduziram a estratégia de aumento de dados para obter estimativas *a posteriori* para os parâmetros no modelo N-Tobit com erro de medida. Em outra direção, Wang (1998) derivou estimadores do método dos momentos e de máxima verossimilhança consistentes para tal modelo (veja também Wang (2002) e Wang and Hsiao (2007)). Em Wang (1998), as estimativas de máxima verossimilhança foram obtidas usando o método de Newton-Raphson. Wang (1998) também discutiu a identificabilidade do modelo N-Tobit com erro nas covariáveis e mostrou que, sob normalidade, o modelo é unicamente reduzido a um modelo livre de erros que faz com que a inferência no modelo original seja realizada mais facilmente. Em contrapartida, o modelo Tobit não é robusto o suficiente para acomodar observações atípicas. Um modelo robusto para variáveis dependentes censuradas é introduzido em Arellano-Valle *et al.* (2012), que estende o modelo proposto por Tobin (1958) assumindo uma distribuição t-Student para o erro do modelo, aqui chamado de modelo t-Tobit. O estimador de máxima verossimilhança para o modelo t-Tobit não possui forma fechada. Deste modo, Arellano-Valle *et al.* (2012) também desenvolveram um algoritmo do tipo EM para obter as estimativas de máxima verossimilhança do modelo t-Tobit. Mais ainda, tais autores obtiveram a distribuição normal assintótica conjunta para os estimadores de máxima verossimilhança do modelo t-Tobit. Massuia *et al.* (2014) também trabalham com um modelo de regressão linear censurado assumindo a distribuição t-Student para os erros e fazem uma análise de diagnóstico. Outras generalizações para o modelo Tobit foram introduzidas por Lachos *et al.* (2011) e Matos *et al.* (2013) para modelar variáveis dependentes repetidas ou longitudinais que também são censuradas. Lachos *et al.* (2011) e Matos *et al.* (2013) consideram um modelo robusto tipo Tobit com efeitos mistos assumindo uma distribuição t multivariada para modelar conjuntamente os erros e o comportamento dos efeitos aleatórios.

Neste trabalho, um modelo tipo Tobit com caudas pesadas que inclui covariáveis medidas com erros é introduzido. Como em Polasek e Krause (1993) e Wang (1998), considera-se uma estrutura aditiva para os erros de medida. Especificamente, propõe-se uma extensão robusta para o modelo N-Tobit com erro de medida modelando conjuntamente o comportamento aleatório do modelo, os erros de medida e as variáveis latentes com uma distribuição t multivariada. Como em Wang (1998), mostra-se que, sob certas condições de identificabilidade, o modelo proposto pode ser unicamente reduzido para um modelo de regressão censurado condicional livre de erros com erros aleatórios heterocedásticos e covariáveis aleatórias observadas, que são não correlacionadas. Também será mostrado que o modelo proposto pode ser representado hierarquicamente em termos de um modelo N-Tobit com erros de medida, onde a incerteza sobre a heterocesdasticidade é modelada por meio de uma distribuição gama. Considerando-se tal representação hierárquica, desenvolve-se um algoritmo do tipo EM para obter-se os estimadores de máxima verossimilhança. Seus erros padrão serão obtidos utilizando a matriz de informação observada. Uma estratégia de aumento de dados também é introduzida para obter-se amostras das distribuições *a posteriori*. Um estudo Monte Carlo é realizado para avaliar o comportamento dos estimadores de máxima verossimilhança obtidos para o modelo proposto. Também avalia-se os estimadores bayesianos média e mediana *a posteriori* obtidos usando o modelo proposto e alguns outros modelos tipo Tobit previamente introduzidos na literatura, que são o Tobit, t-Tobit e N-Tobit com erros nas covariáveis. Como subproduto, também obtém-se alguns resultados relacionados à inferência bayesiana no modelo t-Tobit introduzido por Arellano-Valle *et al.* (2012). Os cenários considerados incluem diferentes proporções de censura, tamanhos amostrais, valores para a restrição de identificabilidade e também diferentes graus de liberdade. O modelo proposto também é ajustado para analisar os valores de gastos ambulatoriais disponíveis no conjunto de dados de *2001 Medical Expenditure Panel Survey*. Esse conjunto de dados foi previamente analisado por Marchenko e Genton (2012) utilizando o modelo de Heckman t-Student. Trabalhos anteriores tem mostrado que usualmente a covariável renda não é medida com precisão. Portanto, diferente do que foi considerado por Marchenko e Genton (2012), será assumido que que a renda individual é uma variável explicativa medida com erro.

Este trabalho está assim organizado. O Capítulo 2 apresenta formalmente o modelo

Tobit livre de erros e discute-se aspectos de inferência bayesiana no modelo t-Tobit. Com isto, é apresentada uma extensão de resultados obtidos em Arellano-Valle *et al.* (2012). No Capítulo 3 os modelos tipo Tobit com erros nas covariáveis são apresentados e é introduzido o modelo t-Tobit com erros de medida. Usando a estratégia de aumento de dados (van Dyk e Meng, 2001, por exemplo) propõe-se uma estratégia para se amostrar da distribuição *a posteriori* no modelo proposto. O método de máxima verossimilhança é estudado no Capítulo 4, onde é construído um algoritmo do tipo EM para aproximá-lo. A matriz de informação de Fisher observada é obtida, o que permite obter-se uma aproximação para intervalos de confiança dos parâmetros de interesse. No Capítulo 5 são feitos estudos Monte Carlo para avaliar o comportamento dos estimadores bayesianos e de máxima verossimilhança sob o modelo proposto em diferentes cenários. Entre outras coisas, o efeito nas estimativas é investigado quando amostras são geradas com diferentes proporções de censura, graus de liberdade e tamanhos de amostra. No caso bayesiano, também tem-se como meta comparar as estimativas obtidas usando o modelo proposto com aquelas obtidas utilizando os modelos Tobit, t-Tobit e N-Tobit com erros nas covariáveis. Um estudo de caso é feito no Capítulo 6, onde dados de gastos ambulatoriais são analisados e ajusta-se a eles o modelo proposto. Por fim, no Capítulo 7 são apresentados comentários finais sobre o trabalho aqui desenvolvido com sugestões de propostas futuras.

Ao longo deste trabalho, denota-se, respectivamente, por $\phi_k(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$ e $\Phi_k(\cdot \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$ a função de densidade de probabilidade (fdp) e a função de distribuição acumulada (fda) de uma distribuição $N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ (distribuição normal k-variada com vetor de médias $\boldsymbol{\mu}$ e matriz de covariância $\boldsymbol{\Sigma}$). Por simplicidade, quando $\boldsymbol{\mu} = \mathbf{0}$ e $\boldsymbol{\Sigma} = \boldsymbol{I}_k$, onde $\mathbf{0}$ e $\boldsymbol{I}_k$ denotam um vetor de zeros e a matriz identidade de dimensão k, respectivamente, a fdp e a fda serão denotadas, respectivamente, por $\phi_k(\mathbf{x})$ e $\Phi_k(\cdot)$. Também será assumido que, se $\mathbf{x} \sim t_k(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$, isto é, $\mathbf{x}$ tem distribuição t k-variada com parâmetro de localização $\boldsymbol{\mu}$, matriz de dispersão $\boldsymbol{\Sigma}$ e graus de liberdade ν, então sua fdp será denotada por $t_k(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$, isto é, $t_k(\mathbf{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu) = |\boldsymbol{\Sigma}|^{-1/2} C_{k,\nu} \left(\nu + q(\mathbf{x})\right)^{-(\nu+k)/2}$, $\mathbf{x} \in \mathbb{R}^k$, onde $C_{k,\nu} = \Gamma\left((\nu+k)/2\right) / \Gamma\left(\nu/2\right) \pi^{-k/2} \nu^{\nu/2}$ e $q(\mathbf{x}) = (\mathbf{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})$, e considere ainda $m_{(\mathbf{x},k,\nu)} = \frac{\nu+k}{\nu+q(\mathbf{x})}$ e $h_{(\nu,k)} = \frac{1}{\nu+k}$. Como no caso normal, $T_k(\cdot \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$ denota a fda associada à distribuição $t_k(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$ e $t_k(\mathbf{x} \mid \nu)$ e $T_k(\cdot \mid \nu)$ denotam, respectivamente, a fdp e a fda de uma distribuição t-Student com grau de liberdade ν e quando $\boldsymbol{\mu} = \mathbf{0}$ e $\boldsymbol{\Sigma} = \boldsymbol{I}_k$.

Capítulo 2

Modelos tipo Tobit

A especificação de um modelo linear censurado inicia-se com uma relação linear dada por

$$\eta_i = \beta_1 + \boldsymbol{\beta}_2' \mathbf{x}_i + u_i, \tag{2.1}$$

$i = 1, \ldots, n$, onde $\eta_i \in \mathbb{R}$ é a variável resposta e $\mathbf{x}_i \in \mathbb{R}^k$ é o vetor de variáveis explicativas medido sem erro para o i-ésimo indivíduo, $(\beta_1, \boldsymbol{\beta}_2') \in \mathbb{R}^{1+k}$ são os parâmetros de regressão e $u_i \in \mathbb{R}$ é um erro aleatório. Um modelo de regressão censurado tipo Tobit pressupõe que a variável dependente η_i não é completamente observada, isto é, observa-se $y_i = \eta_i$ quando $\eta_i > 0$ e $y_i = 0$ caso contrário, o que é equivalente a escrever

$$y_i = \max\{\eta_i, 0\}, \tag{2.2}$$

$i = 1, \ldots, n$. As equações (2.1) e (2.2) definem o usual (livre de erros) modelo de regressão linear censurado. Esse tipo de modelo foi aplicado para analisar um problema econômico por Tobin (1958).

O modelo Tobit é então determinado pela distribuição das variáveis aleatórias latentes u_i, $i = 1, \ldots, n$. Há muitas opções para especificar-se tal distribuição. Devido a estrutura linear da relação (2.1), pode-se considerar distribuições que mantem a distribuição sob transformações lineares. Nesse sentido, distribuições elípticas são opções naturais. Essa classe contem muitas distribuições simétricas que podem comportar caudas leves ou pesadas. A família de distribuições normais independentes ou mistura na escala de distribuições normais é um exemplo. Diz-se que u_i pertence à família de distribuições normais independentes se

$$u_i \stackrel{d}{=} g_i^{-1/2} Z_i, \tag{2.3}$$

onde $\stackrel{d}{=}$ denota igualdade em distribuição, $Z_i \sim N(0, \sigma_u)$ e g_i é uma variável aleatória de mistura, que é positiva, independente de Z_i e possui um parâmetro ν que indexa sua distribuição, para todo $i = 1, \ldots, n$. Dado g_i, u_i possui uma distribuição normal com média 0 e variância $g_i^{-1}\sigma_u$, $i = 1, \ldots, n$. A distribuição marginal de u_i depende de qual distribuição se assume para a variável g_i.

Quando a distribuição de g_i é degenerada atribuindo massa em 1, isto é, se $P(g_i = 1) = 1$ para todo $i = 1, \ldots, n$, então u_i possui distribuição normal com média 0 e variância σ_u. Consequentemente, o modelo obtido é o modelo originalmente apresentado por Tobin (1958) e é denominado na literatura por modelo Tobit. No entanto, sabe-se que a distribuição normal não é robusta o suficiente para modelar o comportamento de variáveis aleatórias com caudas pesadas. Ao assumir que g_i possui uma distribuição gama com parâmetros de forma e escala iguais a $\nu/2$ e $2/\nu$, respectivamente, então u_i possui uma distribuição t-Student com parâmetro de localização 0, parâmetro de escala σ_u e graus de liberdade $\nu > 0$. Como consequência, obtem-se em (2.1) uma distribuição t-Student para η_i. Esse foi o modelo introduzido por Arellano-Valle *et al.* (2012) e, neste trabalho, será chamado de modelo t-Tobit. Outras distribuições podem ser atribuídas para g_i. Neste trabalho, o foco é apenas nos casos normal e t, apresentando resultados relacionados à estatística bayesiana no modelo t-Tobit. Com isto, estende-se alguns resultados apresentados em Arellano-Valle *et al.* (2012).

2.1 O modelo Tobit

O modelo Tobit, introduzido por Tobin (1958), assume em (2.3) que $P(g_i = 1) = 1$ para todo $i = 1, \ldots, n$, o que equivale afirmar que

$$u_i \sim N(0, \sigma_u). \tag{2.4}$$

O modelo Tobit pode ser representado como segue

$$\left.\begin{array}{rcl} y_i & = & \max\{\eta_i, 0\}, \\ \eta_i & \sim & N_1(\beta_1 + \boldsymbol{\beta}_2'\mathbf{x}_i, \sigma_u), \end{array}\right\} \tag{2.5}$$

para todo $i = 1, \ldots, n$.

Denote por $\boldsymbol{\theta} = (\beta_1, \boldsymbol{\beta}_2, \sigma_u)$ o vetor de parâmetros do modelo Tobit e por $\mathbf{D}_o = (\mathbf{D}_{o1}, \ldots, \mathbf{D}_{on})$ os dados observados para todos os n indivíduos, onde $\mathbf{D}_{oi} = (y_i, d_i, \mathbf{x}_i)$ são os dados observados para o i-ésimo indivíduo e $d_i = \mathbf{1}(y_i > 0)$ é o indicador de censura para o i-ésimo indivíduo sendo igual a 1 se sua variável resposta é não censurada. Logo, a função de log-verossimilhança para uma amostra de n observações independentes do modelo Tobit é dada por

$$\begin{aligned} \log f(\mathbf{D}_o \mid \boldsymbol{\theta}) &= \sum_{i=1}^{n} d_i \log \phi_1 \left(y_i \mid \beta_1 + \boldsymbol{\beta}_2' \mathbf{x}_i, \sigma_u\right) \\ &\quad + \sum_{i=1}^{n} (1 - d_i) \log \Phi_1 \left(0 \mid \beta_1 + \boldsymbol{\beta}_2' \mathbf{x}_i, \sigma_u\right). \end{aligned} \tag{2.6}$$

A inferência no modelo Tobit com enfoque frequentista e bayesiano já foi considerada em diversos trabalhos como, por exemplo, Olsen (1978), Maddala (1983), Amemiya (1984, 1985), Carriquiry *et al.* (1987), Sweeting (1987), Greene (1990), Chib (1992), Cowles *et al.* (1996), Hamilton (1999), Austin (2002), dentre outros.

2.2 Função de verossimilhança do modelo t-Tobit

Considere o modelo t-Tobit, isto é, assuma que em (2.3), $g_i \sim G(\nu/2, 2/\nu)$ para todo $i = 1, \ldots, n$, onde $G(a, b)$ denota uma distribuição gama com média e variância dadas, respectivamente, por ab e ab^2. Consequentemente, tem-se que

$$u_i \sim t_1 \left(0, \sigma_u, \nu\right), \tag{2.7}$$

onde $\sigma_u > 0$ é o parâmetro de escala e $\nu > 0$ denota os graus de liberdade e, sendo assim, as equações (2.1), (2.2) e (2.7) definem o modelo t-Tobit.

Seja $\boldsymbol{\theta} = (\beta_1, \boldsymbol{\beta}_2, \sigma_u, \nu)$ e denote por $\mathbf{D}_o = (\mathbf{D}_{o1}, \ldots, \mathbf{D}_{on})$ os dados observados para todos os n indivíduos, onde $\mathbf{D}_{oi} = (y_i, d_i, \mathbf{x}_i)$ são os dados observados para o i-ésimo indivíduo e $d_i = \mathbf{1}(y_i > 0)$ é o indicador de censura para o i-ésimo indivíduo sendo igual a 1 se sua variável resposta é não censurada. Logo, a função de log-verossimilhança para uma amostra de n observações independentes é dada por

$$\log f(\mathbf{D}_o \mid \boldsymbol{\theta}) = \sum_{i=1}^{n} d_i \log t_1 \left(y_i \mid \beta_1 + \boldsymbol{\beta}_2' \mathbf{x}_i, \sigma_u, \nu\right)$$

$$+\sum_{i=1}^{n}(1-d_i)\log T_1\left(0 \mid \beta_1+\boldsymbol{\beta}_2'\mathbf{x}_i,\sigma_u,\nu\right). \tag{2.8}$$

A inferência no modelo t-Tobit via estimador de máxima verossimilhança foi considerada por Arellano-Valle *et al.* (2012), onde foi desenvolvido um algoritmo do tipo EM para obter-se uma aproximação do estimador $\widehat{\boldsymbol{\theta}}$ de $\boldsymbol{\theta}$. Tais autores obtiveram a distribuição normal assintótica de $\widehat{\boldsymbol{\theta}}$. Neste trabalho apresenta-se o enfoque bayesiano para realizar inferência no modelo t-Tobit.

2.2.1 Inferência bayesiana no modelo t-Tobit

Uma estratégia que facilita a inferência bayesiana sob modelos complexos é a técnica de aumento de dados. Tal estratégia foi considerada em modelos tipo Tobit por Chib (1992) e Cowles *et al.* (1996). Tal técnica consiste em incluir variáveis latentes ou dados não observados no modelo para simplificar os procedimentos computacionais. Para uma explicação mais detalhada sobre as técnicas de aumento de dados veja van Dyk e Meng (2001) e referências mencionadas nele.

Nesta seção será desenvolvido um algoritmo para amostrar da distribuição *a posteriori* para $\boldsymbol{\theta}$ no modelo t-Tobit baseado nas técnicas MCMC e aumento de dados. Para isso, considera-se a representação estocástica da distribuição t-Student em termos de uma mistura de escala da distribuição normal onde a medida misturadora é a distribuição gama $G(\nu/2, 2/\nu)$, como foi feito em (2.3). Consequentemente, tem-se

$$\eta_i \stackrel{d}{=} \beta_1+\boldsymbol{\beta}_2'\mathbf{x}_i+g_i^{-1/2}Z_i, \tag{2.9}$$

onde g_i e Z_i são variáveis aleatórias independentes tais que

$$\begin{aligned} g_i &\sim G(\nu/2,2/\nu),\\ Z_i &\sim N_1\left(0,\sigma_u\right), \end{aligned}$$

$i=1,\ldots,n$.

As variáveis $g_1,\ldots,g_n$ são variáveis latentes. Quando $\nu\to\infty$, para cada i, $g_i\to 1$ com probabilidade 1, isto é, η_i tem uma distribuição normal no caso limite, ou seja, se $\nu=\infty$. Consequentemente, sob essa suposição, o algoritmo aqui introduzido com as devidas modificações também pode ser usado para amostrar da distribuição *a posteriori* no modelo Tobit.

Nos modelos de regressão do tipo Tobit as respostas originais $\eta_1, \ldots, \eta_n$ também são variáveis latentes, onde somente as respostas não censuradas são completamente observadas. Sob o paradigma bayesiano, variáveis latentes recebem tratamento similar ao dados aos parâmetros. Logo, os valores censurados de y_i assim como os dados aumentados g_i podem ser obtidos amostrando-se de suas distribuições *a posteriori*. Para esse fim, um esquema MCMC é considerado e as distribuições condicionais completas *a posteriori* (dccp) devem ser obtidas.

Para que a especificação do modelo fique completa, *a priori*, assume-se que $\beta_1 \sim N_1(\mu_1, \sigma_1)$, $\boldsymbol{\beta}_2 \sim N_k(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$, $\sigma_u \sim IG(\alpha_u, \beta_u)$ e $\nu \sim TE(\lambda; 2)$, onde $IG(a, b)$ é a distribuição gama-inversa com parâmetros $a > 0$ e $b > 0$ e $TE(\lambda; L)$ denota a distribuição exponencial com média $\lambda > 0$ que possui massa acima de L.

Para estabelecer notação, seja $\mathbf{D}_c = (\mathbf{D}_o, \mathbf{D}_l)$ o conjunto de dados completos, onde $\mathbf{D}_o$ é o conjunto de dados observados como descrito anteriormente e $\mathbf{D}_l = (\mathbf{D}_{l1}, \ldots, \mathbf{D}_{ln})$, com $\mathbf{D}_{li} = (\eta_i, g_i)$, o conjunto de variáveis latentes, $i = 1, \ldots, n$.

Considerando a representação estocástica em (2.9), as respostas não observadas de η_i são variáveis aleatórias independentes com uma distribuição t-Student truncada acima de zero. Consequentemente, dado que y_i é uma observação censurada, a dccp de η_i é

$$\eta_i \mid g_i, y_i = 0, d_i = 0, \mathbf{x}_i, \boldsymbol{\theta} \sim TN(\beta_1 + \boldsymbol{\beta}_2' \mathbf{x}_i, g_i^{-1}\sigma_u; 0), \tag{2.10}$$

onde $TN(M, V; a)$ denota a distribuição normal univariada com média M e variância V com massa abaixo de a. A dccp de g_i é uma distribuição gama dada por

$$g_i \mid \eta_i, D_{oi}, \boldsymbol{\theta} \sim G\left(\frac{1+\nu}{2}, \frac{2\sigma_u}{(1-d_i)a_{1i}^2 + d_i a_{2i}^2 + \sigma_u \nu}\right), \tag{2.11}$$

onde $a_{1i} = \eta_i - (\beta_1 + \boldsymbol{\beta}_2' \mathbf{x}_i)$ e $a_{2i} = y_i - (\beta_1 + \boldsymbol{\beta}_2' \mathbf{x}_i)$, $i = 1, \ldots, n$.

Seja $\boldsymbol{\theta} = (\beta_1, \boldsymbol{\beta}_2, \sigma_u, \nu)$ e denote por $\boldsymbol{\theta}_{-a}$ a parte de $\boldsymbol{\theta}$ sem a componente a. Exceto para o grau de liberdade ν, as distribuições condicionais completas *a posteriori* de cada componente de $\boldsymbol{\theta}$ tem forma fechada e são dadas por

$$\beta_1 \mid \boldsymbol{\theta}_{-\beta_1}, \mathbf{D}_c \sim N_1\left(H_1 \left\{\sigma_1^{-1}\mu_1 + \sigma_u^{-1}\sum_{i=1}^{n} g_t\left[(1-d_i)\eta_i + d_i y_i - \boldsymbol{\beta}_2' \mathbf{x}_i\right]\right\}, H_1\right),$$

$$\boldsymbol{\beta}_2 \mid \boldsymbol{\theta}_{-\boldsymbol{\beta}_2}, \mathbf{D}_c \sim N_k\left(\mathbf{H}_2 \left\{\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2 + \sigma_u^{-1}\sum_{i=1}^{n} g_t\left[(1-d_i)\eta_i + d_i y_i - \beta_1\right]\mathbf{x}_i\right\}, \mathbf{H}_2\right),$$

$$\sigma_u \mid \boldsymbol{\theta}_{-\sigma_u}, \mathbf{D}_c \;\sim\; IG\left(\alpha_u + \frac{n}{2}, \beta_u + \frac{1}{2}\sum_{i=1}^{n} g_i\left[(1-d_i)a_{1i}^2 + d_i a_{2i}^2\right]\right),$$

em que

$$H_1^{-1} \;=\; \sigma_1^{-1} + \sigma_u^{-1}\textstyle\sum_{i=1}^{n} g_i,$$

$$\mathbf{H}_2^{-1} \;=\; \boldsymbol{\Sigma}_2^{-1} + \sigma_u^{-1}\textstyle\sum_{i=1}^{n} g_i \mathbf{x}_i \mathbf{x}_i'.$$

A dccp de ν não possui forma fechada e seu núcleo é

$$\pi(\nu \mid \boldsymbol{\theta}_{-\nu}, \mathbf{D}_c) \propto \left[\frac{(2/\nu)^{-\nu/2}}{\Gamma(\nu/2)}\right]^n \exp\left(-\nu\left[\frac{1}{2}\left(\sum_{i=1}^{n} g_i - \log(g_i)\right) + \frac{1}{\lambda}\right]\right),\ \nu > 2. \quad (2.12)$$

A dccp de η_i dada em (2.10) segue de maneira direta da hipótese. Como uma consequência, amostras da distribuição *a posteriori* de η_i, g_i e $\boldsymbol{\theta}_{-\nu}$ são obtidas usando o amostrador de Gibbs. O algoritmo de amostragem Metropolis por rejeição adaptativa (ARMS) é usado para amostrar da distribuição *a posteriori* de ν. Detalhes sobre o algoritmo ARMS podem ser encontrados em Gilks *et al.* (1995) e Gilks e Wild (1992), por exemplo.

Esse modelo foi implementado e será comparado no Capítulo 5 aos modelos Tobit e aos modelos com erros nas covariáveis N-Tobit e t-Tobit, que serão introduzidos no próximo capítulo.

Capítulo 3

Modelos tipo Tobit com erros nas covariáveis na família de distribuições normais independentes

A especificação de um modelo linear censurado com erros nas covariáveis também se inicia com uma relação linear dada por

$$\eta_i = \beta_1 + \boldsymbol{\beta}_2' \boldsymbol{\xi}_i + \boldsymbol{\beta}_3' \mathbf{b}_i + u_i, \tag{3.1}$$

$i = 1, \ldots, n$, onde $\eta_i \in \mathbb{R}$ é a variável resposta, $(\boldsymbol{\xi}_i, \mathbf{b}_i) \in \mathbb{R}^{k+l}$ são as variáveis explicativas, $(\beta_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3) \in \mathbb{R}^{1+k+l}$ são os parâmetros de regressão e $u_i \in \mathbb{R}$ é o erro aleatório. Da mesma forma que no modelo linear censurado sem erros nas covariáveis tem-se que a variável resposta observada y_i é definida por (2.2), ou seja, $y_i = \max\{\eta_i, 0\}$, $i = 1, \ldots, n$. As equações (3.1) e (2.2) definem o usual (livre de erros) modelo de regressão linear censurado quando todas as variáveis explicativas, $\boldsymbol{\xi}_i$ e $\mathbf{b}_i$, são observadas exatamente. Neste capítulo, assume-se que a variável explicativa $\mathbf{b}_i$ é exatamente medida, enquanto a variável independente $\boldsymbol{\xi}_i$ não é exatamente observada. Ao invés disso observa-se a covariável $\mathbf{x}_i \in \mathbb{R}^k$ que é igual à verdadeira covariável $\boldsymbol{\xi}_i$ acrescida de um erro de medida aleatório $\mathbf{v}_i \in \mathbb{R}^k$, isto é,

$$\mathbf{x}_i = \boldsymbol{\xi}_i + \mathbf{v}_i, \tag{3.2}$$

$i = 1, \ldots, n$. O modelo censurado linear com erros nas covariáveis é então definido pelas equações (3.1), (2.2) e (3.2).

Tipicamente, em modelos com erros de medida assume-se que u_i e $\mathbf{v}_i$ são erros aleatórios não correlacionados, enquanto a covariável latente $\boldsymbol{\xi}_i$ pode ser tanto considerada

como fixa (modelo funcional) quanto como uma variável aleatória a qual é não correlacionada com os erros u_i e $\mathbf{v}_i$ (modelo estrutural). Neste capítulo, a maioria dos resultados serão obtidos sob o modelo estrutural. Nesse caso, é conveniente formular conjuntamente as relações lineares (3.1) e (3.2) como

$$\begin{pmatrix} \eta_i \\ \mathbf{x}_i \end{pmatrix} = \begin{pmatrix} \beta_1 + \boldsymbol{\beta}_3' \mathbf{b}_i \\ \mathbf{0} \end{pmatrix} + \begin{pmatrix} 1 & \mathbf{0}' & \boldsymbol{\beta}_2' \\ \mathbf{0} & \boldsymbol{I}_k & \boldsymbol{I}_k \end{pmatrix} \begin{pmatrix} u_i \\ \mathbf{v}_i \\ \boldsymbol{\xi}_i \end{pmatrix}, \tag{3.3}$$

$i = 1, \ldots, n$. Logo, cada modelo Tobit com erro de medida estrutural é determinado pela distribuição conjunta das variáveis aleatórias latentes u_i, $\mathbf{v}_i$ e $\boldsymbol{\xi}_i$. Há muitas opções para especificar essa distribuição conjunta. Assim como ocorreu com (2.1) nos modelos Tobit sem erros nas covariáveis, devido a estrutura linear da relação (3.3), pode-se considerar distribuições multivariadas com a propriedade de manter a família de distribuições sob transformações lineares. Nesse sentido, distribuições elípticas multivariadas também são opções naturais. Similar ao caso univariado, a classe de distribuições elípticas contem muitas distribuições multivariadas simétricas que podem modelar variáveis aleatórias com caudas leves ou pesadas. A subclasse de distribuições normais independentes multivariada é um exemplo. Diz-se que $(u_i, \mathbf{v}_i, \boldsymbol{\xi}_i)'$ pertence à família de distribuições normais independentes multivariada se

$$\begin{pmatrix} u_i \\ \mathbf{v}_i \\ \boldsymbol{\xi}_i \end{pmatrix} \stackrel{d}{=} \begin{pmatrix} 0 \\ \mathbf{0} \\ \boldsymbol{\mu}_\xi \end{pmatrix} + g_i^{-1/2} \mathbf{Z}_i, \tag{3.4}$$

onde $\mathbf{Z}_i$ é um vetor aleatório normalmente distribuído com vetor de médias $\mathbf{0}$, matriz de variância bloco diagonal $\boldsymbol{\Omega} = \text{bdiag}(\sigma_u, \boldsymbol{\Sigma}_v, \boldsymbol{\Sigma}_\xi)$, em que $\boldsymbol{\Sigma}_v$ e $\boldsymbol{\Sigma}_\xi$ são matrizes de variância e g_i é uma variável aleatória de mistura a qual é positiva e independente de $\mathbf{Z}_i$ e cuja distribuição é indexada por um parâmetro ν, $i = 1, \ldots, n$. Condicionalmente em g_i, o vetor $\mathbf{r}_i = (u_i, \mathbf{v}_i', \boldsymbol{\xi}_i')'$ possui uma distribuição normal multivariada com vetor de médias $\boldsymbol{\omega} = (0, \mathbf{0}', \boldsymbol{\mu}_\xi')'$ e matriz de variância $g_i^{-1}\boldsymbol{\Omega}$, $i = 1, \ldots, n$. Se a distribuição de g_i é degenerada atribuindo massa em 1 para todo $i = 1, \ldots, n$, então $\mathbf{r}_i$ possui distribuição normal multivariada e obtém-se o modelo apresentado por Wang (1998), com a diferença que aqui também considera a presença de covariáveis medidas sem erro (ver detalhes na Seção 3.1). No entanto, mesmo no caso multivariado, sabe-se que a distribuição

normal não é robusta o suficiente para modelar conjuntamente variáveis aleatórias com caudas pesadas. Estendendo o modelo de Wang (1998), neste trabalho, considera-se em (3.4) que g_i possui distribuição gama com parâmetros de forma e escala iguais a $\nu/2$ e $2/\nu$, respectivamente. Isso é equivalente a afirmar que $\mathbf{r}_i$ possui uma distribuição t multivariada, que é uma generalização da distribuição normal multivariada e é capaz de modelar caudas pesadas. Como consequência, em (3.3) obtem-se uma distribuição t multivariada para $(\eta_i, \mathbf{x}_i')'$. Este modelo é detalhadamente introduzido na Seção 3.2.

Na próxima seção, o modelo proposto por Wang (1998), aqui chamado de modelo N-Tobit com erros nas covariáveis, é apresentado mais detalhadamente.

3.1 O modelo N-Tobit com erros nas covariáveis

O modelo N-Tobit com erro de medida, definido por Wang (1998), assume em (3.4) que $P(g_i = 1) = 1$ para todo $i = 1, \ldots, n$, o que é equivalente afirmar que u_i, $\mathbf{v}_i$ e $\boldsymbol{\xi}_i$ em (3.3) são quantidades aleatórias independentes e normalmente distribuídas com médias 0, $\mathbf{0}$, $\boldsymbol{\mu}_\xi$ e variâncias σ_u, $\boldsymbol{\Sigma}_v$ e $\boldsymbol{\Sigma}_\xi$, respectivamente (veja também Wang (2002)). O modelo N-Tobit com erros nas covariáveis pode ser representado como segue

$$\left.\begin{array}{rcl} y_i & = & \max\{\eta_i, 0\}, \\ \begin{pmatrix}\eta_i \\ \mathbf{x}_i\end{pmatrix} \mid \mathbf{b}_i & \overset{ind.}{\sim} & N_{1+k}\left(\begin{pmatrix}\beta_1 + \boldsymbol{\beta}_2'\boldsymbol{\mu}_\xi + \boldsymbol{\beta}_3'\mathbf{b}_i \\ \boldsymbol{\mu}_\xi\end{pmatrix}, \begin{pmatrix}\sigma_u + \boldsymbol{\beta}_2'\boldsymbol{\Sigma}_\xi\boldsymbol{\beta}_2 & \boldsymbol{\beta}_2'\boldsymbol{\Sigma}_\xi \\ \boldsymbol{\Sigma}_\xi\boldsymbol{\beta}_2 & \boldsymbol{\Sigma}_v + \boldsymbol{\Sigma}_\xi\end{pmatrix}\right), \end{array}\right\} \quad (3.5)$$

para todo $i = 1, \ldots, n$. Para ser mais preciso, o modelo em (3.5) é obtido integrando a função de densidade de probabilidade (fdp) de $\mathbf{r}_i = (u_i, \mathbf{v}_i, \boldsymbol{\xi}_i)$ em relação ao vetor de covariáveis latentes (verdadeiras) $\boldsymbol{\xi}_i$.

Assumindo

$$\left.\begin{array}{rcl} \gamma_1 & = & \beta_1 + \boldsymbol{\beta}_2'\boldsymbol{\Sigma}_v(\boldsymbol{\Sigma}_v + \boldsymbol{\Sigma}_\xi)^{-1}\boldsymbol{\mu}_\xi, \\ \boldsymbol{\gamma}_2 & = & (\boldsymbol{\Sigma}_\xi + \boldsymbol{\Sigma}_v)^{-1}\boldsymbol{\Sigma}_\xi\boldsymbol{\beta}_2, \\ \boldsymbol{\gamma}_3 & = & \boldsymbol{\beta}_3, \\ \sigma_w & = & \sigma_u + \boldsymbol{\beta}_2'\boldsymbol{\Sigma}_v(\boldsymbol{\Sigma}_\xi + \boldsymbol{\Sigma}_v)^{-1}\boldsymbol{\Sigma}_\xi\boldsymbol{\beta}_2, \\ \boldsymbol{\mu}_x & = & \boldsymbol{\mu}_\xi, \\ \boldsymbol{\Sigma}_x & = & \boldsymbol{\Sigma}_\xi + \boldsymbol{\Sigma}_v, \end{array}\right\} \quad (3.6)$$

o modelo em (3.5) pode ser reescrito da seguinte forma

$$\left.\begin{array}{rcl} y_i & = & \max\{\eta_i, 0\}, \\ \begin{pmatrix}\eta_i \\ \mathbf{x}_i\end{pmatrix} \mid \mathbf{b}_i & \overset{ind.}{\sim} & N_{1+k}\left(\begin{pmatrix}\gamma_1 + \boldsymbol{\gamma}_2'\boldsymbol{\mu}_x + \boldsymbol{\gamma}_3'\mathbf{b}_i \\ \boldsymbol{\mu}_x\end{pmatrix}, \begin{pmatrix}\sigma_w + \boldsymbol{\gamma}_2'\boldsymbol{\Sigma}_x\boldsymbol{\gamma}_2 & \boldsymbol{\gamma}_2'\boldsymbol{\Sigma}_x \\ \boldsymbol{\Sigma}_x\boldsymbol{\gamma}_2 & \boldsymbol{\Sigma}_x\end{pmatrix}\right), \end{array}\right\} \quad (3.7)$$

para todo $i = 1, \ldots, n$ e, considerando-se a decomposição marginal de $f(\eta_i, \mathbf{x}_i|\mathbf{b}_i)$, dada por $f(\eta_i, \mathbf{x}_i|\mathbf{b}_i) = f(\eta_i|\mathbf{x}_i, \mathbf{b}_i)f(\mathbf{x}_i|\mathbf{b}_i)$, induz-se a seguinte representação reduzida de (3.7):

$$\left.\begin{array}{rcl} y_i & = & \max\{\eta_i, 0\}, \\ \eta_i \mid \mathbf{b}_i, \mathbf{x}_i & \sim & N_1\left(\gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x}_i + \boldsymbol{\gamma}_3'\mathbf{b}_i, \sigma_w\right), \\ \mathbf{x}_i \mid \mathbf{b}_i & \sim & N_k(\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x), \end{array}\right\} \quad (3.8)$$

$i = 1, \ldots, n$.

Pode-se notar de (3.8) que a relação linear original em (3.1) foi reduzida a uma relação linear "livre de erros" dada por $\eta_i = E(\eta_i \mid \mathbf{x}_i, \mathbf{b}_i) + w_i$, em que $E(\eta_i \mid \mathbf{x}_i, \mathbf{b}_i) = \gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x}_i + \boldsymbol{\gamma}_3'\mathbf{b}_i$ e o erro $w_i = \eta_i - E(\eta_i \mid \mathbf{x}_i, \mathbf{b}_i) \sim N_1(0, \sigma_w)$, que é, por hipótese, independente de $\mathbf{x}_i$. A diferença para o modelo livre de erro é que $\mathbf{x}_i$ é, agora, um componente aleatório. Esse resultado também foi encontrado por Wang (1998, 2002) porém de uma maneira diferente.

No entanto, pode-se observar claramente da expressão (3.6) que a relação entre os $(k+1)(k+2)+l$ diferentes parâmetros originais em $(\beta_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \sigma_u, \boldsymbol{\mu}_\xi, v(\boldsymbol{\Sigma}_\xi), v(\boldsymbol{\Sigma}_v))$, onde $v(\boldsymbol{\Sigma})$ denota o vetor formado pelos $k(k+1)/2$ diferentes elementos de uma matriz simétrica $\boldsymbol{\Sigma}$ e os $(k+1)(k/2+2)+l$ diferentes parâmetros reduzidos em $(\gamma_1, \boldsymbol{\gamma}_2, \boldsymbol{\gamma}_3, \sigma_w, \boldsymbol{\mu}_x, v(\boldsymbol{\Sigma}_x))$ não é uma relação um-a-um e, portanto, as formulações (3.5) e (3.7) para o modelo N-Tobit com erro de medida não são completamente equivalentes. Isso é devido ao fato de que o modelo estrutural dado pela segunda expressão em (3.5) é não identificável (Fuller, 1987; Gleser, 1992). Consequentemente, precisa-se de $k(k+1)/2$ condições adicionais na estrutura paramétrica original para tornar o modelo (3.5) identificável e, consequentemente, equivalente à formulação em (3.7). Essas condições de identificabilidade podem ser especificadas de diferentes maneiras, dependendo da natureza do problema e da informação disponível. Por exemplo, em muitos casos a reamostragem ou a validação contínua de dados pode ser usada para conhecer ou estimar a variância $\boldsymbol{\Sigma}_v$ dos erros de medida.

Além do mais, o conhecimento de $\boldsymbol{\Sigma}_v$ permite conhecer a chamada razão de confiabilidade $\boldsymbol{\Lambda} = \boldsymbol{\Sigma}_\xi \boldsymbol{\Sigma}_x^{-1}$. Isso porque $\boldsymbol{\Sigma}_\xi = \boldsymbol{\Sigma}_x - \boldsymbol{\Sigma}_v$, que implica que $\boldsymbol{\Lambda} = \boldsymbol{I}_k - \boldsymbol{\Sigma}_v \boldsymbol{\Sigma}_x^{-1}$ e a matriz de variância $\boldsymbol{\Sigma}_x$ pode ser estimada consistentemente pela matriz de variância amostral $\mathbf{S}_x = n^{-1} \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})'$, onde $\bar{\mathbf{x}} = n^{-1} \sum_{i=1}^n \mathbf{x}_i$. Por sua vez, conhecer a matriz de razão de confiabilidade $\boldsymbol{\Lambda}$ é equivalente a conhecer a chamada matriz de razão de ruído-sinal $\boldsymbol{\Delta} = \boldsymbol{\Sigma}_\xi^{-1} \boldsymbol{\Sigma}_v$ pois $\boldsymbol{\Lambda} = (\boldsymbol{I}_k + \boldsymbol{\Delta})^{-1}$. Como em Wang (1998), assume-se ao longo deste trabalho que $\boldsymbol{\Delta}$ é conhecida, de modo que $\boldsymbol{\Sigma}_v = \boldsymbol{\Sigma}_\xi \boldsymbol{\Delta}$, reduzindo, assim, os parâmetros desconhecidos originais para $(\beta_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \sigma_u, \boldsymbol{\mu}_\xi, v(\boldsymbol{\Sigma}_\xi))$ e obtendo uma relação um-a-um com os parâmetros reduzidos $(\gamma_1, \boldsymbol{\gamma}_2, \boldsymbol{\gamma}_3, \sigma_w, \boldsymbol{\mu}_x, v(\boldsymbol{\Sigma}_x))$. Portanto, invertendo o sistema (3.6) tem-se que

$$\left.\begin{array}{rcl} \beta_1 &=& \gamma_1 - \boldsymbol{\mu}_x' \boldsymbol{\Delta} \boldsymbol{\gamma}_2, \\ \boldsymbol{\beta}_2 &=& (\mathbf{I}_k + \boldsymbol{\Delta}) \boldsymbol{\gamma}_2, \\ \boldsymbol{\beta}_3 &=& \boldsymbol{\gamma}_3, \\ \sigma_u &=& \sigma_w - \boldsymbol{\gamma}_2' \boldsymbol{\Sigma}_x \boldsymbol{\Delta} \boldsymbol{\gamma}_2, \\ \boldsymbol{\mu}_\xi &=& \boldsymbol{\mu}_x, \\ \boldsymbol{\Sigma}_\xi &=& \boldsymbol{\Sigma}_x (\mathbf{I}_k + \boldsymbol{\Delta})^{-1}, \\ \boldsymbol{\Sigma}_v &=& \boldsymbol{\Sigma}_x [\mathbf{I}_k - (\mathbf{I}_k + \boldsymbol{\Delta})^{-1}]. \end{array}\right\} \tag{3.9}$$

Pode-se observar que a escolha de $\boldsymbol{\Delta}$ não é completamente arbitrária. Como σ_u é positivo o valor de $\boldsymbol{\Delta}$ deve ser tal que $\boldsymbol{\gamma}_2' \boldsymbol{\Sigma}_x \boldsymbol{\Delta} \boldsymbol{\gamma}_2$ seja estritamente menor que σ_w. A partir de (3.9) nota-se também que as estimações de $\boldsymbol{\beta}_3$ e $\boldsymbol{\mu}_\xi$ não são afetadas com a escolha de $\boldsymbol{\Delta}$.

Do ponto de vista bayesiano, a não identificabilidade não deveria causar uma dificuldade uma vez que pode-se eliciar distribuições *a priori* próprias e informativas para todos os parâmetros. No entanto, como mencionado em Swartz *et al.* (2004), na presença de não identificabilidade, a informação dos dados pode não dominar a informação *a priori*, mesmo se são consideradas grandes amostras. Logo, inferências *a posteriori* pobres poderiam ser feitas. A não identificabilidade na verossimilhança também pode ocasionar inferências *a posteriori* pobres se elas são obtidas através de ajustes de modelos baseados em MCMC como mostrado em Gelfand e Sahu (1999). Neste trabalho, para fins de comparação, a sugestão em Wang (1998) será seguida, isto é, será considerado que $\boldsymbol{\Delta}$ é fixo e conhecido.

Como notado por Wang (1998), outro aspecto relevante no modelo N-Tobit com erros nas covariáveis, que segue claramente de sua representação condicional marginal dada em

(3.8), é que para todo $i = 1, \ldots, n$, a distribuição condicional de $\eta_i \mid \mathbf{x}_i, \mathbf{b}_i$ não depende de $(\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x)$. Consequentemente, a estimação de máxima verossimilhança desses parâmetros não é afetada pelo processo de censura. De fato, o estimador de máxima verossimilhança de $(\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x)$ é dado pelos correspondentes estimadores de momentos amostrais, $(\bar{\mathbf{x}}, \mathbf{S}_x)$, uma vez que $\mathbf{x}_i \mid \mathbf{b}_i \sim N_k(\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x)$, como representado em (3.8).

Para ser mais preciso, denote por $\boldsymbol{\theta}$ o vetor formado pelos $(k+1)(k/2+2)+l$ diferentes parâmetros em $(\gamma_1, \boldsymbol{\gamma}_2, \boldsymbol{\gamma}_3, \sigma_w, \boldsymbol{\mu}_x, v(\boldsymbol{\Sigma}_x))$ e por $\mathbf{D}_{\mathrm{o}} = (\mathbf{D}_{o1}, \ldots, \mathbf{D}_{on})$ os dados observados para todos os n indivíduos, onde $\mathbf{D}_{oi} = (y_i, d_i, \mathbf{x}_i, \mathbf{b}_i)$ são os dados observados para o i-ésimo sujeito e $d_i = \mathbf{1}(y_i > 0)$ é a função indicadora de censura para o i-ésimo sujeito sendo igual a 1 se sua variável resposta é não censurada, $i = 1, \ldots, n$. Assim, considerando (3.8) a função de log-verossimilhança para o modelo N-Tobit com erro de medida para uma amostra de n observações independentes pode ser expressa como

$$\begin{aligned} \log f(\mathbf{D}_{\mathrm{o}} \mid \boldsymbol{\theta}) &= \sum_{i=1}^{n} \log \phi_k(\mathbf{x}_i \mid \boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x) + \sum_{i=1}^{n} d_i \log \phi_1\left(y_i \mid \gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x}_i + \boldsymbol{\gamma}_3'\mathbf{b}_i, \sigma_w\right) \\ &\quad + \sum_{i=1}^{n} (1 - d_i) \log \Phi_1\left(0 \mid \gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x}_i + \boldsymbol{\gamma}_3'\mathbf{b}_i, \sigma_w\right). \end{aligned} \tag{3.10}$$

A função de verossimilhança em (3.10) se parece com aquela obtida para o modelo Tobit "livre de erros", onde, agora, o primeiro termo representa a contribuição devido ao comportamento aleatório da covariável observada $\mathbf{x}_i$.

3.2 O modelo t-Tobit com erros nas covariáveis

Nesta seção será introduzido o modelo com erro de medida t-Tobit. Para isso, considere, em (3.4), que $g_i \sim G(\nu/2, 2/\nu)$, $i = 1, \ldots, n$, o que é equivalente a considerar em (3.3) uma distribuição conjunta t multivariada para u_i, $\mathbf{v}_i$ e $\boldsymbol{\xi}_i$. Mais precisamente, assuma que $\mathbf{r}_i$, $i = 1, \ldots, n$, são vetores aleatórios independentes e identicamente distribuídos (iid) seguindo uma distribuição t-Student $(1 + 2k)$-variada com vetor de locação $\boldsymbol{\omega}$, matriz de dispersão bloco diagonal $\boldsymbol{\Omega}$ e graus de liberdade $\nu > 0$, que é denotado por

$$\mathbf{r}_i \overset{iid}{\sim} t_{1+2k}(\boldsymbol{\omega}, \boldsymbol{\Omega}, \nu), \tag{3.11}$$

$i = 1, \ldots, n$. Para uma revisão das principais propriedades de distribuições t-Student multivariadas veja, por exemplo, Arellano-Valle e Bolfarine (1995) e Fang *et al.* (1990).

Sob a suposição (3.11) tem-se que $E(\mathbf{r}_i) = \boldsymbol{\omega}$ para $\nu > 1$ e $Var(\mathbf{r}_i) = \nu/(\nu-2)\boldsymbol{\Omega}$ para $\nu > 2$. Como $\boldsymbol{\Omega}$ é bloco diagonal, essa última propriedade implica, sob a suposição (3.11), que u_i, $\mathbf{v}_i$ e $\boldsymbol{\xi}_i$ são quantidades aleatórias não correlacionadas, mas não são independentes para valores finitos de ν. Mais ainda, como dito anteriormente, a distribuição t multivariada é fechada sob transformações lineares. Isto é, de (3.11) tem-se, para qualquer vetor fixo $\mathbf{a} \in \mathbb{R}^m$ e matriz $\mathbf{B} \in \mathbb{R}^{m\times(1+2k)}$, que $\mathbf{a} + \mathbf{B}\mathbf{r}_i \sim t_m(\mathbf{a} + \mathbf{B}\boldsymbol{\omega}, \mathbf{B}\boldsymbol{\Omega}\mathbf{B}', \nu)$. Portanto, considerando (2.2) e usando essa última propriedade na relação linear conjunta (3.3), obtem-se o modelo t-Tobit com erros nas covariáveis estrutural definido por

$$\left.\begin{array}{rcl} y_i & = & \max\{\eta_i, 0\}, \\ \begin{pmatrix}\eta_i \\ \mathbf{x}_i\end{pmatrix} \mid \mathbf{b}_i & \overset{ind.}{\sim} & t_{1+k}\left(\begin{pmatrix}\beta_1 + \boldsymbol{\beta}_2'\boldsymbol{\mu}_\xi + \boldsymbol{\beta}_3'\mathbf{b}_i \\ \boldsymbol{\mu}_\xi\end{pmatrix}, \begin{pmatrix}\sigma_u + \boldsymbol{\beta}_2'\boldsymbol{\Sigma}_\xi\boldsymbol{\beta}_2 & \boldsymbol{\beta}_2'\boldsymbol{\Sigma}_\xi \\ \boldsymbol{\Sigma}_\xi\boldsymbol{\beta}_2 & \boldsymbol{\Sigma}_v + \boldsymbol{\Sigma}_\xi\end{pmatrix}, \nu\right), \end{array}\right\} \quad (3.12)$$

$i = 1, \ldots, n$.

Ao considerar-se em (3.12) que $\nu = 1$, tem-se um modelo censurado com erro de medida Cauchy e, quando $\nu \to \infty$, o modelo (3.12) converge para um modelo N-Tobit com erros nas covariáveis dado em (3.5). Portanto, pode-se afirmar que o modelo t-Tobit com erro de medida fornece uma generalização robusta do modelo N-Tobit com erro de medida, uma vez que pode-se modelar observações cujas distribuições possuem mais curtose que a da normal.

Similar ao modelo estrutural N-Tobit mostrado em (3.5), o modelo estrutural t-Tobit dado na expressão (3.12) é também não identificável porque a versão t estrutural dada pela segunda expressão de (3.12) não o é (Arellano-Valle e Bolfarine (1996); Bolfarine e Arellano-Valle (1998)). Portanto, para se garantir a identificabilidade na função de verossimilhança dada em (3.12), assume-se novamente que $\boldsymbol{\Delta} = \boldsymbol{\Sigma}_\xi^{-1}\boldsymbol{\Sigma}_v$ é uma matriz conhecida.

Assim como no modelo N-Tobit, do segundo termo de (3.12) obtem-se as distribuições condicionais de $\eta_i|\mathbf{x}_i, \mathbf{b}_i$ e a marginal de $\mathbf{x}_i$ (veja por exemplo Arellano-Valle e Bolfarine, 1995). Consequentemente, o modelo estrutural t-Tobit pode também ser reduzido hierar-

quicamente e assume a seguinte forma

$$\left.\begin{array}{rcl} y_i & = & \max\{\eta_i, 0\}, \\ \eta_i \mid \mathbf{b}_i, \mathbf{x}_i & \sim & t_1\left(\gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x}_i + \boldsymbol{\gamma}_3'\mathbf{b}_i, \left(\frac{\nu+q(\mathbf{x}_i)}{\nu+k}\right)\sigma_w, \nu+k\right), \\ \mathbf{x}_i \mid \mathbf{b}_i & \sim & t_k(\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x, \nu), \end{array}\right\} \tag{3.13}$$

para todo $i = 1, \ldots, n$, onde os parâmetros reduzidos $(\gamma_1, \boldsymbol{\gamma}_2, \boldsymbol{\gamma}_3, \sigma_w, \boldsymbol{\mu}_x, v(\boldsymbol{\Sigma}_x))$ são os mesmos dados na expressão (3.6) e $q(\mathbf{x}_i) = (\mathbf{x}_i - \boldsymbol{\mu}_x)'\boldsymbol{\Sigma}_x^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_x)$. Assim como no modelo N-Tobit com erros nas covariáveis, as $k(k+1)/2$ condições necessárias para a identificabilidade do modelo com estrutura paramétrica dada em (3.12) são necessárias para garantir a equivalência com a representação do modelo estrutural t-Tobit dada em (3.13). Em particular, essa equivalência ocorre quando se assume que $\boldsymbol{\Delta} = \boldsymbol{\Sigma}_\xi^{-1}\boldsymbol{\Sigma}_v$ é uma matriz conhecida.

Por outro lado, diferentemente do que foi observado no modelo estrutural N-Tobit, condicionalmente em $\mathbf{x}_i$, para o modelo proposto, segue de (3.13) que a distribuição condicional de $\eta_i \mid \mathbf{x}_i, \mathbf{b}_i$ depende de $(\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x)$ através da função quadrática $q(\mathbf{x}_i)$. Essa dependência desaparece quando $\nu \to \infty$ e, portanto, o modelo em (3.13) torna-se o modelo em (3.8). Mais ainda, dessa representação hierárquica tem-se que

$$E(\eta_i \mid \mathbf{b}_i, \mathbf{x}_i) = \gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x}_i + \boldsymbol{\gamma}_3'\mathbf{b}_i, \quad \nu + k > 1$$

e

$$Var(\eta_i \mid \mathbf{b}_i, \mathbf{x}_i) = \left(\frac{\nu + q(\mathbf{x}_i)}{\nu + k - 2}\right)\sigma_w, \quad \nu + k > 2.$$

Em outras palavras, em (3.13) encontra-se uma estrutura similar ao modelo t-Tobit "livre de erros" com uma relação linear $\eta_i = \gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x}_i + \boldsymbol{\gamma}_3'\mathbf{b}_i + w_i$, mas, nesse caso, tem-se que, condicionalmente nas covariáveis observadas, os erros aleatórios w_i são heterocedásticos e condicionalmente distribuídos tais que

$$w_i \mid \mathbf{b}_i, \mathbf{x}_i \overset{ind.}{\sim} t_1\left(0, \left(\frac{\nu + q(\mathbf{x}_i)}{\nu + k}\right)\sigma_w, \nu + k\right),$$

$i = 1, \ldots, n$. Como $E(w_i \mid \mathbf{b}_i, \mathbf{x}_i) = 0$ para $\nu + k > 1$, tem-se que $E(w_i\mathbf{x}_i \mid \mathbf{b}_i) = E(\mathbf{x}_i E(w_i \mid \mathbf{b}_i, \mathbf{x}_i)) = 0$ e, portanto, w_i e $\mathbf{x}_i$ são quantidades aleatórias não correlacionadas. Consequentemente, o modelo (3.13) difere do modelo t-Tobit usual introduzido por Arellano-Valle *et al.* (2012) em que todas as variáveis explicativas são consideradas

quantidades conhecidas. Se a variável medida com erro não está na equação de regressão, então $\eta_i = \gamma_1 + \boldsymbol{\beta}_3'\mathbf{b}_i + w_i$, onde $w_i \sim t_1(0, \sigma_w, \nu)$, que é exatamente o modelo proposto por Arellano-Valle *et al.* (2012).

Reconhecendo essas similaridades e diferenças, o processo de inferência no modelo proposto se torna mais simples uma vez que sob as condições de identificabilidade, há uma relação um-a-um entre os parâmetros do modelo (3.12) e aqueles no modelo reduzido dado em (3.13) a qual é dada em (3.9). Logo, sob as condições de identificabilidade consideradas, obtem-se de (3.13) que a função de log-verossimilhança do modelo t-Tobit com erro de medida estrutural para $\boldsymbol{\theta}$, os $(k+1)(k/2+2)+l+1$ diferentes parâmetros em $(\gamma_1, \boldsymbol{\gamma}_2, \boldsymbol{\gamma}_3, \sigma_w, \boldsymbol{\mu}_x, v(\boldsymbol{\Sigma}_x), \nu)$, baseada em uma amostra de n observações independentes pode ser expressa como

$$\begin{aligned}\log f(\mathbf{D}_o \mid \boldsymbol{\theta}) \;=\; & \sum_{i=1}^{n} \log t_k(\mathbf{x}_i \mid \boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x, \nu) \\ & + \sum_{i=1}^{n} d_i \log t_1\left(y_i \mid \gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x}_i + \boldsymbol{\gamma}_3'\mathbf{b}_i, \left(\frac{\nu + q(\mathbf{x}_i)}{\nu + k}\right)\sigma_w, \nu + k\right) \\ & + \sum_{i=1}^{n} (1 - d_i) \log T_1\left(0 \mid \gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x}_i + \boldsymbol{\gamma}_3'\mathbf{b}_i, \left(\frac{\nu + q(\mathbf{x}_i)}{\nu + k}\right)\sigma_w, \nu + k\right).\end{aligned} \tag{3.14}$$

Em (3.14), tem-se novamente que o primeiro termo corresponde a uma contribuição das covariáveis aleatórias $\mathbf{x}_i$, as quais, nesse caso, afetam outros termos da função de log-verossimilhança através da forma quadrática $q(\mathbf{x}_i)$. Esse efeito desaparece quando $\nu \to \infty$ e (3.14) converge para a função de log-verossimilhança do modelo estrutural N-Tobit em (3.10). Para o modelo proposto pode-se notar que, assim como para $\boldsymbol{\beta}_3$ e $\boldsymbol{\mu}_\xi$ (ver Seção 3.1), a estimação de ν não é afetada pela escolha de $\boldsymbol{\Delta}$.

3.2.1 Inferência bayesiana no modelo t-Tobit com erros nas covariáveis

Nesta seção, similarmente ao que foi feito para o modelo t-Tobit (Capítulo 2, Seção 2.2.1), um algoritmo para se obter amostras das distribuições *a posteriori* dos parâmetros do modelo t-Tobit com erro de medida é desenvolvido baseado nas técnicas MCMC e aumento de dados. Para isso, considera-se a representação estocástica da distribuição t-Student multivariada obtendo, assim, a estrutura hierárquica dada em (3.15). Com

isso, obtem-se, de (3.13), a seguinte formulação hierárquica do modelo t-Tobit com erro de medida:

$$\left.\begin{array}{rcl} y_i & = & \max\{\eta_i, 0\}, \\ \eta_i \mid \mathbf{b}_i, \mathbf{x}_i, g_i & \overset{ind.}{\sim} & N_1\left(\gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x}_i + \boldsymbol{\gamma}_3'\mathbf{b}_i, g_i^{-1}\sigma_w\right), \\ \mathbf{x}_i \mid \mathbf{b}_i, g_i & \overset{ind.}{\sim} & N_k\left(\boldsymbol{\mu}_x, g_i^{-1}\boldsymbol{\Sigma}_x\right), \\ g_i \mid \mathbf{b}_i & \overset{iid}{\sim} & G\left(\nu/2, 2/\nu\right), \end{array}\right\} \tag{3.15}$$

$i = 1, \ldots, n$, onde g_i são variáveis aleatórias latentes representando os fatores de mistura de escala.

Para completar a especificação do modelo assume-se, *a priori*, que $\boldsymbol{\Delta}$ é uma quantidade conhecida e que $\gamma_1 \sim \mathrm{N}_1(\mu_1, \sigma_1)$, $\boldsymbol{\gamma}_2 \sim \mathrm{N}_k(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$, $\boldsymbol{\gamma}_3 \sim \mathrm{N}_{k_1}(\boldsymbol{\mu}_3, \boldsymbol{\Sigma}_3)$, $\sigma_w \sim \mathrm{IG}(\alpha_w, \beta_w)$, $\boldsymbol{\mu}_x \sim \mathrm{N}_k(\boldsymbol{\mu}_4, \boldsymbol{\Sigma}_4)$, $\boldsymbol{\Sigma}_x \sim \mathrm{IW}(m_1, \boldsymbol{\Psi}_1)$ e $\nu \sim \mathrm{TE}(\lambda; 2)$, em que $\mathrm{IW}(m_1, \boldsymbol{\Psi}_1)$ denota a distribuição Wishart-Inversa, onde $\boldsymbol{\Psi}_1$ é uma matriz positiva definida de ordem k e $m_1 > 0$. Como consequência, as distribuições *a priori* para os parâmetros de regressão no modelo original são

$$\begin{aligned} \beta_1 \mid \boldsymbol{\beta}_2 &\sim N_1(\mu_1 - \boldsymbol{\mu}_4'\boldsymbol{\Delta}(\mathbf{I}_k + \boldsymbol{\Delta})^{-1}\boldsymbol{\beta}_2, \sigma_1 + [\boldsymbol{\Delta}(\mathbf{I}_k + \boldsymbol{\Delta})^{-1}\boldsymbol{\beta}_2]'\boldsymbol{\Sigma}_4\boldsymbol{\Delta}(\mathbf{I}_k + \boldsymbol{\Delta})^{-1}\boldsymbol{\beta}_2), \\ \boldsymbol{\beta}_2 &\sim N_k((\mathbf{I}_k + \boldsymbol{\Delta})\boldsymbol{\mu}_2, (\mathbf{I}_k + \boldsymbol{\Delta})\boldsymbol{\Sigma}_2(\mathbf{I}_k + \boldsymbol{\Delta})'). \end{aligned}$$

Considerando a distribuição condicional de η_i apresentada em (3.15) pode-se observar que as respostas não observadas de η_i também são variáveis aleatórias independentes com uma distribuição t-Student com massa abaixo de zero. Logo, condicional a uma resposta censurada de y_i, tem-se que a distribuição condicional completa *a posteriori* (dccp) de η_i sob o modelo proposto é

$$\eta_i \mid g_i, y_i = 0, d_i = 0, \mathbf{x}_i, \mathbf{b}_i, \boldsymbol{\theta} \sim \mathrm{TN}(\gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x}_i + \boldsymbol{\gamma}_3'\mathbf{b}_i, g_i^{-1}\sigma_w; 0). \tag{3.16}$$

Além disso, a dccp de g_i é uma distribuição gama dada por

$$g_i \mid \eta_i, D_{oi}, \boldsymbol{\theta} \sim G\left(\frac{1 + k + \nu}{2}, \frac{2\sigma_w}{(1 - d_i)a_{1i}^2 + d_i a_{2i}^2 + \sigma_w q(\mathbf{x}_i) + \sigma_w \nu}\right), \tag{3.17}$$

onde $a_{1i} = \eta_i - (\gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x}_i + \boldsymbol{\gamma}_3'\mathbf{b}_i)$ e $a_{2i} = y_i - (\gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x}_i + \boldsymbol{\gamma}_3'\mathbf{b}_i)$.

Como definido anteriormente, seja $\boldsymbol{\theta} = (\gamma_1, \boldsymbol{\gamma}_2, \boldsymbol{\gamma}_3, \sigma_w, \boldsymbol{\mu}_x, v(\boldsymbol{\Sigma}_x), \nu)$ e denote por $\boldsymbol{\theta}_{-a}$ o vetor $\boldsymbol{\theta}$ sem a componente a. Exceto para o grau de liberdade ν, as dccp das componentes de $\boldsymbol{\theta}$ também possuem forma fechada e são dadas por

$$
\begin{aligned}
\gamma_1 \mid \boldsymbol{\theta}_{-\gamma_1}, \mathbf{D}_c &\sim N_1\left(H_1\left\{\sigma_1^{-1}\mu_1 + \sigma_w^{-1}\sum_{i=1}^{n} g_i\left[(1-d_i)\eta_i + d_i y_i - \boldsymbol{\gamma}_2'\mathbf{x}_i - \boldsymbol{\gamma}_3'\mathbf{b}_i\right]\right\}, H_1\right), \\
\boldsymbol{\gamma}_2 \mid \boldsymbol{\theta}_{-\boldsymbol{\gamma}_2}, \mathbf{D}_c &\sim N_k\left(\mathbf{H}_2\left\{\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2 + \sigma_w^{-1}\sum_{i=1}^{n} g_i\left[(1-d_i)\eta_i + d_i y_i - \gamma_1 - \boldsymbol{\gamma}_3'\mathbf{b}_i\right]\mathbf{x}_i\right\}, \mathbf{H}_2\right), \\
\boldsymbol{\gamma}_3 \mid \boldsymbol{\theta}_{-\boldsymbol{\gamma}_3}, \mathbf{D}_c &\sim N_{k_1}\left(\mathbf{H}_3\left\{\boldsymbol{\Sigma}_3^{-1}\boldsymbol{\mu}_3 + \sigma_w^{-1}\sum_{i=1}^{n} g_i\left[(1-d_i)\eta_i + d_i y_i - \gamma_1 - \boldsymbol{\gamma}_2'\mathbf{x}_i\right]\mathbf{b}_i\right\}, \mathbf{H}_3\right), \\
\sigma_w \mid \boldsymbol{\theta}_{-\sigma_w}, \mathbf{D}_c &\sim IG\left(\alpha_w + \frac{n}{2}, \beta_w + \frac{1}{2}\sum_{i=1}^{n} g_i\left[(1-d_i)a_{1i}^2 + d_i a_{2i}^2\right]\right), \\
\boldsymbol{\mu}_x \mid \boldsymbol{\theta}_{-\boldsymbol{\mu}_x}, \mathbf{D}_c &\sim N_k\left(\mathbf{H}_4\left(\boldsymbol{\Sigma}_4^{-1}\boldsymbol{\mu}_4 + \boldsymbol{\Sigma}_x^{-1}\sum_{i=1}^{n} g_i\mathbf{x}_i\right), \mathbf{H}_4\right), \\
\boldsymbol{\Sigma}_x \mid \boldsymbol{\theta}_{-\boldsymbol{\Sigma}_x}, \mathbf{D}_c &\sim IW\left(m_1 + n, \boldsymbol{\Psi}_1 + \sum_{i=1}^{n} g_i(\mathbf{x}_i - \boldsymbol{\mu}_x)(\mathbf{x}_i - \boldsymbol{\mu}_x)'\right),
\end{aligned}
$$

em que

$$
\begin{aligned}
H_1^{-1} &= \sigma_1^{-1} + \sigma_w^{-1}\textstyle\sum_{i=1}^{n} g_i, \\
\mathbf{H}_2^{-1} &= \boldsymbol{\Sigma}_2^{-1} + \sigma_w^{-1}\textstyle\sum_{i=1}^{n} g_i\mathbf{x}_i\mathbf{x}_i', \\
\mathbf{H}_3^{-1} &= \boldsymbol{\Sigma}_3^{-1} + \sigma_w^{-1}\textstyle\sum_{i=1}^{n} g_i\mathbf{b}_i\mathbf{b}_i', \\
\mathbf{H}_4^{-1} &= \boldsymbol{\Sigma}_4^{-1} + \boldsymbol{\Sigma}_x^{-1}\textstyle\sum_{i=1}^{n} g_i.
\end{aligned}
$$

A dccp para o grau de liberdade ν não possui forma fechada e seu núcleo é o mesmo que foi obtido sob o modelo t-Tobit, ou seja,

$$
\pi(\nu \mid \boldsymbol{\theta}_{-\nu}, \mathbf{D}_c) \propto \left[\frac{(2/\nu)^{-\nu/2}}{\Gamma(\nu/2)}\right]^n \exp\left(-\nu\left[\frac{1}{2}\left(\sum_{i=1}^{n} g_i - \log(g_i)\right) + \frac{1}{\lambda}\right]\right),\ \nu > 2. \quad (3.18)
$$

Apesar do núcleo de $\pi(\nu \mid \boldsymbol{\theta}_{-\nu}, \mathbf{D}_c)$ em (3.18) ser o mesmo de (Capítulo 2, Seção 2.2.1, Equação 2.12), a inferência sobre o grau de liberdade ν não é a mesma em modelos t-Tobit com e sem erro.

A dccp de η_i dada em (3.16) segue de maneira direta da hipótese. Maiores detalhes a respeito de como as dccp são calculadas podem ser encontradas no Apêndice A. Como consequência, amostras das distribuições *a posteriori* de η_i, g_i e $\boldsymbol{\theta}_{-\nu}$ também são obtidas usando o amostrador de Gibbs. O algoritmo de ARMS também é utilizado para obter-se amostras da distribuição *a posteriori* de ν. Amostras das distribuições *a posteriori*

dos parâmetros no modelo original que são dados em (3.12) são obtidas diretamente das expressões em (3.9) substituindo-se os valores gerados de $\boldsymbol{\theta}$.

No próximo capítulo se discute aspectos de inferência clássica, sob o método de máxima verossimilhança, no modelo t-Tobit com erros nas covariáveis.

Capítulo 4

Alguns aspectos de inferência clássica no modelo t-Tobit com erro nas covariáveis

Um dos métodos mais utilizados para estimar parâmetros na escola clássica de Estatística é o método da máxima verossimilhança. Em modelos de regressão censurada, no entanto, a obtenção de formas fechadas para tais estimadores não é possível e, geralmente, métodos computacionais são usados para a sua obtenção. Por exemplo, para o modelo t-Tobit, Arellano-Valle *et al.* (2012) propõem um algoritmo do tipo EM para obter os estimadores de máxima verossimilhança (EMV) enquanto que, no modelo N-Tobit com erro de medida, Wang (1998) propõe um método de Newton-Raphson para aproximá-los.

O algoritmo EM, proposto por Dempster *et al.* (1977), possui várias características interessantes tais como estabilidade da convergência monótona com o aumento do número de iterações aumentando a verossimilhança e a simplicidade de sua implementação. No entanto, a estimação de máxima verossimilhança no modelo (3.13) é complexa e o algoritmo EM é menos aconselhável devido às dificuldades computacionais no passo M, como, por exemplo, encontrar o ponto de máximo em uma função de várias variáveis. Para lidar com esse problema, aplica-se uma extensão do algoritmo EM, chamada de algoritmo ECM (Meng e Rubin, 1993), que compartilha com o algoritmo EM várias características importantes como, por exemplo, o aumento da verossimilhança a cada iteração. Além disso, usualmente, ele possui uma taxa de convergência mais rápida que a do algoritmo EM. Essencialmente, o que difere o algoritmo ECM do algoritmo EM é o passo M que, no algoritmo ECM, ele é substituído por uma sequência de passos de maximizações condi-

cionais (CM) onde cada parâmetro é individualmente maximizado condicionalmente aos outros parâmetros os quais são fixados.

Com o objetivo de construir um algoritmo ECM para o modelo proposto, o fato de uma distribuição t-Student multivariada poder ser representada como uma mistura de escala da distribuição normal multivariada tende a facilitar a construção do algoritmo e, por isso, a representação dada na expressão (3.15) também será utilizado aqui (veja, por exemplo, Arellano-Valle e Bolfarine, 1995). Da representação hierárquica em (3.15), um algoritmo ECM para obter-se os estimadores de máxima verossimilhança (EMV) do vetor $\boldsymbol{\theta} = (\gamma_1, \boldsymbol{\gamma}_2, \boldsymbol{\gamma}_3, \sigma_w, \boldsymbol{\mu}_x, v(\boldsymbol{\Sigma}_x), \nu)$, pode ser implementado sem muita complexidade. A propriedade de invariância dos EMV e a condição de identificabilidade assumida permitem obter os EMV dos parâmetros originais. Para isso, basta considerar as expressões dadas em (3.9).

Na próxima seção o algoritmo ECM para aproximar os EMV para os parâmetros do modelo proposto será construído.

4.1 Algoritmo ECM para o modelo t-Tobit com erros nas covariáveis

Para construir o algoritmo ECM para o modelo proposto algumas notações adicionais são necessárias para descrever a log-verossimilhança completa obtida da expressão (3.15), do Capítulo 3.

Como descrito anteriormente seja $\mathbf{D}_c = (\mathbf{D}_o, \mathbf{D}_l)$ o conjunto de dados completos, onde $\mathbf{D}_o$ são os dados observados e $\mathbf{D}_l = (\mathbf{D}_{l1}, \ldots, \mathbf{D}_{ln})$, onde $\mathbf{D}_{li} = (\eta_i, g_i)$ são os dados latentes (dados faltantes hipotéticos). A função de log-verossimilhança completa para $\boldsymbol{\theta}$, induzida por (3.15) é $l^c(\boldsymbol{\theta} \mid \mathbf{D}_c) = \sum_{i=1}^n l_i^c(\boldsymbol{\theta} \mid \mathbf{D}_{oi}, \mathbf{D}_{li})$, onde

$$\begin{aligned} l_i^c(\boldsymbol{\theta} \mid \mathbf{D}_{oi}, \mathbf{D}_{li}) &= \log \phi_1(\eta_i \mid \mu(\mathbf{x}_i), g_i^{-1}\sigma_w) + \log \phi_k(\mathbf{x}_i \mid \boldsymbol{\mu}_x, g_i^{-1}\boldsymbol{\Sigma}_x) \\ &\quad - \log \Gamma\left(\frac{\nu}{2}\right) + \frac{\nu}{2}\log\left(\frac{\nu}{2}\right) + \frac{\nu}{2}\log g_i - \frac{\nu}{2} g_i \\ &\propto -\frac{1}{2}\log \sigma_w - \frac{1}{2\sigma_w} g_i S_{\eta_i} - \frac{1}{2}\log|\boldsymbol{\Sigma}_x| - \frac{1}{2} g_i (\mathbf{x}_i - \boldsymbol{\mu}_x)' \boldsymbol{\Sigma}_x^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_x) \\ &\quad - \log \Gamma\left(\frac{\nu}{2}\right) + \frac{1}{2}\nu \log\left(\frac{\nu}{2}\right) + \frac{\nu}{2}\log g_i - \frac{\nu}{2} g_i, \end{aligned} \tag{4.1}$$

$S_{\eta_i} = \eta_{0i}^2$, $\eta_{0i} = (\eta_i - \mu(\mathbf{x}_i))$ e $\mu(\mathbf{x}_i) = \gamma_1 + \boldsymbol{\gamma}_2' \mathbf{x}_i + \boldsymbol{\gamma}_3' \mathbf{b}_i$.

Dado o valor atual $\widehat{\boldsymbol{\theta}}^{(t)}$ de $\boldsymbol{\theta}$, que corresponde ao vetor formado pelos diferentes componentes de $\left(\widehat{\gamma}_1^{(t)}, \widehat{\boldsymbol{\gamma}}_2^{\prime(t)}, \widehat{\boldsymbol{\gamma}}_3^{(t)}, \widehat{\sigma}_w^{(t)}, \widehat{\boldsymbol{\mu}}_x^{(t)}, v(\widehat{\boldsymbol{\Sigma}}_x^{(t)}), \widehat{\nu}^{(t)}\right)$, o passo E do algoritmo calcula a esperança da função de log-verossimilhança dos dados completos condicional nos dados observados e em $\widehat{\boldsymbol{\theta}}^{(t)}$. Tal esperança condicional é denotada por $Q(\boldsymbol{\theta} \mid \widehat{\boldsymbol{\theta}}^{(t)}) = \sum_{i=1}^{n} Q_i(\boldsymbol{\theta} \mid \widehat{\boldsymbol{\theta}}^{(t)})$ e de (4.1) tem-se que

$$\begin{aligned}
Q_i(\boldsymbol{\theta} \mid \widehat{\boldsymbol{\theta}}^{(t)}) &= E(l_i^c(\boldsymbol{\theta} \mid \mathbf{D}_{oi}, \mathbf{D}_{oi}) \mid \mathbf{D}_{oi}, \widehat{\boldsymbol{\theta}}^{(t)}) \\
&= (1-d_i)E(l_i^c(\boldsymbol{\theta} \mid \mathbf{D}_{oi}, \mathbf{D}_{oi}) \mid \eta_i \leq 0, \mathbf{b}_i, \mathbf{x}_i, \widehat{\boldsymbol{\theta}}^{(t)}) \\
&\quad + d_i E(l_i^c(\boldsymbol{\theta} \mid \mathbf{D}_{oi}, \mathbf{D}_{oi}) \mid y_i, \mathbf{b}_i, \mathbf{x}_i, \widehat{\boldsymbol{\theta}}^{(t)}) \\
&\propto -\frac{1}{2}\log\sigma_w - \frac{1}{2\sigma_w}\widehat{(g_i S_{\eta_i})}^{(t)} - \frac{1}{2}\log|\boldsymbol{\Sigma}_x| - \frac{1}{2}\widehat{(g_i)}^{(t)}(\mathbf{x}_i - \boldsymbol{\mu}_x)'\boldsymbol{\Sigma}_x^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_x) \\
&\quad - \log\Gamma\left(\frac{\nu}{2}\right) + \frac{1}{2}\nu\log\left(\frac{\nu}{2}\right) + \frac{\nu}{2}\widehat{(\log g_i)}^{(t)} - \frac{\nu}{2}\widehat{(g_i)}^{(t)}, \qquad (4.2)
\end{aligned}$$

onde $\widehat{(g_i S_{\eta_i})}^{(t)} = \widehat{(g_i\eta_i^2)}^{(t)} - 2\mu(\mathbf{x}_i)\widehat{(g_i\eta_i)}^{(t)} + (\mu(\mathbf{x}_i))^2\widehat{(g_i)}^{(t)}$ e

$$\left.\begin{aligned}
\widehat{(g_i)}^{(t)} &= (1-d_i)E(g_i \mid \eta_i \leq 0, \mathbf{b}_i, \mathbf{x}_i, \widehat{\boldsymbol{\theta}}^{(t)}) + d_i E(g_i \mid y_i, \mathbf{b}_i, \mathbf{x}_i, \widehat{\boldsymbol{\theta}}^{(t)}), \\
\widehat{(g_i\eta_i)}^{(t)} &= (1-d_i)E(g_i\eta_i \mid \eta_i \leq 0, \mathbf{b}_i, \mathbf{x}_i, \widehat{\boldsymbol{\theta}}^{(t)}) + d_i y_i E(g_i \mid y_i, \mathbf{b}_i, \mathbf{x}_i, \widehat{\boldsymbol{\theta}}^{(t)}), \\
\widehat{(g_i\eta_i^2)}^{(t)} &= (1-d_i)E(g_i\eta_i^2 \mid \eta_i \leq 0, \mathbf{b}_i, \mathbf{x}_i, \widehat{\boldsymbol{\theta}}^{(t)}) + d_i y_i^2 E(g_i \mid y_i, \mathbf{b}_i, \mathbf{x}_i, \widehat{\boldsymbol{\theta}}^{(t)}), \\
\widehat{(\log g_i)}^{(t)} &= (1-d_i)E(\log g_i \mid \eta_i \leq 0, \mathbf{b}_i, \mathbf{x}_i, \widehat{\boldsymbol{\theta}}^{(t)}) + d_i E(\log g_i \mid y_i, b_i, \mathbf{x}_i, \widehat{\boldsymbol{\theta}}^{(t)}).
\end{aligned}\right\} \qquad (4.3)$$

Logo, a partir de (4.2) tem-se

$$\begin{aligned}
Q(\boldsymbol{\theta} \mid \widehat{\boldsymbol{\theta}}^{(t)}) \propto\ & -\frac{n}{2}\log\sigma_w - \frac{1}{2\sigma_w}\sum_{i=1}^{n}\widehat{(g_i S_{\eta_i})}^{(t)} - \frac{n}{2}\log|\boldsymbol{\Sigma}_x| - \frac{1}{2}\sum_{i=1}^{n}\widehat{(g_i)}^{(t)}(\mathbf{x}_i - \boldsymbol{\mu}_x)'\boldsymbol{\Sigma}_x^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_x) \\
& -n\log\Gamma\left(\frac{\nu}{2}\right) + \frac{n}{2}\nu\log\left(\frac{\nu}{2}\right) + \frac{\nu}{2}\sum_{i=1}^{n}\widehat{(\log g_i)}^{(t)} - \frac{\nu}{2}\sum_{i=1}^{n}\widehat{(g_i)}^{(t)}, \qquad (4.4)
\end{aligned}$$

As esperanças condicionais truncadas consideradas em (4.3) são sumarizadas no Lema 2 dado no Apêndice B.

Os passos CM (maximização condicional) maximizam a função $Q(\boldsymbol{\theta} \mid \widehat{\boldsymbol{\theta}}^{(t)})$ definida em (4.4) com respeito a $\boldsymbol{\theta}$, condicionalmente nos dados observados e em $\widehat{\boldsymbol{\theta}}^{(t)}$ e obtêm uma nova estimativa $\widehat{\boldsymbol{\theta}}^{(t+1)}$ da seguinte forma:

$$\widehat{\gamma}_1^{(t+1)} = \left(\sum_{i=1}^{n}\widehat{(g_i)}^{(t)}\right)^{-1}\left(\sum_{i=1}^{n}\widehat{(g_i\eta_i)}^{(t)} - \widehat{\boldsymbol{\gamma}}_2^{\prime(t)}\sum_{i=1}^{n}\widehat{(g_i)}^{(t)}\mathbf{x}_i - \widehat{\boldsymbol{\gamma}}_3^{\prime(t)}\sum_{i=1}^{n}\widehat{(g_i)}^{(t)}\mathbf{b}_i\right),$$

$$
\begin{aligned}
\widehat{\boldsymbol{\gamma}}_2^{(t+1)} &= \left(\sum_{i=1}^{n} \widehat{(g_i)}^{(t)} \mathbf{x}_i \mathbf{x}_i'\right)^{-1} \left(\sum_{i=1}^{n} \widehat{(g_i \eta_i)}^{(t)} \mathbf{x}_i - \widehat{\gamma}_1^{(t+1)} \sum_{i=1}^{n} \widehat{(g_i)}^{(t)} \mathbf{x}_i - \sum_{i=1}^{n} \widehat{(g_i)}^{(t)} \mathbf{x}_i \mathbf{b}_i' \widehat{\boldsymbol{\gamma}}_3^{(t)}\right), \\
\widehat{\boldsymbol{\gamma}}_3^{(t+1)} &= \left(\sum_{i=1}^{n} \widehat{(g_i)}^{(t)} \mathbf{b}_i \mathbf{b}_i'\right)^{-1} \left(\sum_{i=1}^{n} \widehat{(g_i \eta_i)}^{(t)} \mathbf{b}_i - \widehat{\gamma}_1^{(t+1)} \sum_{i=1}^{n} \widehat{(g_i)}^{(t)} \mathbf{b}_i - \sum_{i=1}^{n} \widehat{(g_i)}^{(t)} \mathbf{b}_i \mathbf{x}_i' \widehat{\boldsymbol{\gamma}}_2^{(t+1)}\right), \\
\widehat{\sigma}_w^{(t+1)} &= \frac{1}{n} \sum_{i=1}^{n} \widehat{(g_i \eta_i^2)}^{(t)} - \frac{2}{n} \sum_{i=1}^{n} \widehat{(g_i \eta_i)}^{(t)} \widehat{(\mu(\mathbf{x}_i))}^{(t+1)} + \frac{1}{n} \sum_{i=1}^{n} \widehat{(g_i)}^{(t)} \widehat{(\mu(\mathbf{x}_i))}^{2(t+1)}, \\
\widehat{\boldsymbol{\mu}}_x^{(t+1)} &= \left(\sum_{i=1}^{n} \widehat{(g_i)}^{(t)}\right)^{-1} \sum_{i=1}^{n} \widehat{(g_i)}^{(t)} \mathbf{x}_i, \\
\widehat{\boldsymbol{\Sigma}}_x^{(t+1)} &= \frac{1}{n} \sum_{i=1}^{n} \widehat{(g_i)}^{(t)} (\mathbf{x}_i - \widehat{\boldsymbol{\mu}}_x^{(t+1)})' (\widehat{\boldsymbol{\Sigma}}_x^{(t)})^{-1} (\mathbf{x}_i - \widehat{\boldsymbol{\mu}}_x^{(t+1)}),
\end{aligned}
$$

e

$$
\log\left(\frac{\widehat{\nu}^{(t+1)}}{2}\right) - \psi\left(\frac{\widehat{\nu}^{(t+1)}}{2}\right) = \frac{1}{n} \sum_{i=1}^{n} \widehat{(g_i)}^{(t)} - \frac{1}{n} \sum_{i=1}^{n} \widehat{(\log g_i)}^{(t)} - 1,
$$

onde $\widehat{(\mu(\mathbf{x}_i))}^{(t)} = \widehat{\gamma}_1^{(t)} + \widehat{\boldsymbol{\gamma}}_2^{\prime(t)} \mathbf{x}_i + \widehat{\boldsymbol{\gamma}}_3^{\prime(t)} \mathbf{b}_i$, $\widehat{(\mu(\mathbf{x}_i))}^{2(t+1)} = \left[\widehat{(\mu(\mathbf{x}_i))}^{(t+1)}\right]^2$ e $\psi(x) = \frac{\mathrm{d}}{\mathrm{d}x} \log \Gamma(x)$ denota a função digama. Maiores informações sobre a maximização de $Q(\boldsymbol{\theta} \mid \widehat{\boldsymbol{\theta}}^{(t)})$ podem ser encontradas no Apêndice C.

O algoritmo ECM é iterado até que a distância envolvendo duas avaliações sucessivas da log-verossimilhança dada em (3.14) seja suficientemente pequena, isto é,

$$
|\log f(\mathbf{D}_o \mid \widehat{\boldsymbol{\theta}}^{(t+1)}) / \log f(\mathbf{D}_o \mid \widehat{\boldsymbol{\theta}}^{(t)}) - 1| < \epsilon,
$$

onde $\epsilon \to 0$ é uma escolha arbitrária.

4.2 A matriz de informação observada do modelo t-Tobit com erros nas covariáveis

Nesta seção a matriz de informação observada para o modelo proposto t-Tobit com erros nas covariáveis será calculada. Com ela é possível obter-se os erros padrão assintóticos dos estimadores de máxima verossimilhança do modelo proposto. Primeiramente, constrói-se a matriz de informação observada para o modelo reduzido dado em (3.13) que depende dos parâmetros transformados $\boldsymbol{\theta} = (\gamma_1, \boldsymbol{\gamma}_2, \boldsymbol{\gamma}_3, \sigma_w, \boldsymbol{\mu}_x, v(\boldsymbol{\Sigma}_x), \nu)$.

Se $\mathbf{D}_o = (y, d, \mathbf{x}, \mathbf{b})$ denota um único dado observado sob o modelo proposto, então a contribuição de $\mathbf{D}_o$ para a função de log-verossimilhança, dada em (3.14), é

$$\begin{aligned}\log f(\mathbf{D}_o \mid \boldsymbol{\theta}) &= \log t_k(\mathbf{x} \mid \boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x, \nu) \\ &\quad + d \log t_1\left(y \mid \gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x} + \boldsymbol{\gamma}_3'\mathbf{b}, \frac{\sigma_w}{m_{(\mathbf{x},k,\nu)}}, \nu + k\right) \\ &\quad + (1-d) \log T_1\left(0 \mid \gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x} + \boldsymbol{\gamma}_3'\mathbf{b}, \frac{\sigma_w}{m_{(\mathbf{x},k,\nu)}}, \nu + k\right).\end{aligned} \tag{4.5}$$

Faça $y_0 = y - \mu(\mathbf{x})$, $\mathbf{x}_0 = \mathbf{x} - \boldsymbol{\mu}_x$ e $\mu(\mathbf{x}) = \gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x} + \boldsymbol{\gamma}_3'\mathbf{b}$. A partir da expressão (4.5), o esquema geral a ser desenvolvido é: (i) determina-se a matriz de informação observada de $(y_0, \sigma_w, \mathbf{x}_0, v(\boldsymbol{\Sigma}_x), \nu)$ a partir do segundo diferencial de cada um dos termos da expressão da log-verossimilhança em (4.5); (ii) obtém-se a matriz jacobiana da transformação de $(y_0, \sigma_w, \mathbf{x}_0, v(\boldsymbol{\Sigma}_x), \nu)$ para $\boldsymbol{\theta}$ e, a partir dela e da matriz obtida em (i), encontra-se a matriz de informação observada para $\boldsymbol{\theta}$; (iii) similarmente, obtém-se a matriz jacobiana de transformação de $\boldsymbol{\theta}$ em $(\beta_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \sigma_u, \boldsymbol{\mu}_\xi, v(\boldsymbol{\Sigma}_\xi), \nu)$, que gera a matriz de informação observada para os parâmetros originais do modelo proposto (3.12).

A matriz de informação observada de $(y_0, \sigma_w, \mathbf{x}_0, v(\boldsymbol{\Sigma}_x), \nu)$ é dada por

$$\boldsymbol{I}^o((y_0, \sigma_w, \mathbf{x}_0, v(\boldsymbol{\Sigma}_x), \nu)) = \boldsymbol{I}^x + d\boldsymbol{I}^y + (1-d)\boldsymbol{I}^0,$$

onde as matrizes $\boldsymbol{I}^x$, $\boldsymbol{I}^y$ e $\boldsymbol{I}^0$ são dadas por

$$\boldsymbol{I}^x = -\begin{pmatrix} 0 & 0 & \mathbf{0}' & \mathbf{0}' & 0 \\ 0 & 0 & \mathbf{0}' & \mathbf{0}' & 0 \\ \mathbf{0} & \mathbf{0} & \boldsymbol{I}^x_{\mathbf{x}_0\mathbf{x}_0} & \boldsymbol{I}^x_{\mathbf{x}_0 v(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^x_{\mathbf{x}_0\nu} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{I}^{x'}_{\mathbf{x}_0 v(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^x_{v(\boldsymbol{\Sigma}_x)v(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^x_{v(\boldsymbol{\Sigma}_x)\nu} \\ 0 & 0 & \boldsymbol{I}^{x'}_{\mathbf{x}_0\nu} & \boldsymbol{I}^{x'}_{v(\boldsymbol{\Sigma}_x)\nu} & I^x_{\nu\nu} \end{pmatrix},$$

$$\boldsymbol{I}^y = -\begin{pmatrix} I^y_{y_0y_0} & I^y_{y_0\sigma_w} & \boldsymbol{I}^{y'}_{y_0\mathbf{x}_0} & \boldsymbol{I}^{y'}_{y_0 v(\boldsymbol{\Sigma}_x)} & I^y_{y_0\nu} \\ I^y_{y_0\sigma_w} & I^y_{\sigma_w\sigma_w} & \boldsymbol{I}^{y'}_{\sigma_w\mathbf{x}_0} & \boldsymbol{I}^{y'}_{\sigma_w v(\boldsymbol{\Sigma}_x)} & I^y_{\sigma_w\nu} \\ \boldsymbol{I}^y_{y_0\mathbf{x}_0} & \boldsymbol{I}^y_{\sigma_w\mathbf{x}_0} & \boldsymbol{I}^y_{\mathbf{x}_0\mathbf{x}_0} & \boldsymbol{I}^y_{\mathbf{x}_0 v(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^y_{\mathbf{x}_0\nu} \\ \boldsymbol{I}^y_{y_0 v(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^y_{\sigma_w v(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^{y'}_{\mathbf{x}_0 v(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^y_{v(\boldsymbol{\Sigma}_x)v(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^y_{v(\boldsymbol{\Sigma}_x)\nu} \\ I^y_{y_0\nu} & I^y_{\sigma_w\nu} & \boldsymbol{I}^{y'}_{\mathbf{x}_0\nu} & \boldsymbol{I}^{y'}_{v(\boldsymbol{\Sigma}_x)\nu} & I^y_{\nu\nu} \end{pmatrix}$$

e

$$\boldsymbol{I}^0 = -\begin{pmatrix} I^0_{y_0y_0} & I^y_{y_0\sigma_w} & \boldsymbol{I}^{0'}_{y_0\mathbf{x}_0} & \boldsymbol{I}^{0'}_{y_0v(\boldsymbol{\Sigma}_x)} & I^0_{y_0\nu} \\ I^0_{y_0\sigma_w} & I^0_{\sigma_w\sigma_w} & \boldsymbol{I}^{0'}_{\sigma_w\mathbf{x}_0} & \boldsymbol{I}^{0'}_{\sigma_wv(\boldsymbol{\Sigma}_x)} & I^0_{\sigma_w\nu} \\ \boldsymbol{I}^0_{y_0\mathbf{x}_0} & \boldsymbol{I}^0_{\sigma_w\mathbf{x}_0} & \boldsymbol{I}^0_{\mathbf{x}_0\mathbf{x}_0} & \boldsymbol{I}^0_{\mathbf{x}_0v(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^0_{\mathbf{x}_0\nu} \\ \boldsymbol{I}^0_{y_0v(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^0_{\sigma_wv(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^{0'}_{\mathbf{x}_0v(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^0_{v(\boldsymbol{\Sigma}_x)v(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^0_{v(\boldsymbol{\Sigma}_x)\nu} \\ I^0_{y_0\nu} & I^0_{\sigma_w\nu} & \boldsymbol{I}^{0'}_{\mathbf{x}_0\nu} & \boldsymbol{I}^{0'}_{v(\boldsymbol{\Sigma}_x)\nu} & I^0_{\nu\nu} \end{pmatrix}.$$

As estradas dessas matrizes e maiores detalhes sobre os cálculos para obtê-las podem ser encontrados no Apêndice D.

A matriz jacobiana $\boldsymbol{J}_1$ associada à transformação de $(y_0, \sigma_w, \mathbf{x}_0, v(\boldsymbol{\Sigma}_x), \nu)$ para $\boldsymbol{\theta}$ é dada por

$$\boldsymbol{J}_1 = \begin{pmatrix} -1 & 0 & \mathbf{0}' & \mathbf{0}' & 0 \\ -\mathbf{x} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ -\mathbf{b} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ 0 & 1 & \mathbf{0}' & \mathbf{0}' & 0 \\ \mathbf{0} & \mathbf{0} & -\boldsymbol{I}_k & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{I}_{\frac{k(k+1)}{2}} & \mathbf{0} \\ 0 & 0 & \mathbf{0}' & \mathbf{0}' & 1 \end{pmatrix}.$$

Logo, a matriz de informação observada para $\boldsymbol{\theta}$, $\boldsymbol{I}^o(\boldsymbol{\theta})$, é

$$\boldsymbol{I}^o(\boldsymbol{\theta}) = \boldsymbol{J}_1\boldsymbol{I}^o((y_0, \sigma_w, \mathbf{x}_0, v(\boldsymbol{\Sigma}_x), \nu))\boldsymbol{J}_1'.$$

A matriz jacobiana $\boldsymbol{J}_2$ associada à transformação de $\boldsymbol{\theta}$ em $(\beta_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \sigma_u, \boldsymbol{\mu}_\xi, v(\boldsymbol{\Sigma}_\xi), \nu)$ é, por sua vez, dada por

$$\boldsymbol{J}_2 = \begin{pmatrix} 1 & \mathbf{0}' & \mathbf{0}' & 0 & \mathbf{0}' & \mathbf{0}' & 0 \\ (\boldsymbol{I}_k - \boldsymbol{\Lambda})\boldsymbol{\mu}_\xi & \boldsymbol{\Lambda}' & \mathbf{0} & [(\boldsymbol{I}_k - \boldsymbol{\Lambda})\boldsymbol{\Sigma}_\xi + \boldsymbol{\Sigma}_\xi(\boldsymbol{I}_k - \boldsymbol{\Lambda})']\boldsymbol{\beta}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{I}_l & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ 0 & \mathbf{0}' & \mathbf{0}' & 1 & \mathbf{0}' & \mathbf{0}' & 0 \\ (\boldsymbol{I}_k - \boldsymbol{\Lambda})'\boldsymbol{\beta}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{I}_k & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{D}'\operatorname{vec}(\boldsymbol{\beta}_2\boldsymbol{\beta}_2'(\boldsymbol{I}_k - \boldsymbol{\Lambda})) & \mathbf{0} & \mathbf{D}'(\boldsymbol{\Lambda}' \otimes \boldsymbol{I}_k)\mathbf{D}(\mathbf{D}'\mathbf{D})^{-1} & \mathbf{0} \\ 0 & \mathbf{0}' & \mathbf{0}' & 0 & \mathbf{0}' & \mathbf{0}' & 1 \end{pmatrix},$$

onde $\boldsymbol{\Lambda} = (\boldsymbol{I}_k + \boldsymbol{\Delta})^{-1}$ é a razão de confiabilidade e $\mathbf{D}$ é a matriz de duplicação com dimensão $k^2 \times k(k+1)/2$ tal que $\operatorname{vec}(\mathbf{A}) = \mathbf{D}v(\mathbf{A})$ para qualquer matriz simétrica $\mathbf{A}$ de dimensão $k \times k$. Portanto, a matriz de informação observada para o modelo proposto é

$$\boldsymbol{I}^o(\beta_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \sigma_u, \boldsymbol{\mu}_\xi, v(\boldsymbol{\Sigma}_\xi), \nu) = \boldsymbol{J}_2\boldsymbol{I}^o(\boldsymbol{\theta})\boldsymbol{J}_2'.$$

Pode-se observar das expressões anteriores que $\boldsymbol{J}_2$ não depende do dado observado. Logo, considerando uma amostra de n indivíduos, a matriz de informação observada para o

modelo proposto é

$$\boldsymbol{I}^o(\beta_1, \boldsymbol{\beta}_2, \boldsymbol{\beta}_3, \sigma_u, \boldsymbol{\mu}_\xi, v(\boldsymbol{\Sigma}_\xi), \nu) = \boldsymbol{J}_2 \left(\sum_{i=1}^{n} \boldsymbol{J}_{1i} \boldsymbol{I}^o((y_{0i}, \sigma_w, \mathbf{x}_{0i}, v(\boldsymbol{\Sigma}_x), \nu)) \boldsymbol{J}_{1i}' \right) \boldsymbol{J}_2'.$$

Esse resultado será utilizado nos Capítulos 5 e 6 na construção de intervalos de confiança para os parâmetros do modelo proposto.

Capítulo 5

Estudos Monte Carlo

Nesse capítulo são realizados estudos Monte Carlo para avaliar os estimadores bayesianos média e mediana *a posteriori* obtidos usando o modelo proposto e compará-los àqueles obtidos usando os modelos tipo Tobit previamente introduzidos na literatura, Tobit, t-Tobit e N-Tobit com erros nas covariáveis. Também será feito um estudo Monte Carlo para avaliar o desempenho dos estimadores de máxima verossimilhança, obtidos pelo algoritmo ECM, introduzidos no Capítulo 4. Diferentes cenários serão considerados em ambas as análises. Em cada cenário deseja-se avaliar o efeito:

Cenário 1: De diferentes proporções de censuras.

Cenário 2: De diferentes tamanhos amostrais.

Cenário 3: Da má especificação da condição de identificabilidade $\mathbf{\Delta}$.

Cenário 4: Do grau de liberdade considerado na geração dos dados.

Cenário 5: Da má especificação do grau de liberdade ν.

Em modelos que envolvem a distribuição t-Student tem sido comum a prática de fixar uma grade de valores para o grau de liberdade ν e utilizar algum critério para selecionar o melhor modelo. A defesa para tal estratégia é que pode-se perder robustez sempre que ν é estimado. No entanto, como o processo de estimação pode ser computacionalmente caro, esse procedimento pode ser inviável para uma grade muito fina. Por isso, diferentes graus de liberdade, isto é, a robustez do modelo, também é levada em consideração nos estudos Monte Carlo. No Cenário 4, os conjuntos de dados são gerados assumindo diferentes valores para o grau de liberdade e, assim como para os cenários de 1 a 3, ν também é um

parâmetro a ser estimado. No Cenário 5 considera-se que ν é conhecido e fixo no modelo e o objetivo é avaliar como as estimativas são afetadas se o valor de ν é mal especificado *a priori*. No caso bayesiano, os dados são gerados sob o modelo proposto e como subproduto é feita uma investigação do efeito nas estimativas da escolha equivocada de um dos modelos já propostos na literatura.

Como o objetivo dos estudos é avaliar o comportamento dos estimadores não houve muita preocupação com a escolha prática do valor de $\boldsymbol{\Delta}$, como mencionado no Capítulo 3. Devido a isso, em alguns casos foram obtidos valores de σ_u negativos.

5.1 Caso 1: Avaliando os estimadores bayesianos

Em todos os cenários são consideradas 500 réplicas do modelo proposto com $k_1 = k = 2$, isto é, considera-se duas covariáveis medidas com erro e duas covariáveis medidas sem erro, em que se assume $\beta_1 = -6$, $\boldsymbol{\beta}_2 = (\boldsymbol{\beta}_{21}, \boldsymbol{\beta}_{22})' = (0,6; -0,3)'$, $\boldsymbol{\beta}_3 = (\boldsymbol{\beta}_{31}, \boldsymbol{\beta}_{32})' = (0,5; 2)'$, $\sigma_u = 18$, $\boldsymbol{\Sigma}_\xi = \begin{pmatrix} \boldsymbol{\Sigma}_{\xi 11} & \boldsymbol{\Sigma}_{\xi 12} \\ \boldsymbol{\Sigma}_{\xi 12} & \boldsymbol{\Sigma}_{\xi 22} \end{pmatrix} = \begin{pmatrix} 180 & -0,1 \\ -0,1 & 50 \end{pmatrix}$, $\boldsymbol{\mu}_\xi = (20; 1)'$, e $\nu = 5$. Assume-se também que $\boldsymbol{\Delta} = 0,1\boldsymbol{I}_2$. Logo, $\boldsymbol{\Sigma}_v = 0,1\boldsymbol{\Sigma}_\xi$. Para as covariáveis medidas sem erro $\mathbf{b}_i = (\mathbf{b}_{1,i}; \mathbf{b}_{2,i})'$ considera-se que $\mathbf{b}_{1,i} \sim N(10, 3^2)$ e $\mathbf{b}_{2,i} \sim Bernoulli(0,8)$, $i = 1, \ldots, n$.

No Cenário 1 as amostras são de tamanho $n = 200$ e os dados são gerados assumindo 5%, 25% e 45% de censura. No Cenário 2 a proporção de censura é fixada em aproximadamente 12,26% e as amostras possuem tamanho $n = 50$, 200 e 500. As amostras no Cenário 3 são de tamanho $n = 200$ e com aproximadamente 12,26% de censura. Nesse caso, os modelos com erros nas covariáveis que assumem as distribuições normal e t-Student são ajustados assumindo $\boldsymbol{\Delta} = 0,001\boldsymbol{I}_2$, $0,05\boldsymbol{I}_2$, $0,1\boldsymbol{I}_2$, $0,15\boldsymbol{I}_2$ e $0,3\boldsymbol{I}_2$. Nos Cenários 4 e 5, as amostras possuem tamanho $n = 200$ e os dados são gerados do modelo proposto assumindo $\nu = 2,01$, 5 e ∞. Tais valores também são assumidos como especificações *a priori* para ν no Cenário 5.

Distribuições *a priori* pouco informativas são consideradas para todos os parâmetros assumindo distribuições *a priori* com grandes variâncias. Assume-se as seguintes distribuições *a priori* para os modelos t-Tobit e N-Tobit ambos com erros nas covariáveis:

$\gamma_1 \sim N_1(0, 10^6)$, $\boldsymbol{\gamma}_2 \stackrel{d}{=} \boldsymbol{\gamma}_3 \stackrel{d}{=} \boldsymbol{\mu}_x \sim N_2\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}10^6 & 10^3\\10^3 & 10^6\end{pmatrix}\right)$, $\sigma_w \sim IG(2,0001; 10,001)$ e $\boldsymbol{\Sigma}_x \sim IW\left(6, \begin{pmatrix}(45\times10^5)^{1/2} & (45\times10^5)^{1/4}\\(45\times10^5)^{1/4} & (45\times10^5)^{1/2}\end{pmatrix}\right)$. Como consequência, as distribuições *a priori* para os parâmetros de regressão nos modelos com erros nas covariáveis t-Tobit e N-Tobit originais são

$$\begin{aligned}\beta_1 \mid \boldsymbol{\beta}_2 &\sim N_1\left(0, 10^6 + \boldsymbol{\beta}_2'(\boldsymbol{I}_2+\boldsymbol{\Delta})^{-1}\boldsymbol{\Delta}\begin{pmatrix}10^6 & 10^3\\10^3 & 10^6\end{pmatrix}\boldsymbol{\Delta}(\boldsymbol{I}_2+\boldsymbol{\Delta})^{-1}\boldsymbol{\beta}_2\right),\\ \boldsymbol{\beta}_2 &\sim N_2\left(\begin{pmatrix}0\\0\end{pmatrix}, (\mathbf{I}_2+\boldsymbol{\Delta})\begin{pmatrix}10^6 & 10^3\\10^3 & 10^6\end{pmatrix}(\mathbf{I}_2+\boldsymbol{\Delta})'\right).\end{aligned}$$

Para os modelos t-Tobit e Tobit assume-se que $\beta_1 \sim N_1(0, 10^6)$, $\boldsymbol{\beta}_2 = (\boldsymbol{\beta}_{21}, \boldsymbol{\beta}_{22}, \boldsymbol{\beta}_{31}, \boldsymbol{\beta}_{32})' \sim N_4\left(\begin{pmatrix}0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}10^6 & 10^3 & 10^3 & 10^3\\10^3 & 10^6 & 10^3 & 10^3\\10^3 & 10^3 & 10^6 & 10^3\\10^3 & 10^3 & 10^3 & 10^6\end{pmatrix}\right)$ e $\sigma_u \sim \text{IG}(2,0001, 10,001)$. Para os modelos t-Tobit e t-Tobit com erros nas covariáveis também é assumido que $\nu \sim TE(10^3; 2)$.

Para o MCMC, com o objetivo de obter uma amostra final de tamanho 150, uma cadeia de tamanho 6.491 é executada. Após a convergência ser alcançada, descarta-se os primeiros 5.000 valores como período de *burn-in* e considera-se um *lag* de 10 passos para obter uma amostra não correlacionada vinda da distribuição *a posteriori*.

Neste estudo também são comparados os resultados obtidos pelo modelo proposto (tce) com os modelos Tobit normal (nse), t-Tobit (tse) e N-Tobit com erros nas covariáveis (tce). Em cada réplica Monte Carlo calculou-se as estatísticas LPML e DIC. O logaritmo da verossimilhança pseudomarginal (LPML) é uma estatística resumo das CPO_i's (ordenadas preditivas condicionais). A CPO (Carlin e Louis, 2008) é uma das estatísticas mais usadas como critério de seleção e é derivada da distribuição preditiva *a posteriori*. Para o modelo proposto, não há forma fechada para a CPO_i. No entanto, uma estimativa Monte Carlo da CPO_i pode ser obtida usando uma única amostra MCMC da distribuição *a posteriori* conjunta dos parâmetros do modelo proposto a partir de uma aproximação por média harmônica (Dey *et al.*, 1997). Valores maiores de LPML indicam melhores ajustes. O '*deviance information criterion*' (DIC) foi proposto por Spiegelhater

et al (2002) e é baseado na média *a posteriori* da '*deviance*', que é estimada a partir da função de verossimilhança e de amostras da distribuição *a posteriori* conjunta dos parâmetros do modelo proposto. Valores menores para o DIC indicam melhores ajustes.

5.1.1 O efeito de diferentes proporções de censura nos estimadores bayesianos - Cenário 1

As Tabelas 5.1 e 5.2 apresentam o valor mediano e o erro quadrático médio (mse) para as médias e medianas *a posteriori*, respectivamente, sob o modelo t-Tobit com erro nas covariáveis, N-Tobit com erros nas covariáveis, t-Tobit e Tobit sempre que os dados são gerados assumindo 5%, 25% e 45% de censura. A mediana foi escolhida para representar as estimativas típicas pois são menos susceptíveis a valores atípicos.

Pode-se observar das Tabelas 5.1 e 5.2 que, de maneira geral, melhores estimativas, para cada modelo, são obtidas para os casos onde há menor proporção de censuras. Maiores diferenças em relação aos verdadeiros valores dos parâmetros são obtidas quando ajustam-se às amostras geradas com erros nas covariáveis os modelos que não os consideram. Para os modelos com erros nas covariáveis os parâmetros com piores estimativas são os de escala σ_u e $\mathbf{\Sigma}_\xi$, enquanto os com as melhores estimativas são os relacionados às covariáveis com erro, $\boldsymbol{\beta}_2$. Nos modelos ajustados sem erros nas covariáveis as melhores estimativas foram obtidas para os parâmetros associados às covariáveis $\boldsymbol{\beta}_2$ e $\boldsymbol{\beta}_3$, enquanto as piores estimações foram observadas para o parâmetro de escala σ_u e no intercepto β_1, especialmente no modelo t-Tobit quando a censura observada é de apenas 5%. Em relação ao grau de liberdade ν observa-se valores superestimados para todos os casos, principalmente no modelo t-Tobit. As melhores estimativas foram obtidas quando são ajustados aos dados o modelo proposto. Tanto a média quanto a mediana *a posteriori* fornecem estimativas e resultados próximos aos valores reais dos parâmetros utilizados para geração dos dados.

A Tabela 5.3 apresenta a mediana dos desvios padrão *a posteriori* dos parâmetros dos modelos ajustados em função da proporção de censura na amostra.

Em relação à variabilidade *a posteriori*, de acordo com a Tabela 5.3, observa-se um aumento no desvio-padrão das distribuições *a posteriori* dos parâmetros à medida que a porcentagem de censuras aumenta. Tanto para os modelos com erros nas covariáveis

Tabela 5.1: Mediana e erro quadrático médio (em parênteses) para as médias *a posteriori* dos parâmetros dos modelos t-Tobit com erro nas covariáveis (tce), Tobit normal com erro nas covariáveis (nce), t-Tobit (tse) e Tobit normal (nse) em função da proporção de censura na amostra.

Modelo			tce	nce	tse	nse
Censura	Parâmetro	Real	mediana mse	mediana mse	mediana mse	mediana mse
5%	β_1	-6	-6,078 $(3,436)$	-5,803 $(4,527)$	-4,838 $(9,762 \times 10^4)$	-4,759 $(5,863)$
	$\boldsymbol{\beta}_{21}$	0,6	0,603 $(1,100 \times 10^{-3})$	0,596 $(1,800 \times 10^{-3})$	0,542 $(4,291 \times 10^{-3})$	0,541 $(4,885 \times 10^{-3})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,299 $(3,500 \times 10^{-3})$	-0,299 $(5,800 \times 10^{-3})$	-0,270 $(4,393 \times 10^{-3})$	-0,269 $(5,714 \times 10^{-3})$
	$\boldsymbol{\beta}_{31}$	0,5	0,483 $(1,990 \times 10^{-2})$	0,489 $(2,680 \times 10^{-2})$	0,490 $(2,337 \times 10^{-2})$	0,496 $(2,689 \times 10^{-2})$
	$\boldsymbol{\beta}_{32}$	2	2,042 $(1,018)$	2,058 $(1,198)$	2,180 $(9,768 \times 10^4)$	2,067 $(1,202)$
	σ_u	18	18,937 $(1,208 \times 10^1)$	26,338 $(1,311 \times 10^2)$	32,358 $(1,340 \times 10^3)$	36,979 $(4,379 \times 10^2)$
	$\boldsymbol{\mu}_{\xi_1}$	20	20,096 $(1,323)$	19,996 $(1,675)$		
	$\boldsymbol{\mu}_{\xi_2}$	1	1,036 $(3,568 \times 10^{-1})$	1,034 $(4,407 \times 10^{-1})$		
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	200,361 $(1,040 \times 10^3)$	296,205 $(1,733 \times 10^4)$		
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,743 $(6,179 \times 10^1)$	0,670 $(3,101 \times 10^2)$		
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	64,382 $(2,616 \times 10^2)$	89,945 $(2,945 \times 10^3)$		
	ν	5	6,149 $(5,426)$		14,710 $(9,990 \times 10^1)$	
25%	β_1	-6	-5,930 $(4,184)$	-6,049 $(5,566)$	-4,883 $(5,834)$	-4,881 $(6,531)$
	$\boldsymbol{\beta}_{21}$	0,6	0,603 $(1,462 \times 10^{-3})$	0,600 $(2,507 \times 10^{-3})$	0,543 $(4,528 \times 10^{-3})$	0,544 $(5,021 \times 10^{-3})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,298 $(4,164 \times 10^{-3})$	-0,301 $(6,838 \times 10^{-3})$	-0,272 $(5,305 \times 10^{-3})$	-0,273 $(6,725 \times 10^{-3})$
	$\boldsymbol{\beta}_{31}$	0,5	0,487 $(2,332 \times 10^{-2})$	0,505 $(3,047 \times 10^{-2})$	0,495 $(2,712 \times 10^{-2})$	0,509 $(3,116 \times 10^{-2})$
	$\boldsymbol{\beta}_{32}$	2	2,066 $(1,134)$	2,114 $(1,343)$	2,155 $(1,212)$	2,085 $(1,373)$
	σ_u	18	18,916 $(1,481 \times 10^1)$	25,947 $(1,313 \times 10^2)$	31,932 $(2,321 \times 10^2)$	36,276 $(4,388 \times 10^2)$
	$\boldsymbol{\mu}_{\xi_1}$	20	20,042 $(1,328)$	20,008 $(1,673)$		
	$\boldsymbol{\mu}_{\xi_2}$	1	1,057 $(3,582 \times 10^{-1})$	1,019 $(4,431 \times 10^{-1})$		
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	200,731 $(1,092 \times 10^3)$	296,372 $(1,752 \times 10^4)$		
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,798 $(6,238 \times 10^1)$	0,876 $(3,115 \times 10^2)$		
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	64,549 $(2,711 \times 10^2)$	90,130 $(2,939 \times 10^3)$		
	ν	5	6,334 $(6,218)$		15,657 $(1,174 \times 10^2)$	
45%	β_1	-6	-6,400 $(6,478)$	-6,867 $(9,946)$	-5,516 $(7,538)$	-5,768 $(8,495)$
	$\boldsymbol{\beta}_{21}$	0,6	0,609 $(2,274 \times 10^{-3})$	0,616 $(4,101 \times 10^{-3})$	0,556 $(4,337 \times 10^{-3})$	0,559 $(4,783 \times 10^{-3})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,297 $(5,509 \times 10^{-3})$	-0,304 $(8,804 \times 10^{-3})$	-0,274 $(6,543 \times 10^{-3})$	-0,279 $(7,967 \times 10^{-3})$
	$\boldsymbol{\beta}_{31}$	0,5	0,492 $(2,977 \times 10^{-2})$	0,512 $(4,061 \times 10^{-2})$	0,517 $(3,514 \times 10^{-2})$	0,521 $(4,071 \times 10^{-2})$
	$\boldsymbol{\beta}_{32}$	2	2,057 $(1,414)$	2,229 $(1,794)$	2,249 $(1,584)$	2,209 $(1,784)$
	σ_u	18	18,996 $(2,172 \times 10^1)$	27,178 $(1,905 \times 10^2)$	33,472 $(3,033 \times 10^2)$	38,245 $(5,735 \times 10^2)$
	$\boldsymbol{\mu}_{\xi_1}$	20	20,010 $(1,323)$	20,030 $(1,681)$		
	$\boldsymbol{\mu}_{\xi_2}$	1	1,049 $(3,597 \times 10^{-1})$	1,008 $(4,414 \times 10^{-1})$		
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	202,580 $(1,135 \times 10^3)$	294,933 $(1,743 \times 10^4)$		
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,673 $(6,334 \times 10^1)$	1,006 $(3,104 \times 10^2)$		
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	64,911 $(2,768 \times 10^2)$	89,919 $(2,953 \times 10^3)$		
	ν	5	6,408 $(6,719)$		16,493 $(1,314 \times 10^2)$	

Tabela 5.2: Mediana e erro quadrático médio (em parênteses) para as medianas *a posteriori* dos parâmetros dos modelos t-Tobit com erro nas covariáveis (tce), Tobit normal com erro nas covariáveis (nce), t-Tobit (tse) e Tobit normal (nse) em função da proporção de censura na amostra.

Modelo			tce	nce	tse	nse
Censura	Parâmetro	Real	mediana mse	mediana mse	mediana mse	mediana mse
5%	β_1	-6	-6,058 $(3,463)$	-5,792 $(4,556)$	-4,865 $(9,751 \times 10^4)$	-4,789 $(5,902)$
	$\boldsymbol{\beta}_{21}$	0,6	0,603 $(1,090 \times 10^{-3})$	0,595 $(1,772 \times 10^{-3})$	0,542 $(4,316 \times 10^{-3})$	0,541 $(4,883 \times 10^{-3})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,300 $(3,464 \times 10^{-3})$	-0,298 $(5,777 \times 10^{-3})$	-0,270 $(4,354 \times 10^{-3})$	-0,269 $(5,730 \times 10^{-3})$
	$\boldsymbol{\beta}_{31}$	0,5	0,487 $(2,008 \times 10^{-2})$	0,490 $(2,679 \times 10^{-2})$	0,490 $(2,321 \times 10^{-2})$	0,496 $(2,705 \times 10^{-2})$
	$\boldsymbol{\beta}_{32}$	2	2,070 $(1,026)$	2,069 $(1,201)$	2,156 $(9,759 \times 10^4)$	2,062 $(1,201)$
	σ_u	18	18,778 $(1,149 \times 10^1)$	26,198 $(1,264 \times 10^2)$	32,035 $(1,127 \times 10^3)$	36,709 $(4,266 \times 10^2)$
	$\boldsymbol{\mu}_{\xi_1}$	20	20,090 $(1,316)$	19,980 $(1,678)$		
	$\boldsymbol{\mu}_{\xi_2}$	1	1,031 $(3,577 \times 10^{-1})$	1,050 $(4,438 \times 10^{-1})$		
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	198,355 $(9,653 \times 10^2)$	294,260 $(1,682 \times 10^4)$		
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,850 $(6,047 \times 10^1)$	0,619 $(3,040 \times 10^2)$		
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	63,977 $(2,472 \times 10^2)$	89,295 $(2,888 \times 10^3)$		
	ν	5	5,922 $(4,147)$		13,539 $(8,016 \times 10^1)$	
25%	β_1	-6	-5,988 $(4,161)$	-6,051 $(5,601)$	-4,836 $(5,944)$	-4,815 $(6,618)$
	$\boldsymbol{\beta}_{21}$	0,6	0,603 $(1,470 \times 10^{-3})$	0,601 $(2,496 \times 10^{-3})$	0,542 $(4,573 \times 10^{-3})$	0,544 $(5,073 \times 10^{-3})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,298 $(4,162 \times 10^{-3})$	-0,302 $(6,829 \times 10^{-3})$	-0,272 $(5,337 \times 10^{-3})$	-0,274 $(6,729 \times 10^{-3})$
	$\boldsymbol{\beta}_{31}$	0,5	0,485 $(2,323 \times 10^{-2})$	0,500 $(3,058 \times 10^{-2})$	0,496 $(2,729 \times 10^{-2})$	0,503 $(3,140 \times 10^{-2})$
	$\boldsymbol{\beta}_{32}$	2	2,057 $(1,132)$	2,120 $(1,342)$	2,135 $(1,211)$	2,051 $(1,376)$
	σ_u	18	18,647 $(1,407 \times 10^1)$	25,499 $(1,244 \times 10^2)$	31,650 $(2,221 \times 10^2)$	35,758 $(4,247 \times 10^2)$
	$\boldsymbol{\mu}_{\xi_1}$	20	20,036 $(1,337)$	20,010 $(1,673)$		
	$\boldsymbol{\mu}_{\xi_2}$	1	1,048 $(3,593 \times 10^{-1})$	1,015 $(4,467 \times 10^{-1})$		
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	200,079 $(1,005 \times 10^3)$	294,225 $(1,698 \times 10^4)$		
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,640 $(6,131 \times 10^1)$	0,731 $(3,058 \times 10^2)$		
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	64,130 $(2,551 \times 10^2)$	89,666 $(2,856 \times 10^3)$		
	ν	5	6,100 $(4,783)$		14,468 $(9,480 \times 10^1)$	
45%	β_1	-6	-6,435 $(6,416)$	-6,862 $(9,860)$	-5,429 $(7,611)$	-5,773 $(8,464)$
	$\boldsymbol{\beta}_{21}$	0,6	0,608 $(2,274 \times 10^{-3})$	0,615 $(4,067 \times 10^{-3})$	0,556 $(4,413 \times 10^{-3})$	0,557 $(4,833 \times 10^{-3})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,296 $(5,483 \times 10^{-3})$	-0,303 $(8,760 \times 10^{-3})$	-0,273 $(6,574 \times 10^{-3})$	-0,278 $(7,954 \times 10^{-3})$
	$\boldsymbol{\beta}_{31}$	0,5	0,494 $(2,949 \times 10^{-2})$	0,518 $(4,011 \times 10^{-2})$	0,514 $(3,527 \times 10^{-2})$	0,522 $(4,084 \times 10^{-2})$
	$\boldsymbol{\beta}_{32}$	2	2,050 $(1,423)$	2,250 $(1,789)$	2,235 $(1,576)$	2,187 $(1,795)$
	σ_u	18	18,630 $(2,008 \times 10^1)$	26,900 $(1,802 \times 10^2)$	33,004 $(2,878 \times 10^2)$	37,890 $(5,485 \times 10^2)$
	$\boldsymbol{\mu}_{\xi_1}$	20	19,999 $(1,329)$	20,050 $(1,690)$		
	$\boldsymbol{\mu}_{\xi_2}$	1	1,060 $(3,622 \times 10^{-1})$	1,009 $(4,439 \times 10^{-1})$		
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	200,230 $(1,054 \times 10^3)$	292,820 $(1,693 \times 10^4)$		
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,766 $(6,126 \times 10^1)$	1,161 $(3,052 \times 10^2)$		
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	64,383 $(2,608 \times 10^2)$	89,084 $(2,897 \times 10^3)$		
	ν	5	6,152 $(5,188)$		15,265 $(1,066 \times 10^2)$	

Tabela 5.3: Mediana para os desvios-padrão *a posteriori* dos parâmetros dos modelos t-Tobit com erro nas covariáveis (tce), Tobit normal com erro nas covariáveis (nce), t-Tobit (tse) e Tobit normal (nse) em função da proporção de censura na amostra.

Modelo		tce	nce	tse	nse
Censura	Estatística	mediana	mediana	mediana	mediana
5%	β_1	1,824	1,963	1,851	1,940
	$\boldsymbol{\beta}_{21}$	0,032	0,029	0,026	0,026
	$\boldsymbol{\beta}_{22}$	0,058	0,052	0,048	0,047
	$\boldsymbol{\beta}_{31}$	0,143	0,156	0,147	0,155
	$\boldsymbol{\beta}_{32}$	0,965	1,043	1,012	1,053
	σ_u	3,365	4,075	3,887	3,828
	$\boldsymbol{\mu}_{\xi_1}$	1,151	1,276		
	$\boldsymbol{\mu}_{\xi_2}$	0,645	0,700		
	$\boldsymbol{\Sigma}_{\xi_{11}}$	25,124	29,733		
	$\boldsymbol{\Sigma}_{\xi_{12}}$	8,785	11,515		
	$\boldsymbol{\Sigma}_{\xi_{22}}$	7,682	8,925		
	ν	1,436		5,784	
25%	β_1	1,991	2,173	2,048	2,143
	$\boldsymbol{\beta}_{21}$	0,037	0,034	0,031	0,031
	$\boldsymbol{\beta}_{22}$	0,061	0,056	0,052	0,051
	$\boldsymbol{\beta}_{31}$	0,148	0,164	0,157	0,165
	$\boldsymbol{\beta}_{32}$	1,017	1,114	1,075	1,122
	σ_u	3,657	4,429	4,307	4,329
	$\boldsymbol{\mu}_{\xi_1}$	1,145	1,279		
	$\boldsymbol{\mu}_{\xi_2}$	0,643	0,700		
	$\boldsymbol{\Sigma}_{\xi_{11}}$	25,223	29,545		
	$\boldsymbol{\Sigma}_{\xi_{12}}$	8,804	11,527		
	$\boldsymbol{\Sigma}_{\xi_{22}}$	7,710	9,016		
	ν	1,535		6,175	
45%	β_1	2,388	2,590	2,436	2,542
	$\boldsymbol{\beta}_{21}$	0,045	0,041	0,038	0,038
	$\boldsymbol{\beta}_{22}$	0,067	0,063	0,057	0,057
	$\boldsymbol{\beta}_{31}$	0,164	0,185	0,177	0,185
	$\boldsymbol{\beta}_{32}$	1,151	1,273	1,227	1,281
	σ_u	4,221	5,289	5,287	5,370
	$\boldsymbol{\mu}_{\xi_1}$	1,149	1,269		
	$\boldsymbol{\mu}_{\xi_2}$	0,644	0,699		
	$\boldsymbol{\Sigma}_{\xi_{11}}$	25,474	29,296		
	$\boldsymbol{\Sigma}_{\xi_{12}}$	8,758	11,449		
	$\boldsymbol{\Sigma}_{\xi_{22}}$	7,807	9,012		
	ν	1,592		6,598	

quanto nos modelos sem erros as maiores variabilidades *a posteriori* ocorrem nas distribuições dos parâmetros de escala. Entre os modelos que consideram erros de medida nas covariáveis as distribuições *a posteriori* com menores desvios-padrão ocorrem no modelo proposto e entre os modelos sem erros nas covariáveis as maiores precisões *a posteriori* estão no modelo t-Tobit. De maneira geral os desvios-padrão *a posteriori* de cada parâmetro não se diferenciam tanto entre os modelos ajustados, a não ser para o grau de liberdade ν, que tem menor variância no modelo proposto.

A Tabela 5.4 apresenta as estatísticas DIC (*deviance information criterion*) e a soma dos logaritmos das ordenadas da densidade preditiva condicional (LPML) para cada modelo ajustado em função da proporção de censura presente nas amostras.

Tabela 5.4: Medianas e médias dos valores obtidos de LPML e DIC para os modelos t-Tobit com erro nas covariáveis (tce), Tobit normal com erro nas covariáveis (nce), t-Tobit (tse) e Tobit normal (nse) em função da proporção de censura na amostra.

Modelo		tce		nce		tse		nse	
Censura	Estatística	mediana	média	mediana	média	mediana	média	mediana	média
5%	LPML	-2195,483	-2196,572	-2224,329	-2227,633	-621,593	-626,786	-625,447	-626,553
	DIC	4390,388	4392,682	4442,904	4448,858	1242,054	1251,774	1248,653	1250,264
25%	LPML	-2086,784	-2087,781	-2113,786	-2116,930	-511,211	-511,250	-514,341	-515,858
	DIC	4173,292	4175,057	4220,678	4227,419	1020,905	1021,027	1026,942	1028,813
45%	LPML	-1972,489	-1973,387	-1998,104	-2000,630	-396,251	-395,996	-399,146	-399,618
	DIC	3944,297	3946,163	3991,443	3994,912	790,943	790,423	796,051	796,364

De acordo com a Tabela 5.4 nota-se que o aumento na proporção de censura resulta em um aumento (decrescimento) nas medidas LPML (DIC). Pode-se ver também que, dentre os modelos com erros nas covariáveis, as medidas LPML e DIC indicam que o melhor modelo é o modelo proposto. Porém, levando em conta os modelos ajustados sem erros nas covariáveis, tais medidas mostram que esses modelos são os melhores. No entanto, como observa-se nas Tabelas 5.1 e 5.2, as melhores estimativas são obtidas no modelo proposto e, portanto, se há evidências de que covariáveis podem ter sido medidas com erros é melhor considerar os modelos que levam isso em conta. Também pode-se notar que se o modelo não é bem especificado, apesar de obter estimativas ruins para σ_u, obtém-se também distribuições *a posteriori* para σ_u com alta variância (veja Tabela 5.3). Isso também pode ser uma boa ferramenta auxiliar na seleção do modelo.

5.1.2 O efeito de diferentes tamanhos amostrais nos estimadores bayesianos - Cenário 2

As Tabelas 5.5 e 5.6 apresentam as medianas e os erros quadráticos médio para a média e mediana *a posteriori*, respectivamente, dos parâmetros dos modelos ajustados considerando diferentes tamanhos amostrais.

Das Tabelas 5.5 e 5.6 nota-se que, quanto maior o tamanho da amostra, melhores são as estimativas. O modelo que apresenta melhores resultados é o modelo proposto, t-Tobit com erros nas covariáveis. Assim como no Cenário 1 pode-se observar que, se existem evidências de que as covariáveis foram medidas com erro, deve-se considerar os modelos que levam isso em conta, uma vez que os erros quadráticos médio nos modelos ajustados sem erro nas covariáveis são maiores que nos modelos com erros de medida nas covariáveis. Considerando o maior tamanho amostral observa-se que a estimativa pontual de $\mathbf{\Sigma}_{\xi_{12}}$ ainda é relativamente distante do valor real em ambos os modelos com erros nas covariáveis. Assim como para os outros parâmetros as estimativas obtidas para o grau de liberdade ν tendem a ficar cada vez mais próximas do valor real com o aumento do tamanho da amostra mas é no modelo proposto que o valor estimado fica ainda mais próximo do real. Nota-se também que, para todos os modelos ajustados, os parâmetros que são mal estimados são os de escala.

A Tabela 5.7 apresenta as medianas para os desvios-padrão *a posteriori* dos parâmetros dos modelos ajustados considerando diferentes tamanhos amostrais.

À medida que aumenta-se o tamanho da amostra observa-se, na Tabela 5.7, que os desvios-padrão *a posteriori* diminuem. Dentre os quatro modelos ajustados aos dados nota-se que o modelo proposto é o que apresenta, em geral, distribuições *a posteriori* mais precisas. Para todos os modelos as distribuições marginais *a posteriori* dos parâmetros de escala são as que possuem maior variância.

A Tabela 5.8 apresenta as estatísticas de ajuste de modelo DIC e LPML para os modelos ajustados considerando diferentes tamanhos amostrais.

Na Tabela 5.8 observa-se que, para cada modelo ajustado, o aumento no tamanho das amostras gera um decréscimo (acréscimo) nas estimativas pontuais das medidas LPML (DIC). Dentre os modelos com erros nas covariáveis, tais medidas mostram que o modelo proposto deve ser o escolhido, exceto no caso de amostras com tamanho 20. No entanto,

Tabela 5.5: Mediana e erro quadrático médio (em parênteses) para as médias *a posteriori* dos parâmetros dos modelos t-Tobit com erro nas covariáveis (tce), Tobit normal com erro nas covariáveis (nce), t-Tobit (tse) e Tobit normal (nse) em função do tamanho da amostra.

	Modelo		tce	nce	tse	nse
n	Parâmetro	Real	mediana mse	mediana mse	mediana mse	mediana mse
20	β_1	-6	-8,106 $(5,381 \times 10^5)$	-7,042 $(4,965 \times 10^4)$	-386,733 $(1,150 \times 10^7)$	-5,549 $(3,959 \times 10^5)$
	$\boldsymbol{\beta}_{21}$	0,6	0,614 $(1,629 \times 10^{-2})$	0,613 $(1,745 \times 10^{-2})$	0,568 $(1,515 \times 10^{-2})$	0,558 $(1,594 \times 10^{-2})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,305 $(5,761 \times 10^{-2})$	-0,295 $(6,357 \times 10^{-2})$	-0,276 $(5,170 \times 10^{-2})$	-0,267 $(5,275 \times 10^{-2})$
	$\boldsymbol{\beta}_{31}$	0,5	0,484 $(3,666 \times 10^{-1})$	0,517 $(3,756 \times 10^{-1})$	0,522 $(3,810 \times 10^{-1})$	0,521 $(3,838 \times 10^{-1})$
	$\boldsymbol{\beta}_{32}$	2	2,955 $(5,380 \times 10^5)$	2,334 $(4,953 \times 10^4)$	382,381 $(1,151 \times 10^7)$	2,290 $(3,958 \times 10^5)$
	σ_u	18	15,994 $(4,065 \times 10^2)$	16,298 $(3,137 \times 10^2)$	33,913 $(1,021 \times 10^3)$	29,940 $(5,092 \times 10^2)$
	$\boldsymbol{\mu}_{\xi_1}$	20	20,280 $(1,398 \times 10^1)$	20,156 $(1,553 \times 10^1)$		
	$\boldsymbol{\mu}_{\xi_2}$	1	1,149 $(3,833)$	1,082 $(4,454)$		
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	276,960 $(1,859 \times 10^4)$	298,833 $(3,576 \times 10^4)$		
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	-0,515 $(8,310 \times 10^2)$	-0,312 $(1,692 \times 10^3)$		
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	146,123 $(1,046 \times 10^4)$	150,951 $(1,298 \times 10^4)$		
	ν	5	50,910 $(1,972 \times 10^3)$		2,591 $(1,362 \times 10^3)$	
50	β_1	-6	-6,469 $(9,383 \times 10^1)$	-6,278 $(1,938 \times 10^1)$	-5,442 $(1,800 \times 10^5)$	-4,957 $(2,022 \times 10^1)$
	$\boldsymbol{\beta}_{21}$	0,6	0,601 $(5,209 \times 10^{-3})$	0,591 $(6,876 \times 10^{-3})$	0,540 $(8,235 \times 10^{-3})$	0,537 $(9,025 \times 10^{-3})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,292 $(1,538 \times 10^{-2})$	-0,293 $(1,991 \times 10^{-2})$	-0,262 $(1,563 \times 10^{-2})$	-0,265 $(1,767 \times 10^{-2})$
	$\boldsymbol{\beta}_{31}$	0,5	0,547 $(1,021 \times 10^{-1})$	0,538 $(1,140 \times 10^{-1})$	0,549 $(1,099 \times 10^{-1})$	0,543 $(1,160 \times 10^{-1})$
	$\boldsymbol{\beta}_{32}$	2	2,012 $(7,947 \times 10^1)$	1,948 $(4,773)$	1,961 $(1,797 \times 10^5)$	1,946 $(4,823)$
	σ_u	18	17,188 $(5,018 \times 10^1)$	21,388 $(1,321 \times 10^2)$	29,583 $(1,137 \times 10^3)$	32,455 $(3,834 \times 10^2)$
	$\boldsymbol{\mu}_{\xi_1}$	20	20,026 $(5,268)$	20,091 $(6,462)$		
	$\boldsymbol{\mu}_{\xi_2}$	1	1,043 $(1,384)$	0,979 $(1,688)$		
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	246,646 $(7,652 \times 10^3)$	307,865 $(2,834 \times 10^4)$		
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	1,075 $(3,597 \times 10^2)$	0,941 $(9,864 \times 10^2)$		
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	98,243 $(2,555 \times 10^3)$	109,728 $(5,352 \times 10^3)$		
	ν	5	17,261 $(2,558 \times 10^2)$		29,783 $(5,931 \times 10^2)$	
200	β_1	-6	-5,960 $(3,516)$	-5,751 $(4,666)$	-4,924 $(1,306 \times 10^5)$	-4,844 $(5,645)$
	$\boldsymbol{\beta}_{21}$	0,6	0,602 $(1,145 \times 10^{-3})$	0,595 $(1,974 \times 10^{-3})$	0,540 $(4,667 \times 10^{-3})$	0,539 $(5,308 \times 10^{-3})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,294 $(3,665 \times 10^{-3})$	-0,297 $(6,084 \times 10^{-3})$	-0,269 $(4,706 \times 10^{-3})$	-0,268 $(6,282 \times 10^{-3})$
	$\boldsymbol{\beta}_{31}$	0,5	0,495 $(2,057 \times 10^{-2})$	0,494 $(2,723 \times 10^{-2})$	0,504 $(1,950 \times 10^{-2})$	0,517 $(2,220 \times 10^{-2})$
	$\boldsymbol{\beta}_{32}$	2	2,042 $(1,043)$	2,111 $(1,220)$	1,982 $(1,307 \times 10^5)$	1,977 $(1,479)$
	σ_u	18	18,780 $(1,220 \times 10^1)$	25,782 $(1,190 \times 10^2)$	31,731 $(3,624 \times 10^2)$	36,092 $(4,103 \times 10^2)$
	$\boldsymbol{\mu}_{\xi_1}$	20	20,035 $(1,319)$	19,997 $(1,667)$		
	$\boldsymbol{\mu}_{\xi_2}$	1	1,053 $(3,574 \times 10^{-1})$	1,022 $(4,429 \times 10^{-1})$		
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	200,663 $(1,079 \times 10^3)$	297,542 $(1,745 \times 10^4)$		
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,682 $(6,209 \times 10^1)$	0,874 $(3,120 \times 10^2)$		
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	64,426 $(2,672 \times 10^2)$	89,974 $(2,927 \times 10^3)$		
	ν	5	6,244 $(5,802)$		15,374 $(1,109 \times 10^2)$	
500	β_1	-6	-5,974 $(1,205)$	-5,787 $(1,628)$	-4,708 $(2,736)$	-4,698 $(3,172)$
	$\boldsymbol{\beta}_{21}$	0,6	0,600 $(4,370 \times 10^{-4})$	0,591 $(9,081 \times 10^{-4})$	0,540 $(4,117 \times 10^{-3})$	0,538 $(4,597 \times 10^{-3})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,302 $(1,382 \times 10^{-3})$	-0,294 $(2,483 \times 10^{-3})$	-0,270 $(2,232 \times 10^{-3})$	-0,267 $(2,852 \times 10^{-3})$
	$\boldsymbol{\beta}_{31}$	0,5	0,493 $(5,950 \times 10^{-3})$	0,495 $(7,366 \times 10^{-3})$	0,494 $(6,323 \times 10^{-3})$	0,496 $(7,414 \times 10^{-3})$
	$\boldsymbol{\beta}_{32}$	2	2,032 $(4,634 \times 10^{-1})$	2,087 $(5,667 \times 10^{-1})$	2,067 $(5,062 \times 10^{-1})$	2,072 $(5,606 \times 10^{-1})$
	σ_u	18	18,092 $(4,994)$	26,813 $(9,768 \times 10^1)$	31,909 $(2,008 \times 10^2)$	36,962 $(3,907 \times 10^2)$
	$\boldsymbol{\mu}_{\xi_1}$	20	19,910 $(5,274 \times 10^{-1})$	19,938 $(6,939 \times 10^{-1})$		
	$\boldsymbol{\mu}_{\xi_2}$	1	1,002 $(1,353 \times 10^{-1})$	1,002 $(1,775 \times 10^{-1})$		
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	188,287 $(3,102 \times 10^2)$	298,948 $(1,587 \times 10^4)$		
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,040 $(2,377 \times 10^1)$	0,620 $(1,556 \times 10^2)$		
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	55,379 $(4,994 \times 10^1)$	85,408 $(1,442 \times 10^3)$		
	ν	5	5,318 $(6,908 \times 10^{-1})$		12,701 $(6,375 \times 10^1)$	

Tabela 5.6: Mediana e erro quadrático médio (em parênteses) para as medianas *a posteriori* dos parâmetros dos modelos t-Tobit com erro nas covariáveis (tce), Tobit normal com erro nas covariáveis (nce), t-Tobit (tse) e Tobit normal (nse) em função do tamanho da amostra.

	Modelo		tce	nce	tse	nse
n	Parâmetro	Real	mediana mse	mediana mse	mediana mse	mediana mse
20	β_1	-6	-7,995 $(5,393 \times 10^5)$	-6,840 $(4,948 \times 10^4)$	-383,523 $(1,150 \times 10^7)$	-5,594 $(3,957 \times 10^5)$
	$\boldsymbol{\beta}_{21}$	0,6	0,613 $(1,586 \times 10^{-2})$	0,610 $(1,712 \times 10^{-2})$	0,563 $(1,511 \times 10^{-2})$	0,555 $(1,609 \times 10^{-2})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,302 $(5,696 \times 10^{-2})$	-0,292 $(6,259 \times 10^{-2})$	-0,275 $(5,136 \times 10^{-2})$	-0,268 $(5,214 \times 10^{-2})$
	$\boldsymbol{\beta}_{31}$	0,5	0,493 $(3,636 \times 10^{-1})$	0,499 $(3,708 \times 10^{-1})$	0,506 $(3,693 \times 10^{-1})$	0,520 $(3,830 \times 10^{-1})$
	$\boldsymbol{\beta}_{32}$	2	2,971 $(5,392 \times 10^5)$	2,377 $(4,933 \times 10^4)$	377,027 $(1,150 \times 10^7)$	2,212 $(3,956 \times 10^5)$
	σ_u	18	13,831 $(2,523 \times 10^2)$	14,568 $(2,617 \times 10^2)$	30,268 $(6,247 \times 10^2)$	27,277 $(3,861 \times 10^2)$
	$\boldsymbol{\mu}_{\xi_1}$	20	20,254 $(1,401 \times 10^1)$	20,204 $(1,559 \times 10^1)$		
	$\boldsymbol{\mu}_{\xi_2}$	1	1,139 $(3,771)$	1,099 $(4,496)$		
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	260,199 $(1,432 \times 10^4)$	282,649 $(2,935 \times 10^4)$		
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	-0,077 $(7,024 \times 10^2)$	-0,381 $(1,459 \times 10^3)$		
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	137,509 $(8,735 \times 10^3)$	142,012 $(1,098 \times 10^4)$		
	ν	5	36,094 $(1,005 \times 10^3)$		2,450 $(7,173 \times 10^2)$	
50	β_1	-6	-6,406 $(1,142 \times 10^2)$	-6,245 $(1,935 \times 10^1)$	-5,403 $(1,780 \times 10^5)$	-4,872 $(2,019 \times 10^1)$
	$\boldsymbol{\beta}_{21}$	0,6	0,598 $(5,151 \times 10^{-3})$	0,591 $(6,908 \times 10^{-3})$	0,539 $(8,347 \times 10^{-3})$	0,537 $(9,129 \times 10^{-3})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,291 $(1,539 \times 10^{-2})$	-0,291 $(1,988 \times 10^{-2})$	-0,260 $(1,574 \times 10^{-2})$	-0,267 $(1,782 \times 10^{-2})$
	$\boldsymbol{\beta}_{31}$	0,5	0,547 $(1,013 \times 10^{-1})$	0,542 $(1,135 \times 10^{-1})$	0,551 $(1,093 \times 10^{-1})$	0,546 $(1,161 \times 10^{-1})$
	$\boldsymbol{\beta}_{32}$	2	1,984 $(1,039 \times 10^2)$	1,936 $(4,776)$	1,990 $(1,774 \times 10^5)$	1,940 $(4,808)$
	σ_u	18	16,499 $(4,647 \times 10^1)$	20,515 $(1,177 \times 10^2)$	28,377 $(5,117 \times 10^2)$	31,397 $(3,394 \times 10^2)$
	$\boldsymbol{\mu}_{\xi_1}$	20	20,009 $(5,290)$	20,099 $(6,417)$		
	$\boldsymbol{\mu}_{\xi_2}$	1	1,044 $(1,416)$	0,985 $(1,686)$		
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	240,310 $(6,562 \times 10^3)$	300,099 $(2,567 \times 10^4)$		
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,693 $(3,291 \times 10^2)$	1,134 $(9,241 \times 10^2)$		
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	95,498 $(2,296 \times 10^3)$	107,010 $(4,938 \times 10^3)$		
	ν	5	12,662 $(1,383 \times 10^2)$		23,485 $(3,525 \times 10^2)$	
200	β_1	-6	-5,936 $(3,495)$	-5,778 $(4,697)$	-4,882 $(1,304 \times 10^5)$	-4,837 $(5,698)$
	$\boldsymbol{\beta}_{21}$	0,6	0,602 $(1,139 \times 10^{-3})$	0,595 $(1,984 \times 10^{-3})$	0,540 $(4,705 \times 10^{-3})$	0,539 $(5,328 \times 10^{-3})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,294 $(3,670 \times 10^{-3})$	-0,296 $(6,123 \times 10^{-3})$	-0,270 $(4,698 \times 10^{-3})$	-0,267 $(6,261 \times 10^{-3})$
	$\boldsymbol{\beta}_{31}$	0,5	0,493 $(2,065 \times 10^{-2})$	0,498 $(2,734 \times 10^{-2})$	0,502 $(1,970 \times 10^{-2})$	0,516 $(2,234 \times 10^{-2})$
	$\boldsymbol{\beta}_{32}$	2	2,043 $(1,046)$	2,091 $(1,226)$	1,971 $(1,305 \times 10^5)$	1,949 $(1,479)$
	σ_u	18	18,556 $(1,157 \times 10^1)$	25,616 $(1,141 \times 10^2)$	31,458 $(3,426 \times 10^2)$	35,692 $(3,986 \times 10^2)$
	$\boldsymbol{\mu}_{\xi_1}$	20	20,046 $(1,336)$	20,023 $(1,684)$		
	$\boldsymbol{\mu}_{\xi_2}$	1	1,056 $(3,586 \times 10^{-1})$	1,026 $(4,434 \times 10^{-1})$		
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	199,390 $(9,971 \times 10^2)$	295,514 $(1,693 \times 10^4)$		
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,625 $(6,054 \times 10^1)$	1,092 $(3,058 \times 10^2)$		
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	63,704 $(2,526 \times 10^2)$	89,253 $(2,851 \times 10^3)$		
	ν	5	6,038 $(4,375)$		14,144 $(8,915 \times 10^1)$	
500	β_1	-6	-5,972 $(1,208)$	-5,774 $(1,636)$	-4,719 $(2,762)$	-4,707 $(3,191)$
	$\boldsymbol{\beta}_{21}$	0,6	0,600 $(4,393 \times 10^{-4})$	0,592 $(9,148 \times 10^{-4})$	0,539 $(4,122 \times 10^{-3})$	0,538 $(4,607 \times 10^{-3})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,302 $(1,388 \times 10^{-3})$	-0,294 $(2,480 \times 10^{-3})$	-0,269 $(2,242 \times 10^{-3})$	-0,268 $(2,849 \times 10^{-3})$
	$\boldsymbol{\beta}_{31}$	0,5	0,493 $(5,995 \times 10^{-3})$	0,497 $(7,361 \times 10^{-3})$	0,494 $(6,310 \times 10^{-3})$	0,496 $(7,520 \times 10^{-3})$
	$\boldsymbol{\beta}_{32}$	2	2,029 $(4,684 \times 10^{-1})$	2,088 $(5,705 \times 10^{-1})$	2,065 $(5,072 \times 10^{-1})$	2,080 $(5,600 \times 10^{-1})$
	σ_u	18	18,053 $(4,888)$	26,729 $(9,579 \times 10^1)$	31,766 $(1,979 \times 10^2)$	36,911 $(3,861 \times 10^2)$
	$\boldsymbol{\mu}_{\xi_1}$	20	19,905 $(5,285 \times 10^{-1})$	19,948 $(6,901 \times 10^{-1})$		
	$\boldsymbol{\mu}_{\xi_2}$	1	0,998 $(1,360 \times 10^{-1})$	0,997 $(1,784 \times 10^{-1})$		
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	187,425 $(2,961 \times 10^2)$	298,398 $(1,567 \times 10^4)$		
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,012 $(2,368 \times 10^1)$	0,531 $(1,554 \times 10^2)$		
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	55,154 $(4,791 \times 10^1)$	85,048 $(1,425 \times 10^3)$		
	ν	5	5,258 $(6,080 \times 10^{-1})$		12,231 $(5,750 \times 10^1)$	

Tabela 5.7: Mediana para os desvios-padrão *a posteriori* dos parâmetros dos modelos t-Tobit com erro nas covariáveis (tce), Tobit normal com erro nas covariáveis (nce), t-Tobit (tse) e Tobit normal (nse) em função do tamanho da amostra.

Modelo		tce	nce	tse	nse
n	Parâmetro	mediana	mediana	mediana	mediana
20	β_1	7,189	6,804	12,560	6,762
	$\boldsymbol{\beta}_{21}$	0,114	0,103	0,108	0,094
	$\boldsymbol{\beta}_{22}$	0,194	0,181	0,187	0,167
	$\boldsymbol{\beta}_{31}$	0,506	0,497	0,557	0,505
	$\boldsymbol{\beta}_{32}$	4,701	4,572	13,322	4,549
	σ_u	13,471	13,252	16,882	11,661
	$\boldsymbol{\mu}_{\xi_1}$	3,988	4,016		
	$\boldsymbol{\mu}_{\xi_2}$	2,835	2,856		
	$\boldsymbol{\Sigma}_{\xi_{11}}$	92,733	94,970		
	$\boldsymbol{\Sigma}_{\xi_{12}}$	45,702	47,235		
	$\boldsymbol{\Sigma}_{\xi_{22}}$	46,800	47,046		
	ν	45,887		0,565	
50	β_1	3,925	4,104	3,926	4,025
	$\boldsymbol{\beta}_{21}$	0,065	0,061	0,056	0,055
	$\boldsymbol{\beta}_{22}$	0,113	0,106	0,098	0,097
	$\boldsymbol{\beta}_{31}$	0,297	0,318	0,304	0,316
	$\boldsymbol{\beta}_{32}$	1,941	2,050	2,012	2,063
	σ_u	7,045	7,846	7,550	7,504
	$\boldsymbol{\mu}_{\xi_1}$	2,447	2,576		
	$\boldsymbol{\mu}_{\xi_2}$	1,491	1,563		
	$\boldsymbol{\Sigma}_{\xi_{11}}$	55,049	61,707		
	$\boldsymbol{\Sigma}_{\xi_{12}}$	23,013	25,785		
	$\boldsymbol{\Sigma}_{\xi_{22}}$	20,789	22,058		
	ν	14,122		21,294	
200	β_1	1,858	2,003	1,773	1,852
	$\boldsymbol{\beta}_{21}$	0,033	0,031	0,028	0,028
	$\boldsymbol{\beta}_{22}$	0,059	0,053	0,049	0,048
	$\boldsymbol{\beta}_{31}$	0,142	0,157	0,138	0,144
	$\boldsymbol{\beta}_{32}$	0,970	1,058	1,093	1,139
	σ_u	3,435	4,105	3,968	3,963
	$\boldsymbol{\mu}_{\xi_1}$	1,142	1,274		
	$\boldsymbol{\mu}_{\xi_2}$	0,638	0,698		
	$\boldsymbol{\Sigma}_{\xi_{11}}$	25,314	29,767		
	$\boldsymbol{\Sigma}_{\xi_{12}}$	8,730	11,508		
	$\boldsymbol{\Sigma}_{\xi_{22}}$	7,716	9,011		
	ν	1,477		6,066	
500	β_1	1,066	1,171	1,090	1,156
	$\boldsymbol{\beta}_{21}$	0,021	0,019	0,018	0,017
	$\boldsymbol{\beta}_{22}$	0,037	0,034	0,031	0,031
	$\boldsymbol{\beta}_{31}$	0,077	0,087	0,082	0,087
	$\boldsymbol{\beta}_{32}$	0,642	0,721	0,684	0,726
	σ_u	2,102	2,654	2,519	2,534
	$\boldsymbol{\mu}_{\xi_1}$	0,712	0,810		
	$\boldsymbol{\mu}_{\xi_2}$	0,380	0,435		
	$\boldsymbol{\Sigma}_{\xi_{11}}$	15,556	18,885		
	$\boldsymbol{\Sigma}_{\xi_{12}}$	5,046	7,157		
	$\boldsymbol{\Sigma}_{\xi_{22}}$	4,431	5,370		
	ν	0,715		3,047	

Tabela 5.8: Medianas e médias dos valores obtidos de LPML e DIC para os modelos t-Tobit com erro nas covariáveis (tce), Tobit normal com erro nas covariáveis (nce), t-Tobit (tse) e Tobit normal (nse) em função do tamanho da amostra.

Modelo		tce		nce		tse		nse	
n	Estatística	mediana	média	mediana	média	mediana	média	mediana	média
20	LPML	-224,447	-227,681	-222,509	-222,947	-69,760	-75,589	-61,104	-61,391
	DIC	444,810	450,963	441,020	441,914	135,704	148,210	119,664	119,551
50	LPML	-544,606	-545,421	-549,111	-550,325	-147,740	-149,180	-148,205	-148,330
	DIC	1087,028	1088,853	1094,129	1096,222	293,272	295,534	294,183	294,076
200	LPML	-2154,761	-2154,764	-2181,829	-2185,024	-578,324	-583,764	-582,387	-584,315
	DIC	4309,232	4309,047	4359,496	4363,589	1155,311	1166,131	1162,523	1165,743
500	LPML	-5369,651	-5369,453	-5449,337	-5451,837	-1439,975	-1438,813	-1453,543	-1452,201
	DIC	10739,039	10738,773	10891,066	10895,724	2878,658	2876,212	2903,975	2900,870

dentre os modelos sem erros nas covariáveis tais medidas apontam o modelo mais robusto t-Tobit como sendo o mais apropriado para os dados, a não ser quando $n = 20$. Assim, como a conclusão para o Cenário 1, os resultados obtidos aqui mostram que apenas considerar as medidas LPML e DIC como medidas para a escolha do modelo mais adequado pode levar à decisões errôneas, uma vez que todas as amostras foram geradas sob o modelo proposto e dentre todos os modelos ajustados ele não foi o selecionado por tais medidas.

5.1.3 O efeito da má especificação de $\boldsymbol{\Delta}$ nos estimadores bayesianos - Cenário 3

As Tabelas 5.9 e 5.10 mostram as medianas e os erros quadráticos médio para a média e mediana *a posteriori*, respectivamente, dos parâmetros dos modelos ajustados considerando diferentes especificações para $\boldsymbol{\Delta}$.

As Tabelas 5.9 e 5.10 mostram como se comportam os estimadores de alguns parâmetros quando são especificados diferentes valores para $\boldsymbol{\Delta}$. As estimativas dos parâmetros $\boldsymbol{\beta}_3$, $\boldsymbol{\mu}_\xi$ e do grau de liberdade ν estão omitidas pois, para diferentes valores especificados em $\boldsymbol{\Delta}$, eles não se alteram, como mencionado na Seção 3.2. Quando valores em $\boldsymbol{\Delta}$ que não sejam os reais são especificados, nota-se que os erros quadráticos médio tendem a aumentar à medida que os valores especificados para $\boldsymbol{\Delta}$ se distanciam do real. Para os casos onde os valores especificados para $\boldsymbol{\Delta}$ possuem valores menores (maiores) que os reais ocorre a superestimação (subestimação) dos parâmetros β_1, $\boldsymbol{\beta}_{22}$ e σ_u e a subestimação (superestimação) de $\boldsymbol{\beta}_{21}$. Os parâmetros de $\boldsymbol{\Sigma}_\xi$ são superestimados para quaisquer especificações em $\boldsymbol{\Delta}$ mas nota-se um decréscimo em seus valores à medida que se especifica

Tabela 5.9: Mediana e erro quadrático médio (em parênteses) para as médias *a posteriori* dos parâmetros dos modelos t-Tobit com erro nas covariáveis (tce) e Tobit normal com erro nas covariáveis (nce) em função da especificação de $\boldsymbol{\Delta}$, onde $\boldsymbol{\Delta}$ real é $0,1\boldsymbol{I}_2$.

	Modelo		tce		nce	
$\boldsymbol{\Delta}$	Parâmetro	Real	mediana	mse	mediana	mse
$0,001\boldsymbol{I}_2$	β_1	-6	-4,876	$(4,505)$	-4,697	$(6,037)$
	$\boldsymbol{\beta}_{21}$	0,6	0,548	$(3,742 \times 10^{-3})$	0,542	$(5,176 \times 10^{-3})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,268	$(3,902 \times 10^{-3})$	-0,270	$(6,157 \times 10^{-3})$
	$\boldsymbol{\beta}_{31}$	0,5	0,495	$(2,057 \times 10^{-2})$	0,494	$(2,723 \times 10^{-2})$
	$\boldsymbol{\beta}_{32}$	2	2,042	$(1,043)$	2,111	$(1,220)$
	σ_u	18	25,891	$(7,649 \times 10^{1})$	35,954	$(4,051 \times 10^{2})$
	$\boldsymbol{\mu}_{\xi_1}$	20	20,035	$(1,319)$	19,997	$(1,667)$
	$\boldsymbol{\mu}_{\xi_2}$	1	1,053	$(3,574 \times 10^{-1})$	1,022	$(4,429 \times 10^{-1})$
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	220,509	$(2,497 \times 10^{3})$	326,969	$(2,618 \times 10^{4})$
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,750	$(7,497 \times 10^{1})$	0,960	$(3,767 \times 10^{2})$
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	70,798	$(5,091 \times 10^{2})$	98,872	$(4,029 \times 10^{3})$
	ν	5	6,244	$(5,802)$		
$0,05\boldsymbol{I}_2$	β_1	-6	-5,399	$(3,734)$	-5,215	$(5,084)$
	$\boldsymbol{\beta}_{21}$	0,6	0,575	$(1,722 \times 10^{-3})$	0,568	$(2,873 \times 10^{-3})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,281	$(3,598 \times 10^{-3})$	-0,283	$(5,936 \times 10^{-3})$
	σ_u	18	22,184	$(3,152 \times 10^{1})$	30,915	$(2,346 \times 10^{2})$
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	210,219	$(1,660 \times 10^{3})$	311,711	$(2,143 \times 10^{4})$
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,715	$(6,814 \times 10^{1})$	0,916	$(3,424 \times 10^{2})$
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	67,494	$(3,733 \times 10^{2})$	94,258	$(3,434 \times 10^{3})$
$0,1\boldsymbol{I}_2$	β_1	-6	-5,960	$(3,516)$	-5,751	$(4,666)$
	$\boldsymbol{\beta}_{21}$	0,6	0,602	$(1,145 \times 10^{-3})$	0,595	$(1,974 \times 10^{-3})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,294	$(3,665 \times 10^{-3})$	-0,297	$(6,084 \times 10^{-3})$
	σ_u	18	18,780	$(1,220 \times 10^{1})$	25,782	$(1,190 \times 10^{2})$
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	200,663	$(1,079 \times 10^{3})$	297,542	$(1,745 \times 10^{4})$
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,682	$(6,209 \times 10^{1})$	0,874	$(3,120 \times 10^{2})$
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	64,426	$(2,672 \times 10^{2})$	89,974	$(2,927 \times 10^{3})$
$0,15\boldsymbol{I}_2$	β_1	-6	-6,500	$(3,873)$	-6,276	$(4,808)$
	$\boldsymbol{\beta}_{21}$	0,6	0,629	$(2,066 \times 10^{-3})$	0,622	$(2,538 \times 10^{-3})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,307	$(4,111 \times 10^{-3})$	-0,310	$(6,609 \times 10^{-3})$
	σ_u	18	15,119	$(1,973 \times 10^{1})$	20,403	$(6,237 \times 10^{1})$
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	191,939	$(7,132 \times 10^{2})$	284,606	$(1,420 \times 10^{4})$
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,652	$(5,682 \times 10^{1})$	0,836	$(2,854 \times 10^{2})$
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	61,625	$(1,873 \times 10^{2})$	86,062	$(2,503 \times 10^{3})$
$0,3\boldsymbol{I}_2$	β_1	-6	-8,115	$(8,387)$	-7,857	$(8,594)$
	$\boldsymbol{\beta}_{21}$	0,6	0,712	$(1,382 \times 10^{-2})$	0,704	$(1,302 \times 10^{-2})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,347	$(7,729 \times 10^{-3})$	-0,351	$(1,045 \times 10^{-2})$
	σ_u	18	4,377	$(2,034 \times 10^{2})$	5,611	$(2,462 \times 10^{2})$
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	169,792	$(4,897 \times 10^{2})$	251,767	$(7,531 \times 10^{3})$
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,577	$(4,448 \times 10^{1})$	0,739	$(2,234 \times 10^{2})$
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	54,514	$(5,658 \times 10^{1})$	76,132	$(1,592 \times 10^{3})$

Tabela 5.10: Mediana e erro quadrático médio (em parênteses) para as medianas *a posteriori* dos parâmetros dos modelos t-Tobit com erro nas covariáveis (tce) e Tobit normal com erro nas covariáveis (nce) em função da especificação de $\boldsymbol{\Delta}$, onde $\boldsymbol{\Delta}$ real é $0,1\boldsymbol{I}_2$.

Modelo			tce		nce	
$\boldsymbol{\Delta}$	Parâmetro	Real	mediana	mse	mediana	mse
$0,001\boldsymbol{I}_2$	β_1	-6	-4,876	$(4,504)$	-4,704	$(6,089)$
	$\boldsymbol{\beta}_{21}$	0,6	0,547	$(3,758 \times 10^{-3})$	0,541	$(5,205 \times 10^{-3})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,268	$(3,925 \times 10^{-3})$	-0,269	$(6,205 \times 10^{-3})$
	$\boldsymbol{\beta}_{31}$	0,5	0,493	$(2,065 \times 10^{-2})$	0,498	$(2,734 \times 10^{-2})$
	$\boldsymbol{\beta}_{32}$	2	2,043	$(1,046)$	2,091	$(1,226)$
	σ_u	18	25,610	$(7,244 \times 10^{1})$	35,617	$(3,933 \times 10^{2})$
	$\boldsymbol{\mu}_{\xi_1}$	20	20,046	$(1,336)$	20,023	$(1,684)$
	$\boldsymbol{\mu}_{\xi_2}$	1	1,056	$(3,586 \times 10^{-1})$	1,026	$(4,434 \times 10^{-1})$
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	219,110	$(2,333 \times 10^{3})$	324,740	$(2,547 \times 10^{4})$
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,687	$(7,310 \times 10^{1})$	1,200	$(3,692 \times 10^{2})$
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	70,004	$(4,861 \times 10^{2})$	98,080	$(3,931 \times 10^{3})$
	ν	5	6,038	$(4,375)$		
$0,05\boldsymbol{I}_2$	β_1	-6	-5,364	$(3,725)$	-5,242	$(5,126)$
	$\boldsymbol{\beta}_{21}$	0,6	0,574	$(1,728 \times 10^{-3})$	0,568	$(2,894 \times 10^{-3})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,281	$(3,613 \times 10^{-3})$	-0,282	$(5,980 \times 10^{-3})$
	σ_u	18	21,931	$(2,920 \times 10^{1})$	30,773	$(2,263 \times 10^{2})$
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	208,885	$(1,540 \times 10^{3})$	309,586	$(2,082 \times 10^{4})$
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,655	$(6,644 \times 10^{1})$	1,144	$(3,356 \times 10^{2})$
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	66,737	$(3,548 \times 10^{2})$	93,503	$(3,347 \times 10^{3})$
$0,1\boldsymbol{I}_2$	β_1	-6	-5,936	$(3,495)$	-5,778	$(4,697)$
	$\boldsymbol{\beta}_{21}$	0,6	0,602	$(1,139 \times 10^{-3})$	0,595	$(1,984 \times 10^{-3})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,294	$(3,670 \times 10^{-3})$	-0,296	$(6,123 \times 10^{-3})$
	σ_u	18	18,556	$(1,157 \times 10^{1})$	25,616	$(1,141 \times 10^{2})$
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	199,390	$(9,971 \times 10^{2})$	295,514	$(1,693 \times 10^{4})$
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,625	$(6,054 \times 10^{1})$	1,092	$(3,058 \times 10^{2})$
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	63,704	$(2,526 \times 10^{2})$	89,253	$(2,851 \times 10^{3})$
$0,15\boldsymbol{I}_2$	β_1	-6	-6,478	$(3,845)$	-6,308	$(4,826)$
	$\boldsymbol{\beta}_{21}$	0,6	0,629	$(2,047 \times 10^{-3})$	0,622	$(2,537 \times 10^{-3})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,308	$(4,107 \times 10^{-3})$	-0,309	$(6,643 \times 10^{-3})$
	σ_u	18	14,924	$(2,057 \times 10^{1})$	20,231	$(6,031 \times 10^{1})$
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	190,721	$(6,629 \times 10^{2})$	282,665	$(1,375 \times 10^{4})$
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,598	$(5,540 \times 10^{1})$	1,045	$(2,797 \times 10^{2})$
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	60,934	$(1,759 \times 10^{2})$	85,372	$(2,436 \times 10^{3})$
$0,3\boldsymbol{I}_2$	β_1	-6	-8,088	$(8,318)$	-7,827	$(8,585)$
	$\boldsymbol{\beta}_{21}$	0,6	0,711	$(1,376 \times 10^{-2})$	0,703	$(1,298 \times 10^{-2})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,348	$(7,691 \times 10^{-3})$	-0,350	$(1,046 \times 10^{-2})$
	σ_u	18	4,395	$(2,030 \times 10^{2})$	5,656	$(2,442 \times 10^{2})$
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	168,715	$(5,085 \times 10^{2})$	250,050	$(7,251 \times 10^{3})$
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,529	$(4,337 \times 10^{1})$	0,924	$(2,189 \times 10^{2})$
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	53,903	$(5,245 \times 10^{1})$	75,522	$(1,546 \times 10^{3})$

valores maiores para $\boldsymbol{\Delta}$. No modelo proposto e no caso onde $\boldsymbol{\Delta} = 0,3\boldsymbol{I}_2$, a estimativa de $\boldsymbol{\Sigma}_{\xi_{11}}$ é menor que o valor real. Para o modelo proposto por Wang (1998) observa-se que σ_u é sempre superestimado, exceto quando $\boldsymbol{\Delta} = 0,3I_2$. Em geral, o modelo proposto gerou melhores estimativas que o modelo Tobit normal com erros nas covariáveis. Também nota-se que a mediana e a média *a posteriori* tendem a fornecer estimativas similares entre si, qualquer que seja o valor atribuído a $\boldsymbol{\Delta}$.

A Tabela 5.11 apresenta as medianas para os desvios-padrão *a posteriori* dos parâmetros dos modelos ajustados considerando diferentes especificações para $\boldsymbol{\Delta}$.

Observa-se da Tabela 5.11 que, à medida que os valores especificados em $\boldsymbol{\Delta}$ aumentam, as distribuições marginais *a posteriori* de β_1, $\boldsymbol{\beta}_{21}$ e $\boldsymbol{\beta}_{22}$ passam a ter menores variâncias para cada modelo ajustado. O mesmo ocorre com a distribuição de σ_u no modelo proposto por Wang (1998). Para os parâmetros em $\boldsymbol{\Sigma}_\xi$ o comportamento é o oposto. No modelo proposto o desvio-padrão *a posteriori* de σ_u passa a ficar menor para as especificações em $\boldsymbol{\Delta}$ superiores a $0,05I_2$. Em geral, o modelo que gera distribuições marginais *a posteriori* para seus parâmetros com menor variabilidade é o modelo proposto, exceto para $\boldsymbol{\beta}_2$.

5.1.4 O efeito de ν nos estimadores bayesianos - Cenário 4

A Tabela 5.12 apresenta as estimativas *a posteriori* para os parâmetros do modelo proposto quando amostras são geradas assumindo-se diferentes graus de liberdade ν.

Pela Tabela 5.12 pode-se notar que ambos os estimadores média e mediana *a posteriori* apresentam estimativas pontuais próximas dos valores reais, exceto para $\boldsymbol{\Sigma}_\xi$. Quanto maior o valor do grau de liberdade utilizado para geração dos dados, menores são os erros quadráticos médio e os desvios-padrão *a posteriori* para todos os parâmetros do modelo proposto, exceto para o grau de liberdade. Nota-se, como afirmado na literatura, que em geral há perda de robustez ao estimar ν, mas quando as amostras são geradas sob o modelo Tobit normal com erros nas covariáveis ($\nu = \infty$) tem-se que, ajustando-se o modelo proposto aos dados, o grau de liberdade ν é estimado em 22,172 (20,835) pela média (mediana) *a posteriori*, indicando que o uso de um modelo mais robusto é mais apropriado.

Tabela 5.11: Mediana para os desvios-padrão dos parâmetros dos modelos t-Tobit com erro nas covariáveis (tce) e Tobit normal com erro nas covariáveis (nce) em função da especificação de $\boldsymbol{\Delta}$, onde $\boldsymbol{\Delta}$ real é $0,1\boldsymbol{I}_2$.

Modelo		tce	nce
$\boldsymbol{\Delta}$	Estatística	mediana	mediana
$0,001\boldsymbol{I}_2$	β_1	1,830	1,981
	$\boldsymbol{\beta}_{21}$	0,030	0,028
	$\boldsymbol{\beta}_{22}$	0,053	0,048
	$\boldsymbol{\beta}_{31}$	0,142	0,157
	$\boldsymbol{\beta}_{32}$	0,970	1,058
	σ_u	3,456	3,920
	$\boldsymbol{\mu}_{\xi_1}$	1,142	1,274
	$\boldsymbol{\mu}_{\xi_2}$	0,638	0,698
	$\boldsymbol{\Sigma}_{\xi_{11}}$	27,818	32,711
	$\boldsymbol{\Sigma}_{\xi_{12}}$	9,594	12,647
	$\boldsymbol{\Sigma}_{\xi_{22}}$	8,479	9,902
	ν	1,477	
$0,05\boldsymbol{I}_2$	β_1	1,843	1,990
	$\boldsymbol{\beta}_{21}$	0,032	0,029
	$\boldsymbol{\beta}_{22}$	0,056	0,051
	σ_u	3,407	3,942
	$\boldsymbol{\Sigma}_{\xi_{11}}$	26,520	31,184
	$\boldsymbol{\Sigma}_{\xi_{12}}$	9,146	12,056
	$\boldsymbol{\Sigma}_{\xi_{22}}$	8,083	9,440
$0,1\boldsymbol{I}_2$	β_1	1,858	2,003
	$\boldsymbol{\beta}_{21}$	0,033	0,031
	$\boldsymbol{\beta}_{22}$	0,059	0,053
	σ_u	3,435	4,105
	$\boldsymbol{\Sigma}_{\xi_{11}}$	25,314	29,767
	$\boldsymbol{\Sigma}_{\xi_{12}}$	8,730	11,508
	$\boldsymbol{\Sigma}_{\xi_{22}}$	7,716	9,011
$0,15\boldsymbol{I}_2$	β_1	1,873	2,017
	$\boldsymbol{\beta}_{21}$	0,035	0,032
	$\boldsymbol{\beta}_{22}$	0,061	0,056
	σ_u	3,574	4,391
	$\boldsymbol{\Sigma}_{\xi_{11}}$	24,214	28,473
	$\boldsymbol{\Sigma}_{\xi_{12}}$	8,351	11,008
	$\boldsymbol{\Sigma}_{\xi_{22}}$	7,380	8,619
$0,3\boldsymbol{I}_2$	β_1	1,915	2,060
	$\boldsymbol{\beta}_{21}$	0,039	0,036
	$\boldsymbol{\beta}_{22}$	0,069	0,063
	σ_u	4,456	5,782
	$\boldsymbol{\Sigma}_{\xi_{11}}$	21,420	25,187
	$\boldsymbol{\Sigma}_{\xi_{12}}$	7,387	9,738
	$\boldsymbol{\Sigma}_{\xi_{22}}$	6,529	7,625

Tabela 5.12: Estimativas *a posteriori* para os parâmetros do modelo t-Tobit com erro nas covariáveis quando amostras são geradas sob diferentes graus de liberdade, ν.

Estatística *a posteriori*		média		mediana		desvio-padrão
Parâmetro	Real	mediana	mse	mediana	mse	mediana
β_1	-6	-6,055	$(4,152)$	-6,075	$(4,182)$	1,973
$\boldsymbol{\beta}_{21}$	0,6	0,598	$(1,212 \times 10^{-3})$	0,598	$(1,226 \times 10^{-3})$	0,035
$\boldsymbol{\beta}_{22}$	-0,3	-0,301	$(4,073 \times 10^{-3})$	-0,302	$(4,088 \times 10^{-3})$	0,062
$\boldsymbol{\beta}_{31}$	0,5	0,500	$(2,432 \times 10^{-2})$	0,505	$(2,436 \times 10^{-2})$	0,153
$\boldsymbol{\beta}_{32}$	2	1,931	$(1,133)$	1,944	$(1,129)$	1,044
σ_u	18	19,425	$(2,044 \times 10^{1})$	19,138	$(1,894 \times 10^{1})$	4,005
$\boldsymbol{\mu}_{\xi_1}$	20	19,988	$(1,549)$	19,985	$(1,564)$	1,243
$\boldsymbol{\mu}_{\xi_2}$	1	0,963	$(3,695 \times 10^{-1})$	0,967	$(3,722 \times 10^{-1})$	0,684
$\boldsymbol{\Sigma}_{\xi_{11}}$	180	217,498	$(2,297 \times 10^{3})$	214,892	$(2,124 \times 10^{3})$	30,188
$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,378	$(8,058 \times 10^{1})$	0,160	$(7,813 \times 10^{1})$	10,033
$\boldsymbol{\Sigma}_{\xi_{22}}$	50	68,366	$(4,308 \times 10^{2})$	67,620	$(4,043 \times 10^{2})$	9,019
ν	2,01	2,327	$(2,238 \times 10^{-1})$	2,285	$(1,937 \times 10^{-1})$	0,238
β_1	-6	-5,960	$(3,516)$	-5,936	$(3,495)$	1,858
$\boldsymbol{\beta}_{21}$	0,6	0,602	$(1,145 \times 10^{-3})$	0,602	$(1,139 \times 10^{-3})$	0,033
$\boldsymbol{\beta}_{22}$	-0,3	-0,294	$(3,665 \times 10^{-3})$	-0,294	$(3,670 \times 10^{-3})$	0,059
$\boldsymbol{\beta}_{31}$	0,5	0,495	$(2,057 \times 10^{-2})$	0,493	$(2,065 \times 10^{-2})$	0,142
$\boldsymbol{\beta}_{32}$	2	2,042	$(1,043)$	2,043	$(1,046)$	0,970
σ_u	18	18,780	$(1,220 \times 10^{1})$	18,556	$(1,157 \times 10^{1})$	3,435
$\boldsymbol{\mu}_{\xi_1}$	20	20,035	$(1,319)$	20,046	$(1,336)$	1,142
$\boldsymbol{\mu}_{\xi_2}$	1	1,053	$(3,574 \times 10^{-1})$	1,056	$(3,586 \times 10^{-1})$	0,638
$\boldsymbol{\Sigma}_{\xi_{11}}$	180	200,663	$(1,079 \times 10^{3})$	199,390	$(9,971 \times 10^{2})$	25,314
$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,682	$(6,209 \times 10^{1})$	0,625	$(6,054 \times 10^{1})$	8,730
$\boldsymbol{\Sigma}_{\xi_{22}}$	50	64,426	$(2,672 \times 10^{2})$	63,704	$(2,526 \times 10^{2})$	7,716
ν	5	6,244	$(5,802)$	6,038	$(4,375)$	1,477
β_1	-6	-6,073	$(3,098)$	-6,035	$(3,104)$	1,655
$\boldsymbol{\beta}_{21}$	0,6	0,600	$(8,794 \times 10^{-4})$	0,600	$(8,887 \times 10^{-4})$	0,030
$\boldsymbol{\beta}_{22}$	-0,3	-0,296	$(2,804 \times 10^{-3})$	-0,297	$(2,819 \times 10^{-3})$	0,055
$\boldsymbol{\beta}_{31}$	0,5	0,502	$(2,016 \times 10^{-2})$	0,498	$(2,033 \times 10^{-2})$	0,127
$\boldsymbol{\beta}_{32}$	2	1,976	$(6,862 \times 10^{-1})$	1,985	$(6,867 \times 10^{-1})$	0,857
σ_u	18	16,334	$(9,576)$	16,159	$(1,009 \times 10^{1})$	2,670
$\boldsymbol{\mu}_{\xi_1}$	20	20,038	$(1,079)$	20,036	$(1,079)$	1,008
$\boldsymbol{\mu}_{\xi_2}$	1	0,976	$(2,900 \times 10^{-1})$	0,976	$(2,898 \times 10^{-1})$	0,569
$\boldsymbol{\Sigma}_{\xi_{11}}$	180	173,374	$(3,374 \times 10^{2})$	172,254	$(3,498 \times 10^{2})$	18,553
$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	-0,015	$(3,553 \times 10^{1})$	-0,037	$(3,474 \times 10^{1})$	7,142
$\boldsymbol{\Sigma}_{\xi_{22}}$	50	55,651	$(5,106 \times 10^{1})$	55,287	$(4,690 \times 10^{1})$	5,827
ν	∞	22,172		20,835		7,316

5.1.5 O efeito da má especificação de ν nos estimadores bayesianos - Cenário 5

As Tabelas 5.13 e 5.14 mostram a mediana e o erro quadrático médio das médias e medianas *a posteriori*, respectivamente, dos parâmetros do modelo proposto ajustado sob diferentes especificações do grau de liberdade ν. Diferentemente do estudo apresentado na Seção 5.1.4, aqui ν será assumido conhecido, ou seja, assume-se uma distribuição *a priori* degenerada para ν nos seguintes valores, $\nu = 2,01$; 5 e ∞.

Tabela 5.13: Mediana e erro quadrático médio (em parênteses) para as médias *a posteriori* dos parâmetros do modelo t-Tobit com erro nas covariáveis em função da especificação do grau de liberdade ν.

ν informado			2,01		5		∞	
ν real	Parâmetro	Real	mediana	mse	mediana	mse	mediana	mse
2,01	β_1	-6	-6,043	$(3,857)$	-6,012	$(4,338)$	-5,727	$(2,046 \times 10^{1})$
	$\boldsymbol{\beta}_{21}$	0,6	0,599	$(1,191 \times 10^{-3})$	0,592	$(1,322 \times 10^{-3})$	0,580	$(1,693 \times 10^{-2})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,305	$(4,027 \times 10^{-3})$	-0,298	$(4,281 \times 10^{-3})$	-0,295	$(3,615 \times 10^{-2})$
	$\boldsymbol{\beta}_{31}$	0,5	0,503	$(2,293 \times 10^{-2})$	0,507	$(2,560 \times 10^{-2})$	0,534	$(8,025 \times 10^{-2})$
	$\boldsymbol{\beta}_{32}$	2	1,943	$(1,129)$	1,987	$(1,236)$	2,165	$(3,471)$
	σ_u	18	18,717	$(1,637 \times 10^{1})$	24,228	$(7,588 \times 10^{1})$	50,494	$(3,520 \times 10^{5})$
	$\boldsymbol{\mu}_{\xi_1}$	20	20,003	$(1,546)$	19,915	$(1,752)$	20,052	$(1,535 \times 10^{1})$
	$\boldsymbol{\mu}_{\xi_2}$	1	0,973	$(3,562 \times 10^{-1})$	0,963	$(4,104 \times 10^{-1})$	0,892	$(5,910)$
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	207,477	$(1,495 \times 10^{3})$	284,799	$(1,322 \times 10^{4})$	1081,039	$(4,199 \times 10^{8})$
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,447	$(7,211 \times 10^{1})$	0,334	$(1,603 \times 10^{2})$	-0,621	$(2,044 \times 10^{8})$
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	65,701	$(3,267 \times 10^{2})$	86,187	$(1,549 \times 10^{3})$	274,590	$(1,042 \times 10^{8})$
5	β_1	-6	-5,958	$(3,671)$	-6,034	$(3,357)$	-5,751	$(4,666)$
	$\boldsymbol{\beta}_{21}$	0,6	0,606	$(1,195 \times 10^{-3})$	0,604	$(1,143 \times 10^{-3})$	0,595	$(1,974 \times 10^{-3})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,298	$(3,722 \times 10^{-3})$	-0,294	$(3,663 \times 10^{-3})$	-0,297	$(6,084 \times 10^{-3})$
	$\boldsymbol{\beta}_{31}$	0,5	0,477	$(2,160 \times 10^{-2})$	0,488	$(2,022 \times 10^{-2})$	0,494	$(2,723 \times 10^{-2})$
	$\boldsymbol{\beta}_{32}$	2	1,986	$(1,101)$	2,045	$(1,035)$	2,111	$(1,220)$
	σ_u	18	15,581	$(1,347 \times 10^{1})$	18,248	$(1,048 \times 10^{1})$	25,782	$(1,190 \times 10^{2})$
	$\boldsymbol{\mu}_{\xi_1}$	20	20,091	$(1,378)$	20,060	$(1,310)$	19,997	$(1,667)$
	$\boldsymbol{\mu}_{\xi_2}$	1	1,020	$(3,810 \times 10^{-1})$	1,052	$(3,566 \times 10^{-1})$	1,022	$(4,429 \times 10^{-1})$
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	162,194	$(5,731 \times 10^{2})$	194,066	$(6,850 \times 10^{2})$	297,542	$(1,745 \times 10^{4})$
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,869	$(4,212 \times 10^{1})$	0,781	$(5,775 \times 10^{1})$	0,874	$(3,120 \times 10^{2})$
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	54,727	$(5,167 \times 10^{1})$	62,798	$(2,077 \times 10^{2})$	89,974	$(2,927 \times 10^{3})$
∞	β_1	-6	-5,992	$(3,886)$	-6,132	$(3,436)$	-5,954	$(3,053)$
	$\boldsymbol{\beta}_{21}$	0,6	0,603	$(1,106 \times 10^{-3})$	0,603	$(9,820 \times 10^{-4})$	0,599	$(8,670 \times 10^{-4})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,297	$(3,383 \times 10^{-3})$	-0,298	$(3,018 \times 10^{-3})$	-0,294	$(2,707 \times 10^{-3})$
	$\boldsymbol{\beta}_{31}$	0,5	0,494	$(2,405 \times 10^{-2})$	0,501	$(2,129 \times 10^{-2})$	0,495	$(2,006 \times 10^{-2})$
	$\boldsymbol{\beta}_{32}$	2	1,954	$(8,007 \times 10^{-1})$	1,984	$(7,299 \times 10^{-1})$	1,955	$(6,740 \times 10^{-1})$
	σ_u	18	13,308	$(2,660 \times 10^{1})$	14,475	$(1,748 \times 10^{1})$	17,377	$(8,053)$
	$\boldsymbol{\mu}_{\xi_1}$	20	20,101	$(1,354)$	20,022	$(1,182)$	20,045	$(1,061)$
	$\boldsymbol{\mu}_{\xi_2}$	1	0,963	$(3,439 \times 10^{-1})$	0,977	$(3,069 \times 10^{-1})$	0,967	$(2,895 \times 10^{-1})$
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	140,468	$(1,745 \times 10^{3})$	154,763	$(8,801 \times 10^{2})$	186,976	$(4,092 \times 10^{2})$
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,193	$(2,799 \times 10^{1})$	0,157	$(3,043 \times 10^{1})$	0,326	$(4,069 \times 10^{1})$
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	47,961	$(1,788 \times 10^{1})$	51,201	$(1,730 \times 10^{1})$	58,589	$(9,648 \times 10^{1})$

Das Tabelas 5.13 e 5.14 pode-se notar que, em geral, melhores resultados foram obtidos

Tabela 5.14: Mediana e erro quadrático médio (em parênteses) para as medianas *a posteriori* dos parâmetros do modelo t-Tobit com erro nas covariáveis em função da especificação do grau de liberdade ν.

ν informado			2,01		5		∞	
ν real	Parâmetro	Real	mediana	mse	mediana	mse	mediana	mse
2,01	β_1	-6	-6,051	$(3,869)$	-6,003	$(4,341)$	-5,740	$(2,059 \times 10^{1})$
	$\boldsymbol{\beta}_{21}$	0,6	0,598	$(1,191 \times 10^{-3})$	0,592	$(1,329 \times 10^{-3})$	0,581	$(1,692 \times 10^{-2})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,303	$(4,057 \times 10^{-3})$	-0,297	$(4,310 \times 10^{-3})$	-0,294	$(3,616 \times 10^{-2})$
	$\boldsymbol{\beta}_{31}$	0,5	0,503	$(2,306 \times 10^{-2})$	0,510	$(2,561 \times 10^{-2})$	0,530	$(8,057 \times 10^{-2})$
	$\boldsymbol{\beta}_{32}$	2	1,959	$(1,129)$	2,005	$(1,240)$	2,140	$(3,492)$
	σ_u	18	18,335	$(1,543 \times 10^{1})$	23,787	$(7,074 \times 10^{1})$	49,855	$(3,508 \times 10^{5})$
	$\boldsymbol{\mu}_{\xi_1}$	20	20,003	$(1,548)$	19,929	$(1,755)$	19,989	$(1,672 \times 10^{1})$
	$\boldsymbol{\mu}_{\xi_2}$	1	0,973	$(3,559 \times 10^{-1})$	0,946	$(4,115 \times 10^{-1})$	0,859	$(6,597)$
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	204,690	$(1,366 \times 10^{3})$	282,467	$(1,268 \times 10^{4})$	1077,069	$(4,195 \times 10^{8})$
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,264	$(6,906 \times 10^{1})$	0,273	$(1,566 \times 10^{2})$	-0,911	$(2,042 \times 10^{8})$
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	65,159	$(3,050 \times 10^{2})$	85,415	$(1,491 \times 10^{3})$	272,483	$(1,039 \times 10^{8})$
5	β_1	-6	-5,917	$(3,691)$	-6,002	$(3,355)$	-5,778	$(4,697)$
	$\boldsymbol{\beta}_{21}$	0,6	0,606	$(1,184 \times 10^{-3})$	0,603	$(1,154 \times 10^{-3})$	0,595	$(1,984 \times 10^{-3})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,298	$(3,728 \times 10^{-3})$	-0,295	$(3,687 \times 10^{-3})$	-0,296	$(6,123 \times 10^{-3})$
	$\boldsymbol{\beta}_{31}$	0,5	0,482	$(2,148 \times 10^{-2})$	0,488	$(2,046 \times 10^{-2})$	0,498	$(2,734 \times 10^{-2})$
	$\boldsymbol{\beta}_{32}$	2	1,997	$(1,106)$	2,060	$(1,033)$	2,091	$(1,226)$
	σ_u	18	15,297	$(1,442 \times 10^{1})$	17,971	$(1,024 \times 10^{1})$	25,616	$(1,141 \times 10^{2})$
	$\boldsymbol{\mu}_{\xi_1}$	20	20,093	$(1,395)$	20,053	$(1,317)$	20,023	$(1,684)$
	$\boldsymbol{\mu}_{\xi_2}$	1	1,023	$(3,878 \times 10^{-1})$	1,048	$(3,603 \times 10^{-1})$	1,026	$(4,434 \times 10^{-1})$
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	160,733	$(6,221 \times 10^{2})$	192,466	$(6,378 \times 10^{2})$	295,514	$(1,693 \times 10^{4})$
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,776	$(4,081 \times 10^{1})$	1,024	$(5,653 \times 10^{1})$	1,092	$(3,058 \times 10^{2})$
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	54,245	$(4,623 \times 10^{1})$	62,158	$(1,952 \times 10^{2})$	89,253	$(2,851 \times 10^{3})$
∞	β_1	-6	-5,954	$(3,901)$	-6,105	$(3,433)$	-5,974	$(3,080)$
	$\boldsymbol{\beta}_{21}$	0,6	0,604	$(1,109 \times 10^{-3})$	0,603	$(9,826 \times 10^{-4})$	0,599	$(8,710 \times 10^{-4})$
	$\boldsymbol{\beta}_{22}$	-0,3	-0,297	$(3,399 \times 10^{-3})$	-0,297	$(3,027 \times 10^{-3})$	-0,294	$(2,706 \times 10^{-3})$
	$\boldsymbol{\beta}_{31}$	0,5	0,495	$(2,416 \times 10^{-2})$	0,502	$(2,118 \times 10^{-2})$	0,498	$(2,020 \times 10^{-2})$
	$\boldsymbol{\beta}_{32}$	2	1,981	$(8,064 \times 10^{-1})$	1,982	$(7,256 \times 10^{-1})$	1,972	$(6,777 \times 10^{-1})$
	σ_u	18	13,152	$(2,824 \times 10^{1})$	14,295	$(1,857 \times 10^{1})$	17,277	$(8,154)$
	$\boldsymbol{\mu}_{\xi_1}$	20	20,093	$(1,353)$	20,023	$(1,187)$	20,047	$(1,056)$
	$\boldsymbol{\mu}_{\xi_2}$	1	0,975	$(3,439 \times 10^{-1})$	0,991	$(3,083 \times 10^{-1})$	0,964	$(2,900 \times 10^{-1})$
	$\boldsymbol{\Sigma}_{\xi_{11}}$	180	139,302	$(1,845 \times 10^{3})$	153,051	$(9,421 \times 10^{2})$	185,607	$(3,881 \times 10^{2})$
	$\boldsymbol{\Sigma}_{\xi_{12}}$	-0,1	0,152	$(2,743 \times 10^{1})$	0,019	$(2,961 \times 10^{1})$	0,438	$(3,973 \times 10^{1})$
	$\boldsymbol{\Sigma}_{\xi_{22}}$	50	47,577	$(1,954 \times 10^{1})$	50,848	$(1,667 \times 10^{1})$	58,173	$(9,042 \times 10^{1})$

quando os valores especificados para o grau de liberdade ν são os valores reais, utilizados para geração das amostras. Os piores resultados foram observados quando as amostras foram geradas sob uma distribuição com cauda pesada e ajusta-se a elas o modelo Tobit normal com erros nas covariáveis. Para essas amostras o ideal é que seja informado um valor de ν igual ao valor real pois, assim, tem-se os menores erros quadráticos médio. Quando amostras são geradas considerando-se $\nu = 5$ nota-se, em geral, o mesmo comportamento. Nesse caso, porém, observa-se que, para $\mathbf{\Sigma}_{\xi}$, os melhores resultados surgem quando é informado que o grau de liberdade é igual a $2,01$. Porém, para a média *a posteriori* e observando apenas a estimativa pontual obtida para $\mathbf{\Sigma}_{\xi_{12}}$, o melhor resultado foi obtido quando um valor para ν igual ao valor real, 5, é assumido. Sob esse valor usado para geração das amostras as piores estimativas surgiram quando foi ajustado o modelo que considera $\nu = \infty$. Quando os dados são gerados sob o modelo Tobit normal com erros nas covariáveis as melhores estimativas, em geral, são obtidas quando ajusta-se o modelo correto, exceto para $\mathbf{\Sigma}_{\xi_{12}}$ e $\mathbf{\Sigma}_{\xi_{22}}$, que apresentam menores erros quadráticos médio quando se assume ν iguais a $2,01$ e 5, respectivamente, sendo que as melhores estimativas pontuais são observadas quando o número do grau de liberdade é igual a 5. Para esses dois parâmetros as piores estimativas foram obtidas exatamente quando foi ajustado o modelo correto às amostras.

Se o modelo proposto é assumindo com grau de liberdade especificado em $2,01$ observa-se que, exceto para $\mathbf{\Sigma}_{\xi}$, as estimativas dos parâmetros são aproximadamente iguais para todos os valores considerados nas gerações das amostras. Para $\mathbf{\Sigma}_{\xi_{12}}$ e $\mathbf{\Sigma}_{\xi_{22}}$ os valores mais próximos dos reais foram obtidos quando o modelo real ao gerar os dados era o Tobit normal com erros nas covariáveis, enquanto, para $\mathbf{\Sigma}_{\xi_{11}}$ os melhores resultados foram observados quando o grau de liberdade considerado na geração das amostras é igual 5. Quando o modelo proposto com ν igual a 5 é ajustado aos dados, os melhores resultados foram obtidos quando os valores reais de ν eram iguais a 5 ou ∞, ou seja, quando amostras foram geradas do modelo Tobit normal com erros nas covariáveis. Ao ajustar-se às amostras o modelo mais parcimonioso, as melhores estimativas foram observadas quando, de fato, os dados vieram de tal modelo.

A Tabela 5.15 apresenta a mediana do desvio-padrão *a posteriori* dos parâmetros do modelo proposto ajustado sob diferentes especificações do grau de liberdade ν.

Tabela 5.15: Mediana para os desvios-padrão *a posteriori* dos parâmetros do modelo t-Tobit com erro nas covariáveis em função da especificação do grau de liberdade ν.

ν informado		2,01	5	∞
ν real	Parâmetro	mediana	mediana	mediana
2,01	β_1	1,950	2,117	3,102
	$\boldsymbol{\beta}_{21}$	0,036	0,035	0,030
	$\boldsymbol{\beta}_{22}$	0,063	0,060	0,053
	$\boldsymbol{\beta}_{31}$	0,150	0,165	0,250
	$\boldsymbol{\beta}_{32}$	1,025	1,133	1,687
	σ_u	3,855	4,702	11,481
	$\boldsymbol{\mu}_{\xi_1}$	1,218	1,363	2,417
	$\boldsymbol{\mu}_{\xi_2}$	0,680	0,746	1,244
	$\boldsymbol{\Sigma}_{\xi_{11}}$	28,491	35,871	106,456
	$\boldsymbol{\Sigma}_{\xi_{12}}$	9,626	12,451	40,468
	$\boldsymbol{\Sigma}_{\xi_{22}}$	8,615	10,519	27,556
5	β_1	1,805	1,841	2,003
	$\boldsymbol{\beta}_{21}$	0,035	0,033	0,031
	$\boldsymbol{\beta}_{22}$	0,062	0,059	0,053
	$\boldsymbol{\beta}_{31}$	0,136	0,141	0,157
	$\boldsymbol{\beta}_{32}$	0,929	0,973	1,058
	σ_u	3,066	3,298	4,105
	$\boldsymbol{\mu}_{\xi_1}$	1,096	1,136	1,274
	$\boldsymbol{\mu}_{\xi_2}$	0,627	0,641	0,698
	$\boldsymbol{\Sigma}_{\xi_{11}}$	21,645	23,152	29,767
	$\boldsymbol{\Sigma}_{\xi_{12}}$	7,659	8,567	11,508
	$\boldsymbol{\Sigma}_{\xi_{22}}$	6,829	7,235	9,011
∞	β_1	1,669	1,666	1,655
	$\boldsymbol{\beta}_{21}$	0,034	0,032	0,030
	$\boldsymbol{\beta}_{22}$	0,062	0,059	0,053
	$\boldsymbol{\beta}_{31}$	0,127	0,127	0,127
	$\boldsymbol{\beta}_{32}$	0,861	0,860	0,863
	σ_u	2,568	2,559	2,676
	$\boldsymbol{\mu}_{\xi_1}$	1,023	1,024	1,006
	$\boldsymbol{\mu}_{\xi_2}$	0,578	0,575	0,566
	$\boldsymbol{\Sigma}_{\xi_{11}}$	18,288	18,250	18,610
	$\boldsymbol{\Sigma}_{\xi_{12}}$	6,559	6,783	7,321
	$\boldsymbol{\Sigma}_{\xi_{22}}$	5,806	5,726	5,816

Da Tabela 5.15 nota-se que maiores diferenças nos desvios-padrão *a posteriori* para cada grau de liberdade considerado na geração das amostras foram obtidas quando o grau de liberdade verdadeiro é igual a $2,01$. Os maiores valores surgiram quando foi ajustado o modelo proposto por Wang (1998). Observa-se também que, quando amostras foram geradas sob o modelo proposto com grau de liberdade igual a 5 e este mesmo número de grau de liberdade foi assumido no modelo, os desvios-padrão *a posteriori* dos parâmetros são sempre maiores do que quando os modelos são ajustados com outros graus de liberdade. Quando considera-se os modelos propostos nas amostras geradas sob o modelo Tobit normal com erros nas covariáveis as distribuições marginais *a posteriori* dos parâmetros tem, em geral, menor variabilidade quando o modelo real é ajustado.

A Tabela 5.16 apresenta as medianas e médias das estatísticas LPML e DIC para os modelos ajustados sob diferentes especificações dos graus de liberdade ν e a Tabela 5.17 relaciona o número de vezes que cada modelo foi escolhido de acordo com os valores de LPML e DIC para cada valor especificado para ν.

Tabela 5.16: Medianas e médias dos valores obtidos de LPML e DIC para o modelo t-Tobit com erro nas covariáveis em função da especificação do grau de liberdade ν.

ν informado		2,01		5		∞	
ν real	Estatística	mediana	média	mediana	média	mediana	média
2,01	LPML	-2282,259	-2284,111	-2303,993	-2305,512	-2501,062	-2527,287
	DIC	4564,560	4567,924	4606,853	4609,944	4972,107	5018,340
5	LPML	-2168,083	-2168,025	-2154,364	-2154,225	-2181,829	-2185,024
	DIC	4335,412	4335,691	4308,574	4308,200	4359,496	4363,589
∞	LPML	-2104,552	-2104,537	-2075,542	-2075,690	-2062,211	-2062,002
	DIC	4208,534	4208,592	4151,352	4151,484	4123,929	4123,843

Tabela 5.17: Número de vezes que cada modelo foi escolhido de acordo com o LPML (DIC) para cada especificação do grau de liberdade ν.

Modelo real	Modelo escolhido		
ν	2,01	5	∞
2,01	495 (495)	5 (5)	0 (0)
5	2 (2)	492 (490)	6 (8)
∞	0 (0)	0 (0)	500 (500)

Pela Tabela 5.16 observa-se que, para cada grau de liberdade utilizado na geração das amostras, os valores obtidos pelas medidas LPML e DIC indicam que, em termos medianos, os modelos ajustados com graus de liberdade iguais aos considerados na geração

dos dados são os mais adequados aos dados. Da Tabela 5.17 conclui-se que tanto o DIC quanto o LPML tendem a selecionar corretamente o modelo (acima de 98% das vezes) e, nas raras vezes que não o fazem, mais frequentemente selecionam um modelo que assume um número de graus de liberdade maior que o real.

Na próxima seção será apresentado um estudo Monte Carlo para avaliar a qualidade dos estimadores de máxima verossimilhança no modelo proposto quando o algoritmo ECM, proposto no Capítulo 4, é considerado.

5.2 Estimadores de máxima verossimilhança

Em todos os cenários considerados nese estudo são geradas 5.000 réplicas do modelo proposto com uma covariável medida com erro e uma covariável medida sem erro e assumindo $\beta_1 = -6$, $\boldsymbol{\beta}_2 = 0,6$, $\boldsymbol{\beta}_3 = 0,5$, $\sigma_u = \boldsymbol{\Sigma}_v = 18$ e $\boldsymbol{\mu}_\xi = 20$. Assume-se $\boldsymbol{\Delta} = 1/10$, logo, $\boldsymbol{\Sigma}_\xi = 10\boldsymbol{\Sigma}_v = 180$. Para a covariável medida sem erro $\mathbf{b}_i$ considera-se que $\mathbf{b}_i \sim N(10, 3^2)$, $i = 1, \ldots, n$.

As amostras possuem tamanho $n = 200$ para todos os cenários, exceto para o Cenário 2, onde são consideradas amostras de tamanho 25, 100, 200 e 500. No Cenário 1, os dados são gerados assumindo que 5%, 25% e 45% das respostas são censuradas. A proporção de censura nos Cenários 2 e 3 é de aproximadamente $14,04\%$ e para o Cenário 2 também consideram-se amostras com 45% de respostas censuradas. Os dados são gerados assumindo $\nu = 5$ nos Cenários 1, 2 e 3. No Cenário 4, as amostras são geradas do modelo proposto assumindo $\nu = 2,01$, 5, 30 e ∞ para os quais são observadas, aproximadamente, $17,49\%$, $14,04\%$, $11,81\%$ e $11,25\%$ de observações censuradas. Os mesmos graus de liberdade foram assumidos para gerar os dados no Cenário 5, com exceção de $\nu = 30$. No Cenário 3 o modelo proposto é ajustado assumindo que a condição de identificabilidade $\boldsymbol{\Delta}$ esteja fixada em 1/19, 1/10, 1/3 e 1. Para o Cenário 5, o modelo proposto é ajustado assumindo $\nu = 2,01$, 5 e 30.

O algoritmo ECM foi iniciado com os valores reais dos parâmetros e iterados até que fosse obtido $|\log f(\mathbf{D}_o \mid \widehat{\boldsymbol{\theta}}^{(t+1)})/\log f(\mathbf{D}_o \mid \widehat{\boldsymbol{\theta}}^{(t)}) - 1| < 10^{-8}$. Em média, para o Cenário 4, 79,39, 95,24, 182,55 e 193,81 iterações foram necessárias para a convergência do algoritmo quando amostras foram geradas do modelo proposto com, respectivamente,

$2,01$, 5, 30 e ∞ graus de liberdade. Em geral, a convergência do algoritmo ECM é lenta para cenários com alta proporção de censura e quando pequenas amostras são consideradas. Por exemplo, em média, para o Cenário 1 a convergência é alcançada após $335,74$ e $77,95$ iterações, respectivamente, se os percentuais de censura são 45% e 5%. Para o Cenário 2 com $14,04\%$ de censura, em média, $168,32$ passos do algoritmo são necessários para atingir a convergência se $n = 25$ e apenas $84,80$ passos são necessários se amostras de tamanho 500 são consideradas. A convergência do algoritmo tende a ser rápida se os graus de liberdade são fixos. Em média, para o Cenário 5, o número de passos necessários foi menor que $85,58$ em todos os casos.

Os intervalos de confiança para os estimadores também são obtidos com o uso de propriedades assintóticas dos estimadores de máxima verossimilhança e aproximando-se os erros-padrão considerando-se a matriz de informação de Fisher observada, dada na Seção 4.2.

5.2.1 O efeito de diferentes proporções de censura nos EMV - Cenário 1

A Figura 5.1 apresenta o gráfico do vício (*Bias*) e da raiz do erro quadrático médio (*RMSE*) dos estimadores de máxima verossimilhança em relação às diferentes proporções de censura das amostras.

Nota-se, pela Figura 5.1, que os vícios dos estimadores de $\boldsymbol{\beta}_2$ e $\boldsymbol{\beta}_3$ são próximos de zero para todas as proporções de censura consideradas. Para os estimadores de σ_u e $\boldsymbol{\Sigma}_\xi$ o vício aumenta, em valores absolutos, à medida que a proporção de censura aumenta. Para $\boldsymbol{\mu}_\xi$ e ν o vício se aproxima de zero quando a proporção de censura aumenta. As raízes dos erros quadráticos médios não são fortemente afetadas pela proporção de censura, exceto para β_1 e $\boldsymbol{\Sigma}_\xi$, onde maiores valores são observados à medida que são consideradas amostras com maiores proporções de observações censuradas.

A Figura 5.2 mostra a probabilidade de cobertura dos intervalos de 95% de confiança para os parâmetros do modelo proposto no Cenário 1.

Observa-se, a partir da Figura 5.2, que as probabilidades de cobertura estão bem próximas dos verdadeiros níveis de confiança para todos os parâmetros.

A Figura 5.3 mostra as distribuições dos estimadores de máxima verossimilhança dos

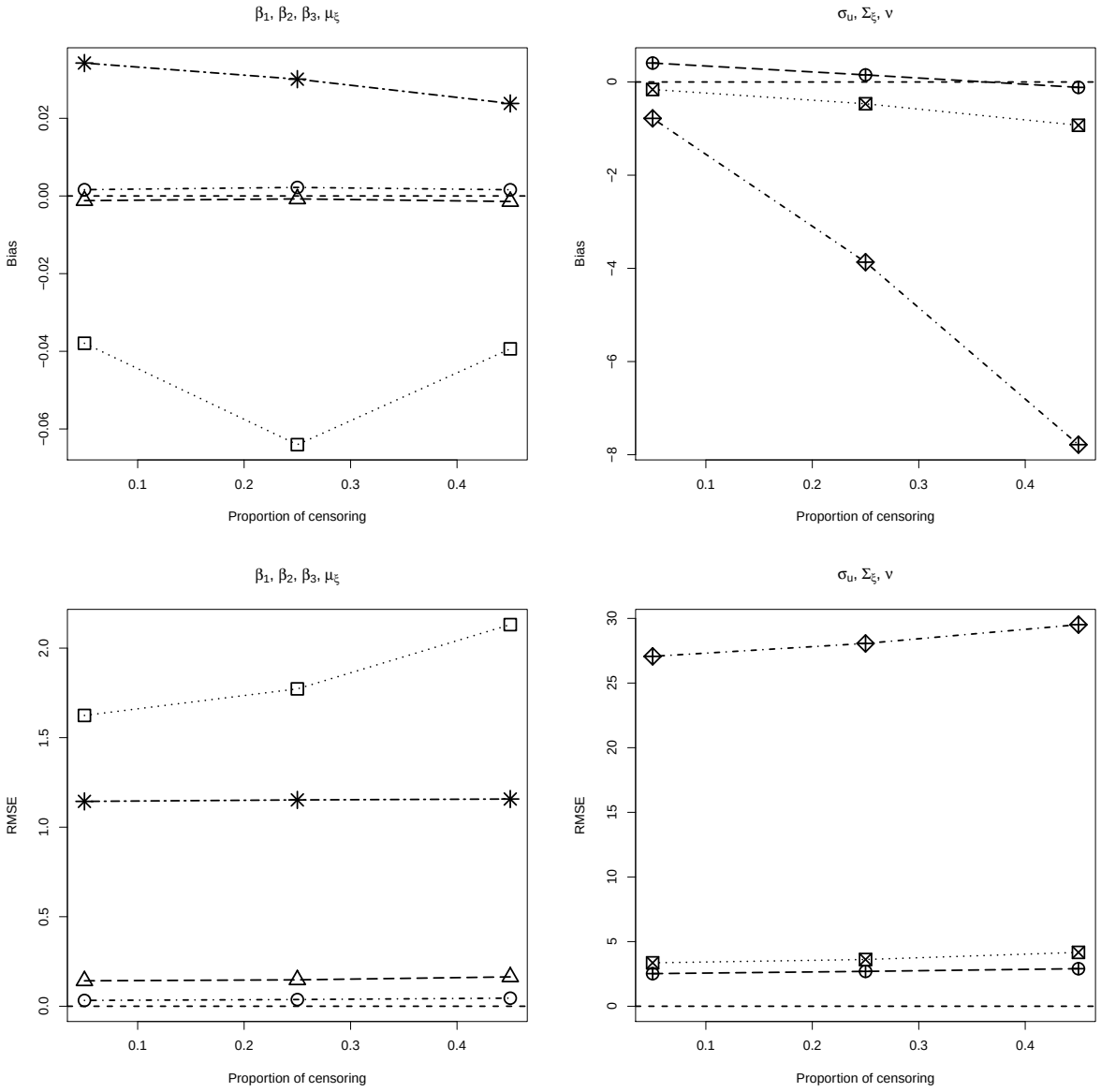

Figura 5.1: Vício (superior) e raiz do erro quadrático médio (inferior) para os estimadores de β_1 (quadrado), $\boldsymbol{\beta}_2$ (círculo), $\boldsymbol{\beta}_3$ (triângulo), $\boldsymbol{\mu}_\xi$ (asterisco) (gráficos à esquerda), σ_u (quadrado), $\boldsymbol{\Sigma}_\xi$ (losango) e ν (círculo) (gráficos à direita) quando amostras são geradas com 5%, 25% e 45% de respostas censuradas.

parâmetros para cada proporção de censura considerada nas amostras.

Da Figura 5.3 nota-se que, em geral, as distribuições dos estimadores possuem maiores variabilidades para os casos onde as proporções de censuras são maiores. Para os parâmetros relacionados à variância da distribuição t-Student, σ_u, $\boldsymbol{\Sigma}_\xi$ e ν, mais de 50% dos valores obtidos pelos seus estimadores estavam subestimados e a subestimação é ainda maior quando as amostras possuem maiores porcentagens de censura.

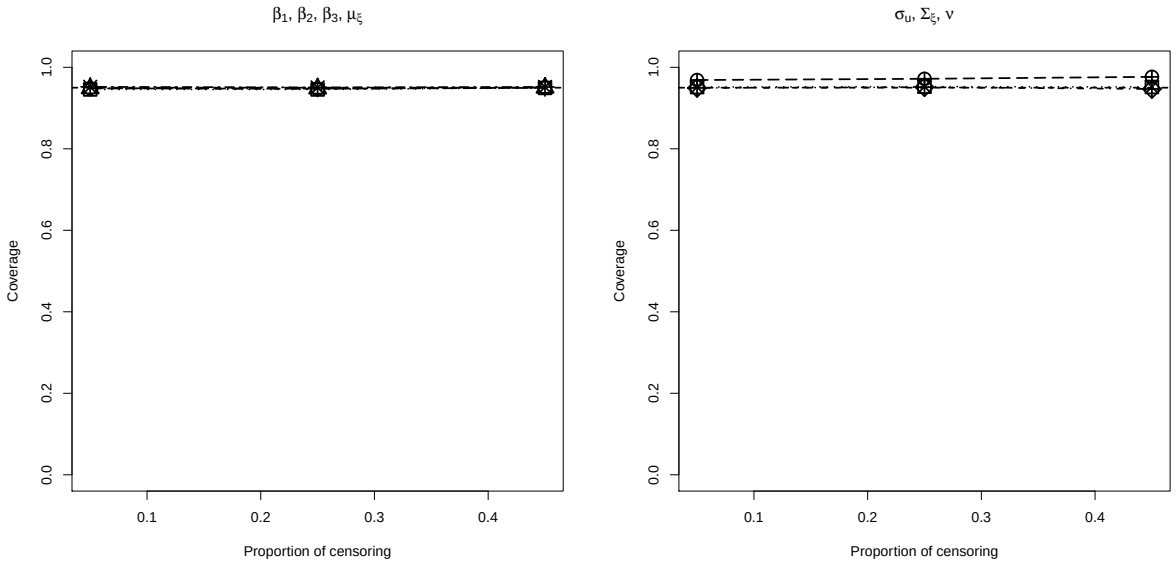

Figura 5.2: Cobertura dos intervalos com 95% de confiança para β_1 (quadrado), $\boldsymbol{\beta}_2$ (círculo), $\boldsymbol{\beta}_3$ (triângulo), $\boldsymbol{\mu}_\xi$ (asterisco) (gráfico à esquerda), σ_u (quadrado), $\boldsymbol{\Sigma}_\xi$ (losango) e ν (círculo) (gráfico à direita) quando amostras são geradas com 5%, 25% e 45% de respostas censuradas.

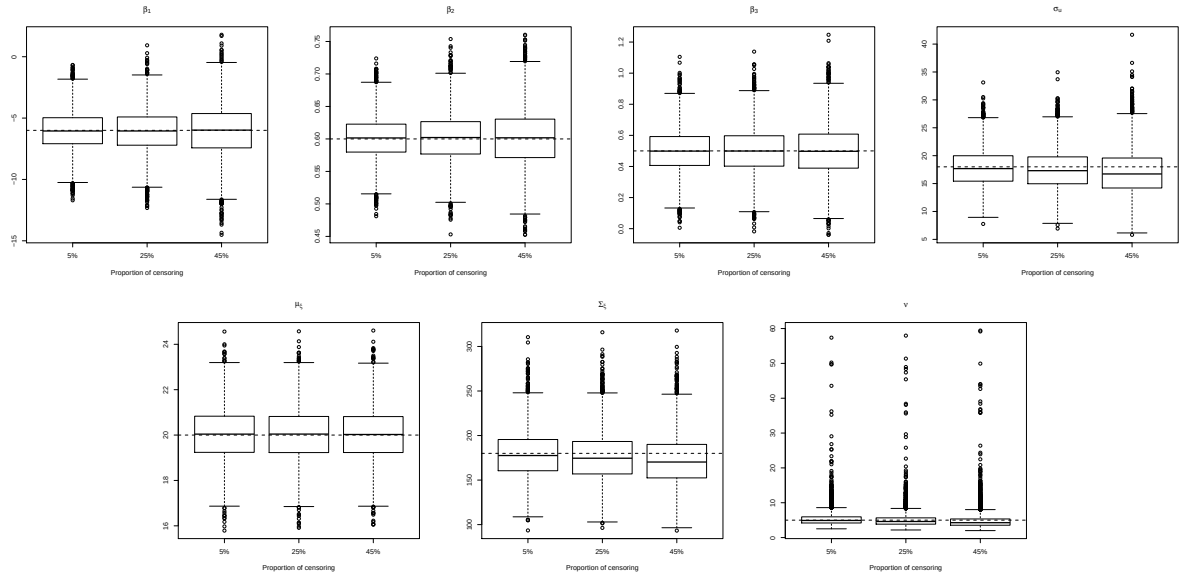

Figura 5.3: Distribuição dos estimadores de máxima verossimilhança de β_1, $\boldsymbol{\beta}_2$, $\boldsymbol{\beta}_3$, σ_u, $\boldsymbol{\mu}_\xi$, $\boldsymbol{\Sigma}_\xi$ e ν quando amostras são geradas com 5%, 25% e 45% de respostas censuradas.

5.2.2 O efeito de diferentes tamanhos amostrais nos EMV - Cenário 2

As Figuras 5.4 e 5.5 apresentam os gráficos dos vícios (*Bias*) e das raízes dos erros quadráticos médio (*RMSE*) dos estimadores de máxima verossimilhança em relação aos

diferentes tamanhos amostrais para amostras com aproximadamente $14,04\%$ e 45% de censura, respectivamente.

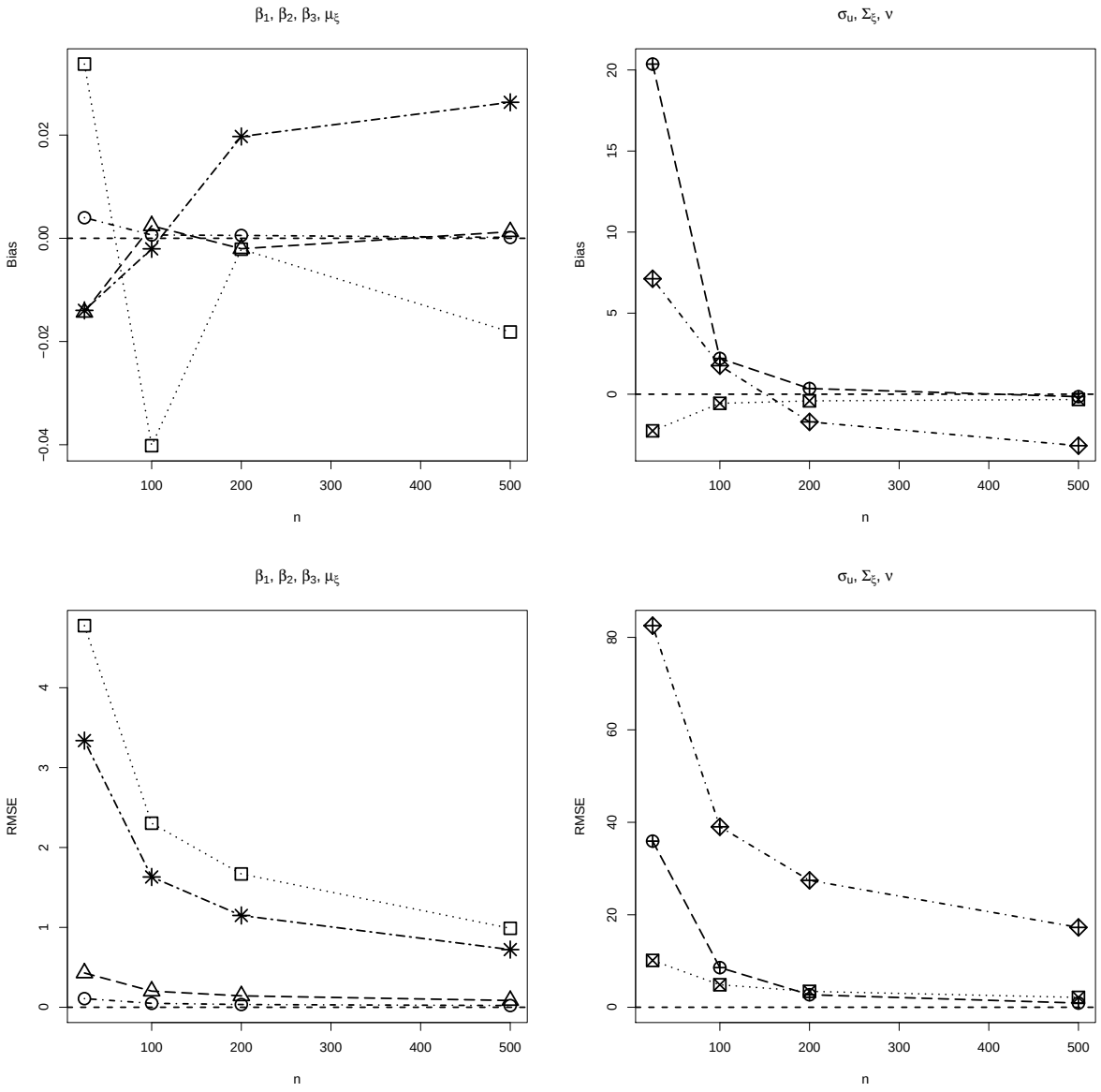

Figura 5.4: Vício (superior) e raiz do erro quadrático médio (inferior) para os estimadores de β_1 (quadrado), $\boldsymbol{\beta}_2$ (círculo), $\boldsymbol{\beta}_3$ (triângulo), $\boldsymbol{\mu}_\xi$ (asterisco) (gráficos à esquerda), σ_u (quadrado), $\boldsymbol{\Sigma}_\xi$ (losango) e ν (círculo) (gráficos à direita) em função do tamanho amostral n e com $14,04\%$ de censura.

As Figuras 5.4 e 5.5 mostram que, para amostras com aproximadamente 14% e 45% de censura, os estimadores de $\boldsymbol{\beta}_2$, $\boldsymbol{\beta}_3$, σ_u e ν tendem a ser não viciados e consistentes. No entanto, os vícios dos estimadores de β_1, $\boldsymbol{\mu}_\xi$ e $\boldsymbol{\Sigma}_\xi$ tendem a ser grandes, em módulo, para todos os tamanhos amostrais. Para tais parâmetros o vício e a raiz do erro quadrático

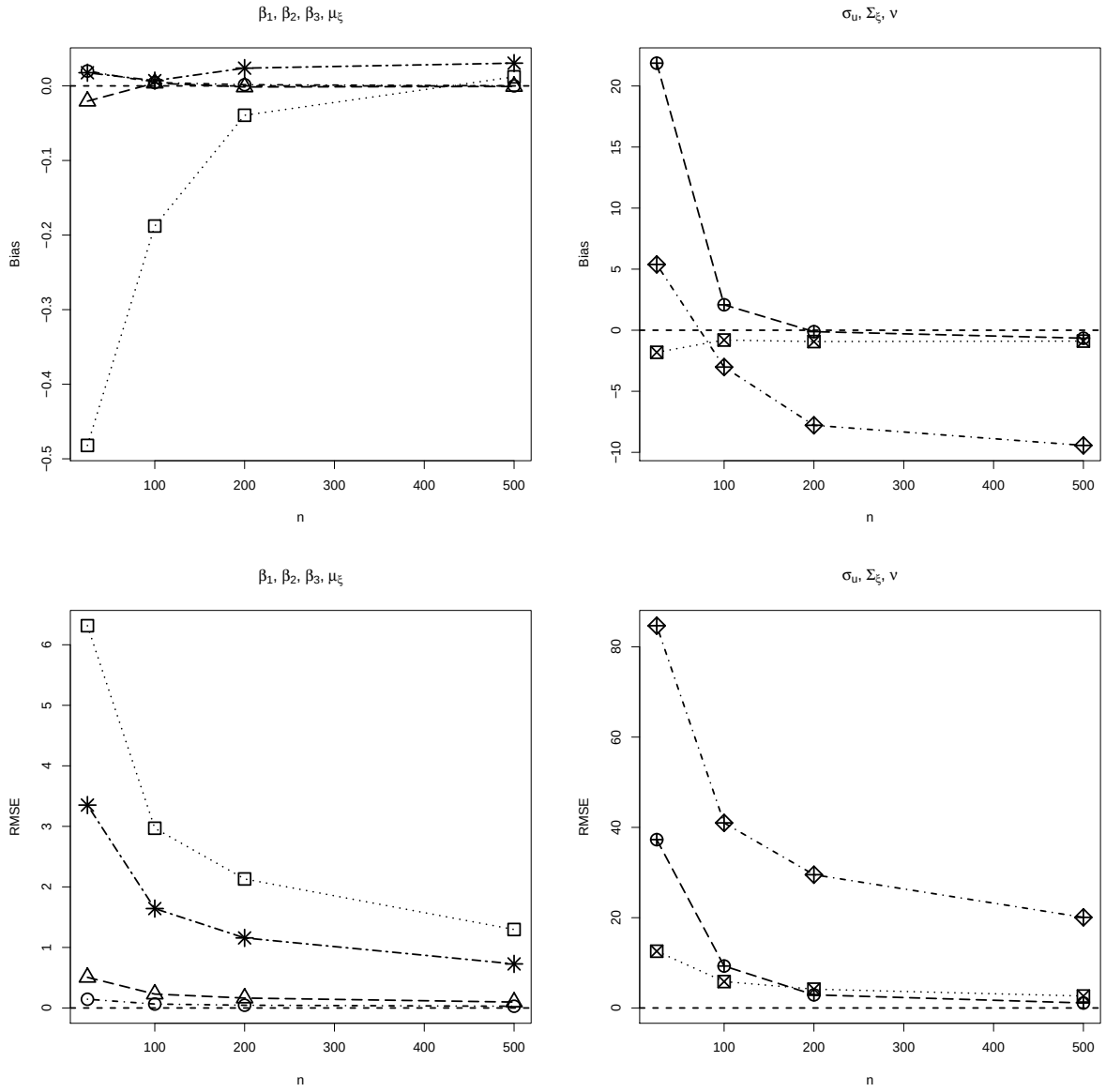

Figura 5.5: Vício (superior) e raiz do erro quadrático médio (inferior) para os estimadores de β_1 (quadrado), $\boldsymbol{\beta}_2$ (círculo), $\boldsymbol{\beta}_3$ (triângulo), $\boldsymbol{\mu}_\xi$ (asterisco) (gráficos à esquerda), σ_u (quadrado), $\boldsymbol{\Sigma}_\xi$ (losango) e ν (círculo) (gráficos à direita) em função do tamanho amostral n e com 45% de censura.

médio não convergem para zero à medida que n aumenta, isto é, os gráficos sugerem que os estimadores desses parâmetros não são consistentes. Comparando as Figuras 5.4 e 5.5 nota-se que se uma amostra de tamanho $n = 25$ com 14% de censura é considerada o valor absoluto do vício do estimador de $\boldsymbol{\Sigma}_\xi$ é muito similar ao observado para uma amostra de tamanho $n = 200$ com 45% de dados censurados. Ambos os casos sugerem que a falta de informação dos dados (pequenas amostras e alta censura) induz a um maior vício nesse

estimador.

As Figuras 5.6 e 5.7 mostram a probabilidade de cobertura dos intervalos de 95% de confiança para os parâmetros do modelo proposto no Cenário 2 quando amostras apresentam $14,04\%$ e 45% de respostas censuradas, respectivamente.

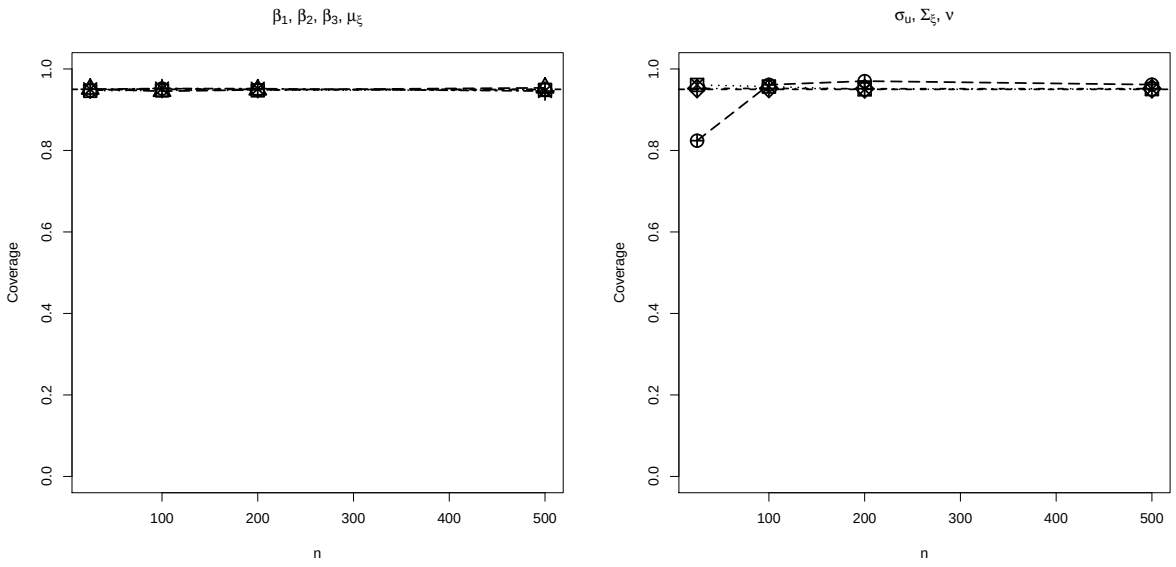

Figura 5.6: Cobertura dos intervalos com 95% de confiança para β_1 (quadrado), $\boldsymbol{\beta}_2$ (círculo), $\boldsymbol{\beta}_3$ (triângulo), $\boldsymbol{\mu}_\xi$ (asterisco) (gráfico à esquerda), σ_u (quadrado), $\boldsymbol{\Sigma}_\xi$ (losango) e ν (círculo) (gráfico à direita) em função do tamanho amostral n e com $14,04\%$ de censura.

Das Figuras 5.6 e 5.7 nota-se que as probabilidades de cobertura estão, em geral, próximas dos verdadeiros níveis de confiança. A exceção ocorre para ν se $n = 25$.

As Figuras 5.8 e 5.9 apresentam as distribuições dos estimadores de máxima verossimilhança dos parâmetros para cada tamanho amostral considerado e em amostras com, respectivamente, $14,04\%$ e 45% de censura.

A partir das Figuras 5.8 e 5.9 pode-se ver que, para todos os parâmetros, as distribuições dos seus respectivos estimadores apresentam menores variabilidades à medida que n aumenta.

5.2.3 O efeito da má especificação de $\boldsymbol{\Delta}$ nos EMV - Cenário 3

A Figura 5.10 apresenta o gráfico dos vícios (*Bias*) e das raízes dos erros quadráticos médio (*RMSE*) dos estimadores de máxima verossimilhança em relação à especificação da condição de identificabilidade $\boldsymbol{\Delta}$.

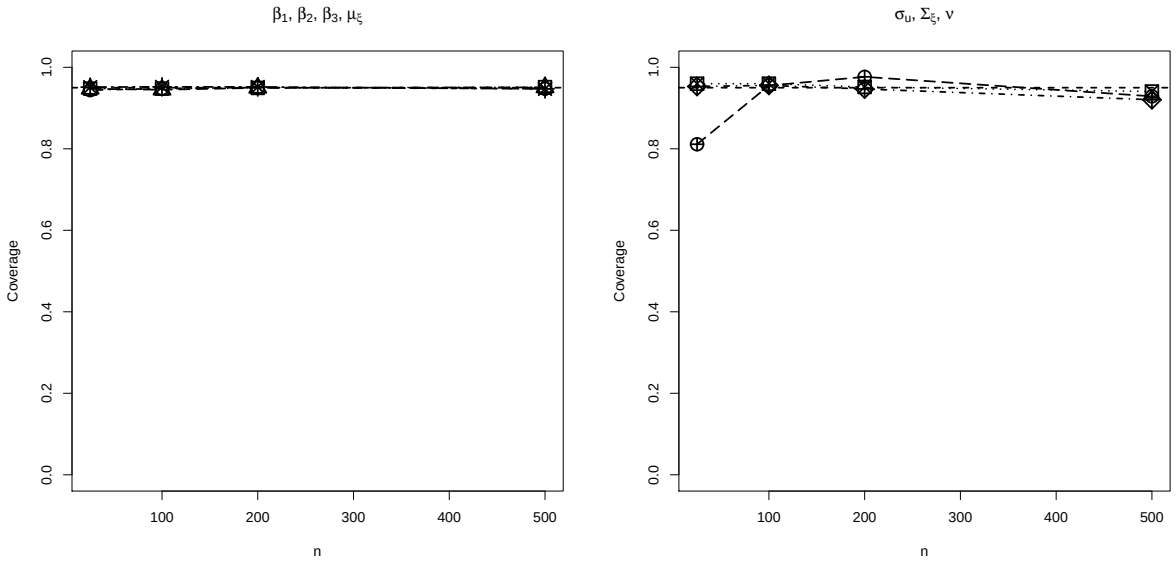

Figura 5.7: Cobertura dos intervalos com 95% de confiança para β_1 (quadrado), $\boldsymbol{\beta}_2$ (círculo), $\boldsymbol{\beta}_3$ (triângulo), $\boldsymbol{\mu}_\xi$ (asterisco) (gráfico à esquerda), σ_u (quadrado), $\boldsymbol{\Sigma}_\xi$ (losango) e ν (círculo) (gráfico à direita) em função do tamanho amostral n e com 45% de censura.

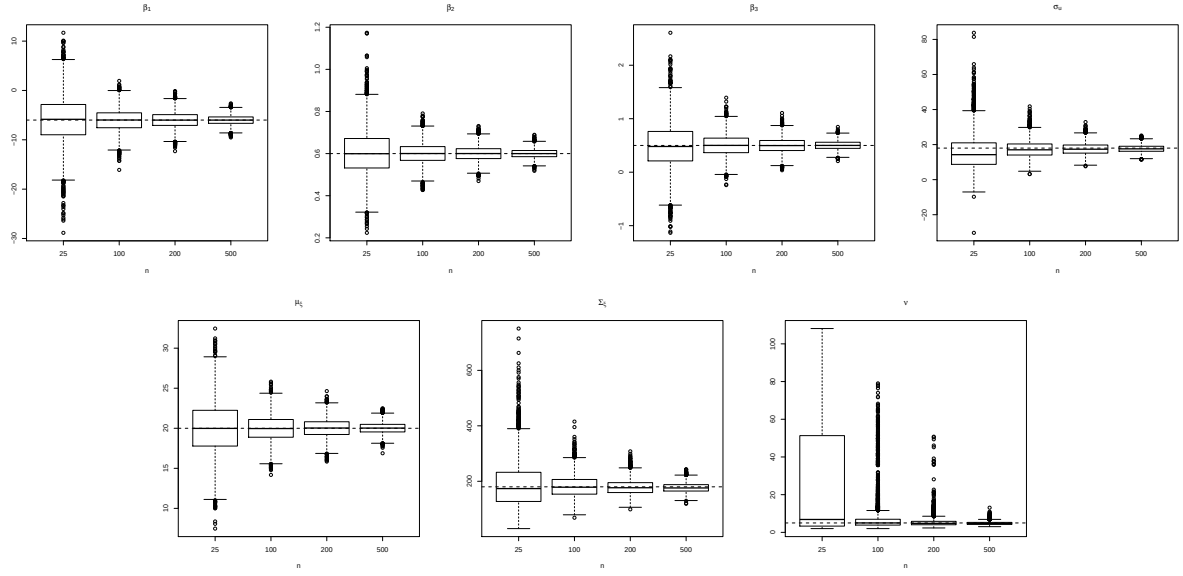

Figura 5.8: Distribuição dos estimadores de máxima verossimilhança de β_1, $\boldsymbol{\beta}_2$, $\boldsymbol{\beta}_3$, σ_u, $\boldsymbol{\mu}_\xi$, $\boldsymbol{\Sigma}_\xi$ e ν em função do tamanho amostral e com $14,04\%$ de censura.

Na Figura 5.10 nota-se que o vício e a raiz do erro quadrático médio dos estimadores de β_1, σ_u e $\boldsymbol{\Sigma}_\xi$ estão ambos distantes de zero se a escolha de $\boldsymbol{\Delta}$ é distante do valor real. Isso significa que a escolha de $\boldsymbol{\Delta}$ tem um importante papel na qualidade das estimativas dos parâmetros. Algumas diretrizes em como escolher um valor apropriado para $\boldsymbol{\Delta}$ podem ser

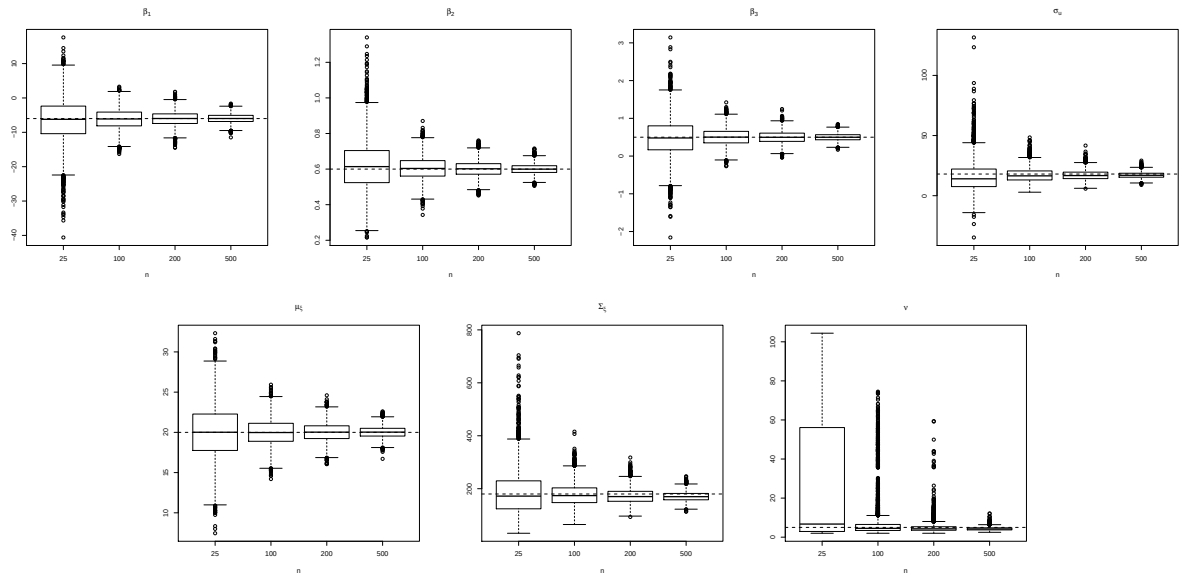

Figura 5.9: Distribuição dos estimadores de máxima verossimilhança de β_1, $\boldsymbol{\beta}_2$, $\boldsymbol{\beta}_3$, σ_u, $\boldsymbol{\mu}_\xi$, $\boldsymbol{\Sigma}_\xi$ e ν em função do tamanho amostral e com 45% de censura.

encontradas em Fuller (1987). Para o estimador de $\boldsymbol{\beta}_2$ não observa-se grandes alterações no vício e raiz do erro quadrático médio com a escolha do valor de $\boldsymbol{\Delta}$. Para os estimadores dos parâmetros restantes já se sabia que não haveriam alterações em suas estimativas e, consequentemente, nos vícios e erros quadrático médios, como mencionado na Seção 3.2.

A Figura 5.11 mostra a probabilidade de cobertura dos intervalos de 95% de confiança para os parâmetros do modelo proposto em relação à especificação de $\boldsymbol{\Delta}$.

Pode-se observar da Figura 5.11 que, escolhendo-se um valor de $\boldsymbol{\Delta}$ distante do valor real, as probabilidades de cobertura dos intervalos de 95% de confiança para β_1, $\boldsymbol{\beta}_2$, σ_u e $\boldsymbol{\Sigma}_\xi$ também estarão distantes do nível nominal considerado para o intervalo.

A Figura 5.12 mostra as distribuições dos EMV de β_1, $\boldsymbol{\beta}_2$, σ_u e $\boldsymbol{\Sigma}_\xi$, que são os parâmetros que são afetados pela escolha do valor especificado de $\boldsymbol{\Delta}$.

A Figura 5.12 revela a importância da escolha da condição de identificabilidade para a matriz $\boldsymbol{\Delta}$. Se o valor de $\boldsymbol{\Delta}$ especificado é menor que o valor real, observa-se que há uma tendência de haver mais estimativas superestimadas para β_1, σ_u e $\boldsymbol{\Sigma}_\xi$. Por outro lado, se $\boldsymbol{\Delta}$ especificado é maior que o real, a tendência é que haja mais estimativas subestimadas para tais parâmetros, ocorrendo o contrário para $\boldsymbol{\beta}_2$. Para o caso em estudo nota-se, também, que assumindo-se $\boldsymbol{\Delta} = 1/3$ ou 1, há um grande percentual de estimativas de σ_u

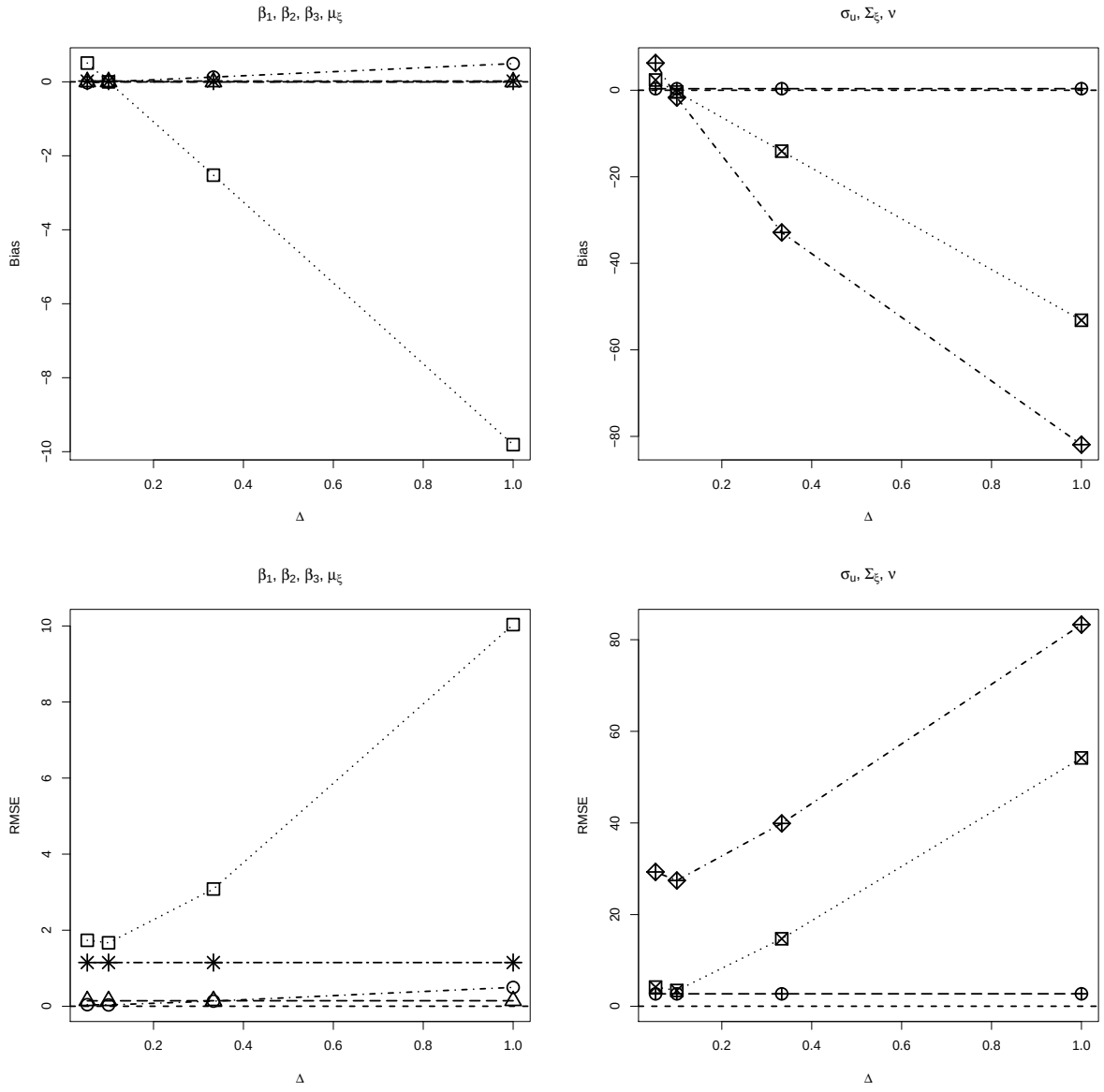

Figura 5.10: Vício (superior) e raiz do erro quadrático médio (inferior) para os estimadores de β_1 (quadrado), $\boldsymbol{\beta}_2$ (círculo), $\boldsymbol{\beta}_3$ (triângulo), $\boldsymbol{\mu}_\xi$ (asterisco) (gráficos à esquerda), σ_u (quadrado), $\boldsymbol{\Sigma}_\xi$ (losango) e ν (círculo) (gráficos à direita) quando amostras são geradas assumindo $\boldsymbol{\Delta} = 0,1$ mas informando que vale 1/19, 1/10, 1/3 e 1.

que são negativas, o que não faz sentido.

5.2.4 O efeito de ν nos EMV - Cenário 4

A Tabela 5.18 mostra o vício, a raiz do erro quadrático médio e a cobertura dos intervalos de 95% de confiança para os estimadores de todos os parâmetros do modelo proposto em cenários em que o grau de liberdade também é estimado.

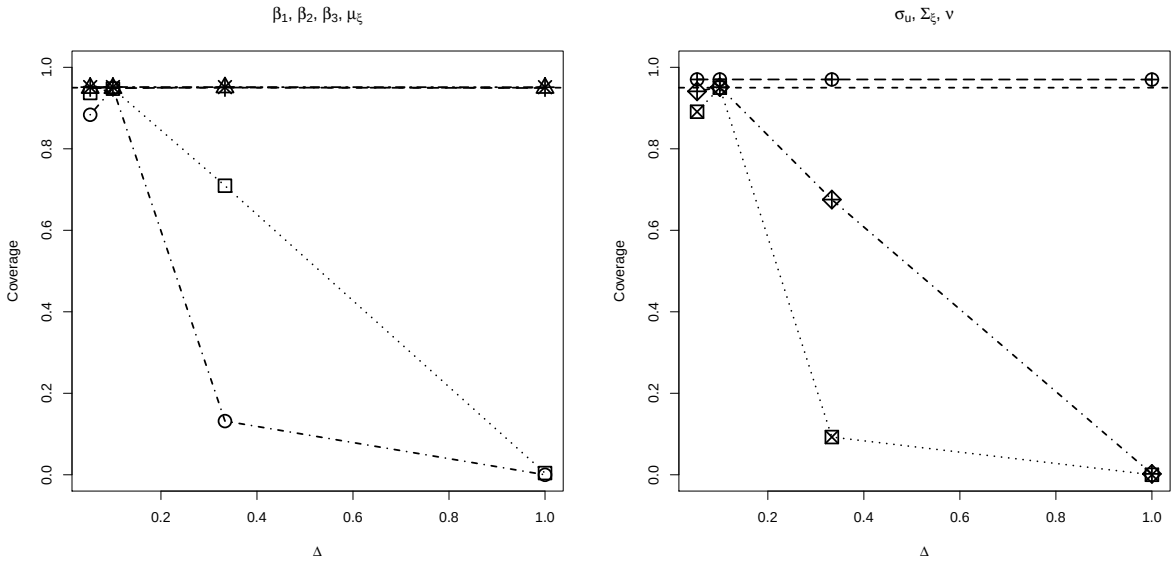

Figura 5.11: Cobertura dos intervalos com 95% de confiança para β_1 (quadrado), $\boldsymbol{\beta}_2$ (círculo), $\boldsymbol{\beta}_3$ (triângulo), $\boldsymbol{\mu}_\xi$ (asterisco) (gráfico à esquerda), σ_u (quadrado), $\boldsymbol{\Sigma}_\xi$ (losango) e ν (círculo) (gráfico à direita) quando amostras são geradas assumindo $\boldsymbol{\Delta} = 0,1$ mas informando que vale 1/19, 1/10, 1/3 e 1.

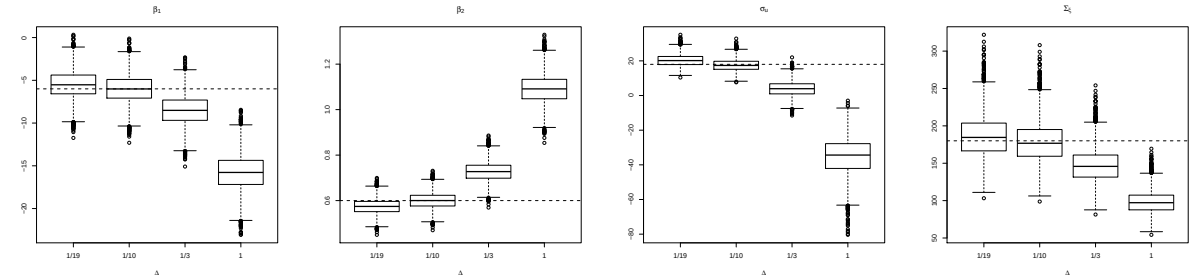

Figura 5.12: Distribuição dos estimadores de máxima verossimilhança de β_1, $\boldsymbol{\beta}_2$, σ_u e $\boldsymbol{\Sigma}_\xi$ quando amostras são geradas assumindo $\boldsymbol{\Delta} = 0,1$ mas informando que vale 1/19, 1/10, 1/3 e 1.

Pode-se notar, da Tabela 5.18, que o estimador de ν é praticamente não viciado e com erro quadrático médio pequeno apenas se o verdadeiro valor de ν é pequeno ($\nu = 2,01$). Se o valor real de ν é 5 e 30 o vício é positivo e aumenta com o verdadeiro valor de ν, assim como ocorre com o erro quadrático médio. Em geral, apenas os estimadores de σ_u e $\boldsymbol{\Sigma}_\xi$ tendem a ser viciados e os vícios aumentam, em módulo, à medida que são consideradas amostras geradas com maiores graus de liberdade. Similarmente ao que foi observado para os Cenários 1 e 2, a cobertura dos intervalos de 95% para todos os parâmetros está próxima do nível de confiança real.

A Figura 5.13 apresenta as distribuições dos estimadores de máxima verossimilhança

Tabela 5.18: Vício, raiz do erro quadrático médio e cobertura dos intervalos com 95% de confiança para β_1, $\boldsymbol{\beta}_2$, $\boldsymbol{\beta}_3$, σ_u, $\boldsymbol{\mu}_\xi$, $\boldsymbol{\Sigma}_\xi$ e ν quando amostras são geradas assumindo ν iguais a $2,01$, 5, 30 e ∞.

Parâmetro	ν	Vício	REQM	Cobertura
β_1	2,01	0,002	1,789	0,9502
	5	-0,002	1,669	0,9476
	30	-0,062	1,524	0,9498
	∞	-0,045	1,476	0,9458
$\boldsymbol{\beta}_2$	2,01	0,001	0,037	0,9514
	5	0,001	0,035	0,9486
	30	0,002	0,031	0,9516
	∞	0,002	0,030	0,9496
$\boldsymbol{\beta}_3$	2,01	-0,001	0,152	0,9510
	5	-0,002	0,143	0,9488
	30	0,003	0,129	0,9492
	∞	-0,001	0,126	0,9518
σ_u	2,01	-0,139	3,757	0,9526
	5	-0,411	3,440	0,9500
	30	-0,610	2,862	0,9484
	∞	-1,090	2,787	0,9324
$\boldsymbol{\mu}_\xi$	2,01	-0,016	1,249	0,9492
	5	0,020	1,149	0,9516
	30	-0,011	1,033	0,9480
	∞	-0,003	1,003	0,9514
$\boldsymbol{\Sigma}_\xi$	2,01	3,300	29,136	0,9466
	5	-1,707	27,445	0,9514
	30	-4,278	21,310	0,9464
	∞	-9,264	20,597	0,9264
ν	2,01	0,070	0,197	0,9276
	5	0,351	2,709	0,9702
	30	3,715	18,607	0,9606

dos parâmetros do modelo proposto quando amostras são geradas sob distintos graus de liberdade ν.

Pode-se notar da Figura 5.13 que, usualmente, as distribuições dos estimadores de máxima verossimilhança tendem a possui menor variância quando amostras são geradas sob maiores valores para ν, ou seja, se os dados vêm de uma distribuição com menor variância há uma tendência de que os estimadores também possuam menores variâncias. O fato inverso ocorre para o estimador do grau de liberdade ν. Para ν observa-se também que há uma tendência de seu estimador apresentar estimativas superestimadas, ou seja, há uma perda de robustez do modelo, uma vez que mesmo em amostras com caudas pesadas pode-se ter uma estimação de ν em um valor alto, que corresponderia a uma distribuição com cauda mais leve que a real. Se os dados são gerados de uma distribuição normal ($\nu = \infty$), as estimativas de máxima verossimilhança para os graus de liberdade tendem a ser altas (obviamente, nesse cenário, ν é subestimado). Observa-se também que

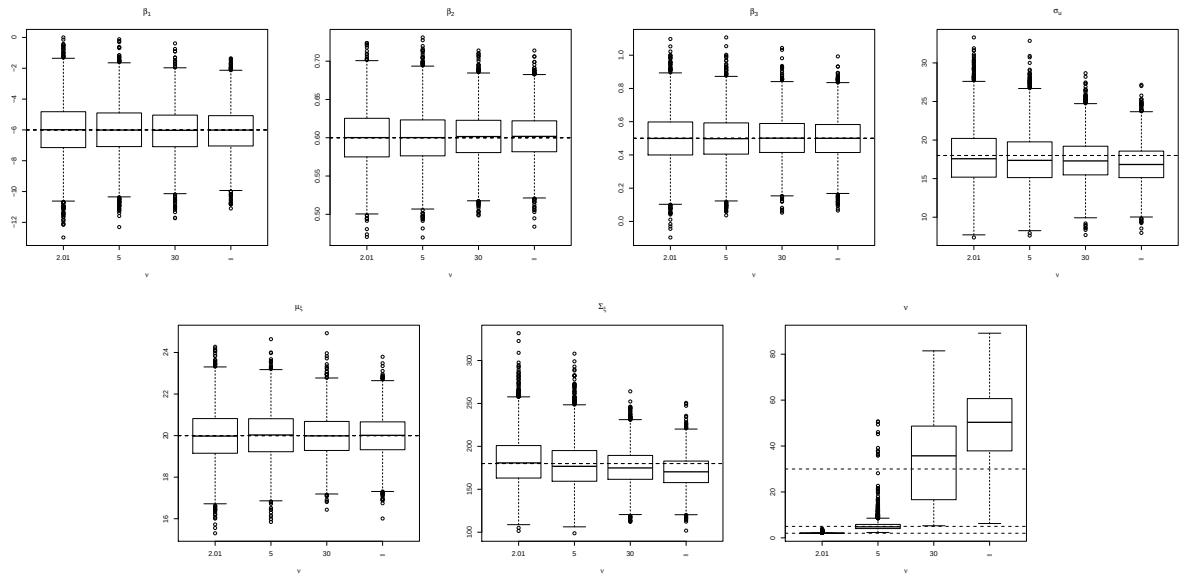

Figura 5.13: Distribuição dos estimadores de máxima verossimilhança de β_1, $\boldsymbol{\beta}_2$, $\boldsymbol{\beta}_3$, σ_u, $\boldsymbol{\mu}_\xi$, $\boldsymbol{\Sigma}_\xi$ e ν quando amostras são geradas assumindo ν iguais a $2,01$, 5, 30 e ∞.

para os estimadores de σ_u e $\boldsymbol{\Sigma}_\xi$ mais de 50% das estimativas obtidas estão subestimadas.

5.2.5 O efeito da má especificação de ν nos EMV - Cenário 5

No Cenário 5 o modelo proposto é ajustado assumindo-se que ν é um parâmetro fixo e conhecido. As Figuras 5.14, 5.15 e 5.16 mostram os vícios e as raízes dos erros quadráticos médios de todos os estimadores sempre que os dados são gerados assumindo-se $\nu = 2,01$, 5 e ∞, respectivamente.

Das Figuras 5.14, 5.15 e 5.16 nota-se que, em geral, os estimadores de todos os parâmetros tendem a apresentar os menores vícios e erros quadráticos médios se o modelo é ajustado fixando-se o grau de liberdade em valores mais próximos do valor real. Observa-se também que os estimadores de β_1, $\boldsymbol{\mu}_\xi$ e $\boldsymbol{\Sigma}_\xi$ são mais sensíveis às especificações de ν.

As Figuras 5.17, 5.18 e 5.19 apresentam as probabilidades de cobertura dos intervalos de 95% de confiança dos parâmetros do modelo quando amostras são geradas assumindo $\nu = 2,01$, 5 e ∞, respectivamente.

Das Figuras 5.17, 5.18 e 5.19 observa-se que, se ν é mal especificado no modelo, as probabilidades de cobertura dos intervalos de confiança para σ_u e $\boldsymbol{\Sigma}_\xi$ são bem menores que o nível de confiança nominal e, em alguns casos, aproximam-se de zero.

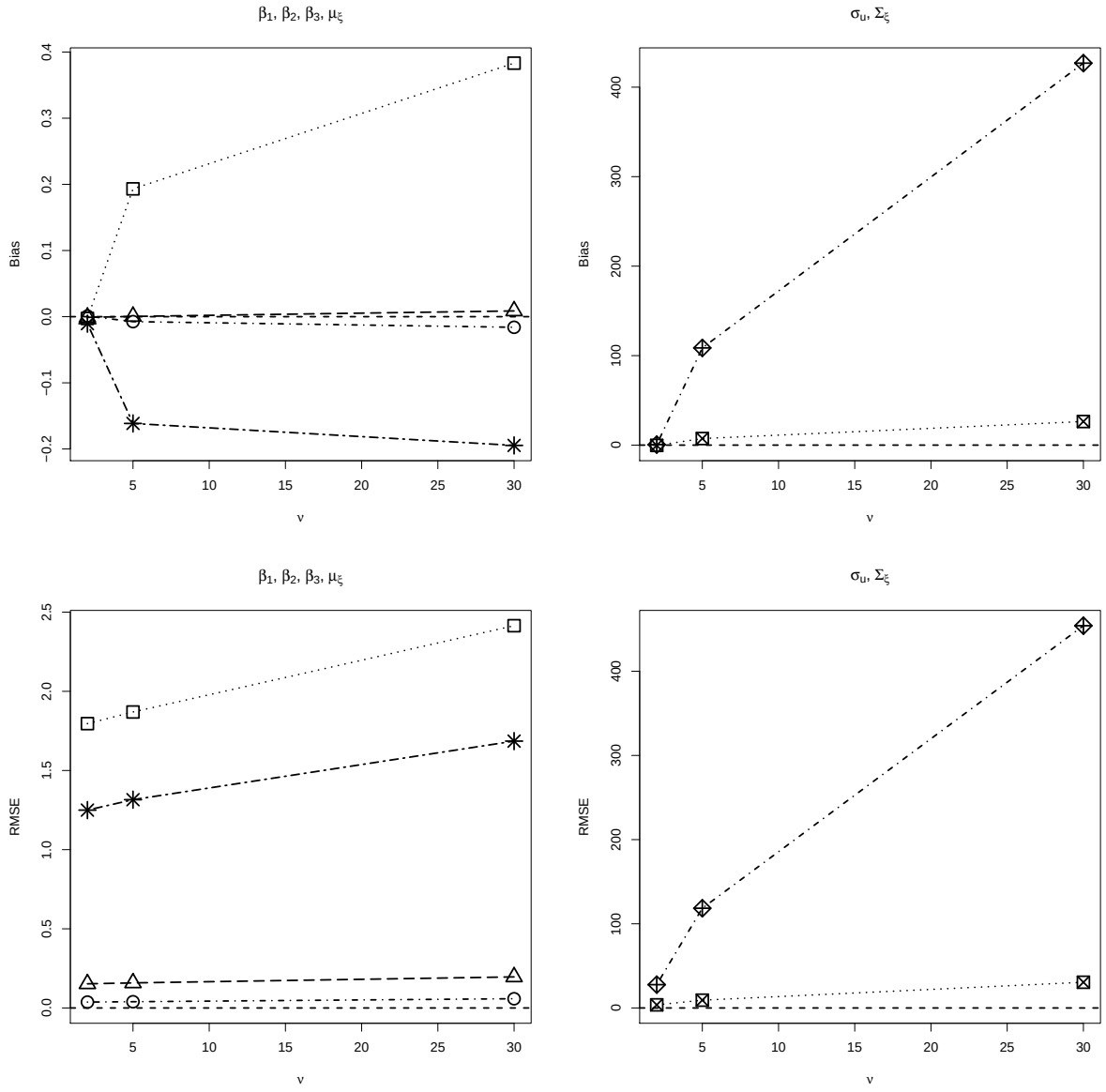

Figura 5.14: Vício (superior) e raiz do erro quadrático médio (inferior) para os estimadores de β_1 (quadrado), $\boldsymbol{\beta}_2$ (círculo), $\boldsymbol{\beta}_3$ (triângulo), $\boldsymbol{\mu}_\xi$ (asterisco) (gráficos à esquerda), σ_u (quadrado) e $\boldsymbol{\Sigma}_\xi$ (losango) (gráficos à direita) quando amostras são geradas assumindo $\nu = 2,01$ e no modelo assume-se ν iguais a $2,01$, 5 e 30.

As Figuras 5.20, 5.21 e 5.22 apresentam as distribuições dos EMV dos parâmetros do modelo proposto quando dados são gerados considerando-se $\nu = 2,01$, 5 e ∞, respectivamente.

Das Figuras 5.20, 5.21 e 5.22 nota-se que as estimativas de σ_u e $\boldsymbol{\Sigma}_u$ são as mais influenciadas pela especificação de ν. Quando um valor de ν maior que o real é especificado seus estimadores fornecem, em geral, superestimativas. Por outro lado, quando valores

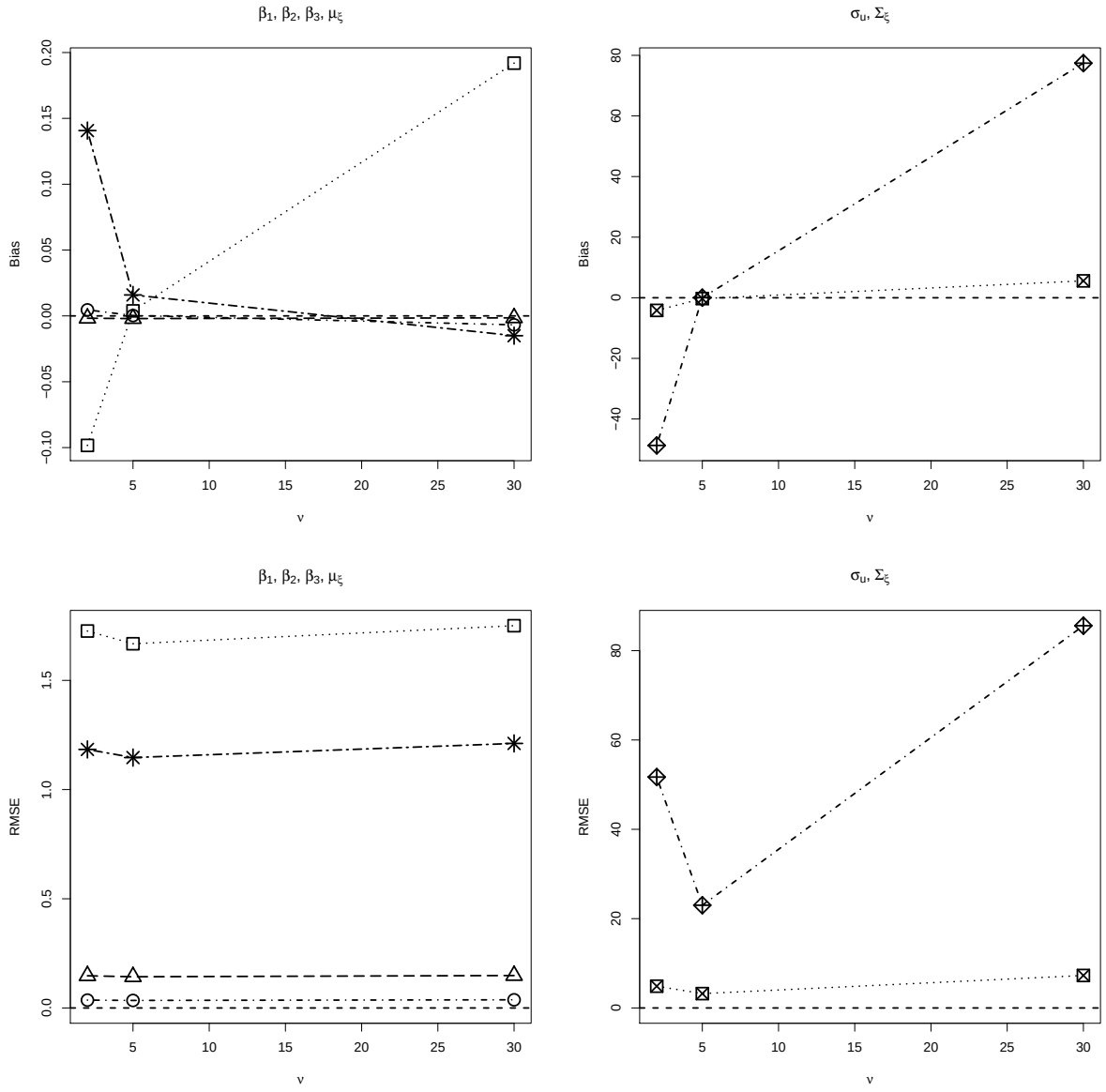

Figura 5.15: Vício (superior) e raiz do erro quadrático médio (inferior) para os estimadores de β_1 (quadrado), $\boldsymbol{\beta}_2$ (círculo), $\boldsymbol{\beta}_3$ (triângulo), $\boldsymbol{\mu}_\xi$ (asterisco) (gráficos à esquerda), σ_u (quadrado) e $\boldsymbol{\Sigma}_\xi$ (losango) (gráficos à direita) quando amostras são geradas assumindo $\nu = 5$ e no modelo assume-se ν iguais a $2,01$, 5 e 30.

de ν menores que o real são especificados, ν tende a ser subestimado.

Comparando os Cenários 4 e 5, conclui-se que há um ganho em estimar-se o grau de liberdade pois, apesar de haver uma perda de robustez, o vício dos estimadores dos outros parâmetros tendem a ser menores do que o observado quando o grau de liberdade ν é mal especificado.

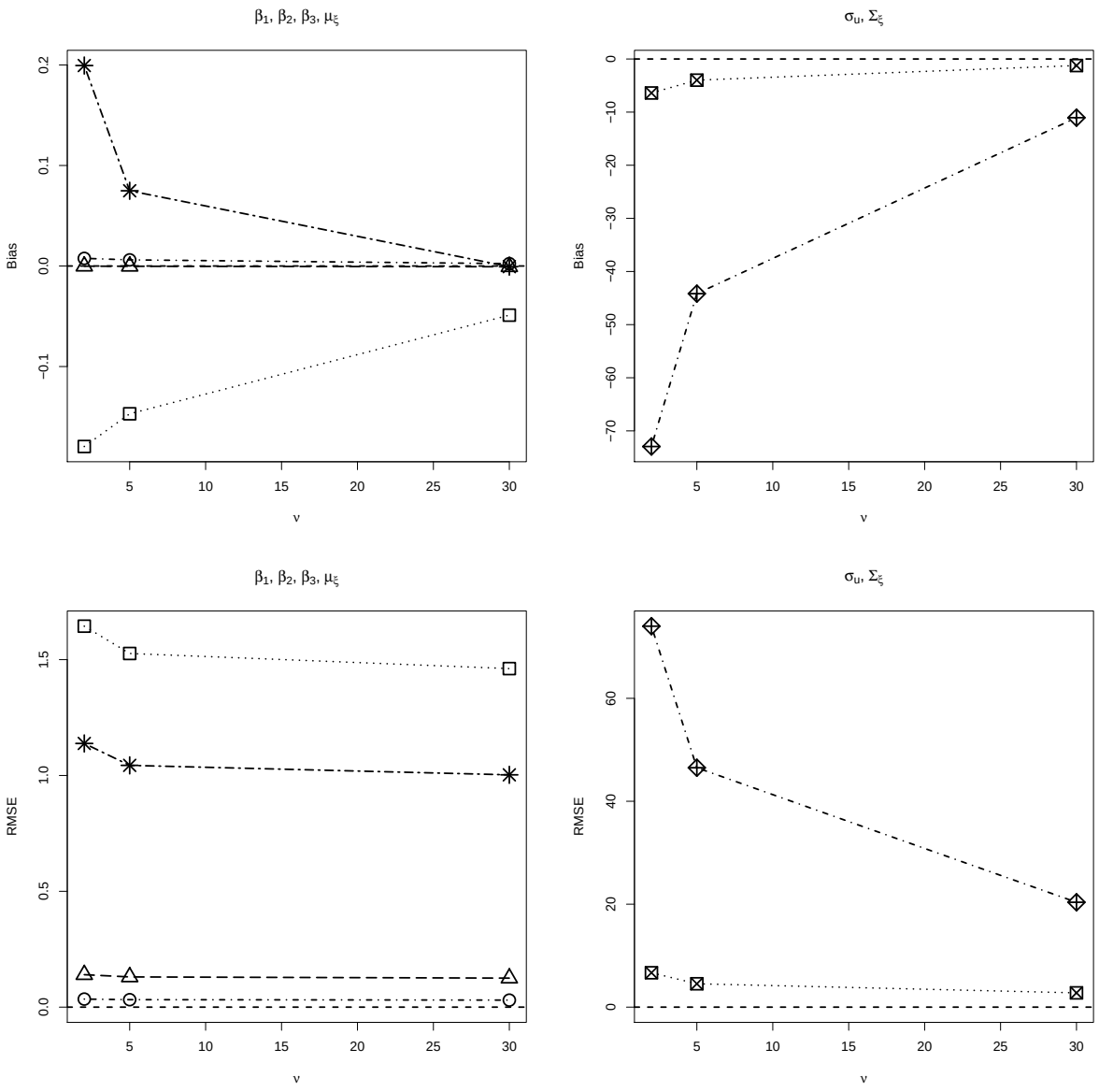

Figura 5.16: Vício (superior) e raiz do erro quadrático médio (inferior) para os estimadores de β_1 (quadrado), $\boldsymbol{\beta}_2$ (círculo), $\boldsymbol{\beta}_3$ (triângulo), $\boldsymbol{\mu}_\xi$ (asterisco) (gráficos à esquerda), σ_u (quadrado) e $\boldsymbol{\Sigma}_\xi$ (losango) (gráficos à direita) quando amostras são geradas assumindo $\nu = \infty$ e no modelo assume-se ν iguais a $2, 01$, 5 e 30.

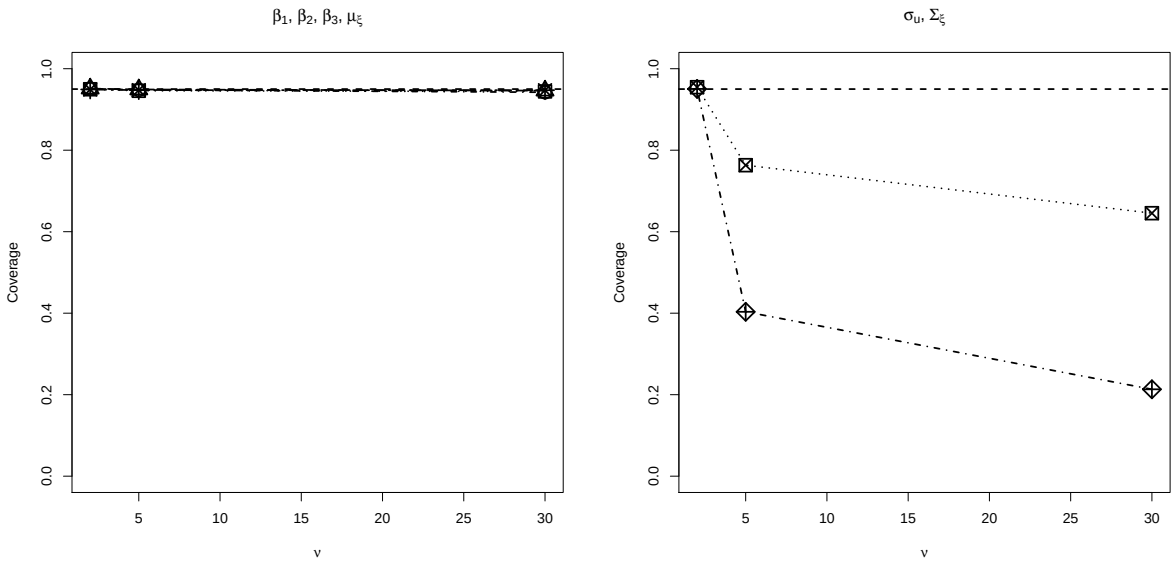

Figura 5.17: Cobertura dos intervalos com 95% de confiança para β_1 (quadrado), $\boldsymbol{\beta}_2$ (círculo), $\boldsymbol{\beta}_3$ (triângulo), $\boldsymbol{\mu}_\xi$ (asterisco) (gráfico à esquerda), σ_u (quadrado) e $\boldsymbol{\Sigma}_\xi$ (losango) (gráfico à direita) quando amostras são geradas assumindo $\nu = 2,01$ e no modelo assume-se ν iguais a $2,01$, 5 e 30.

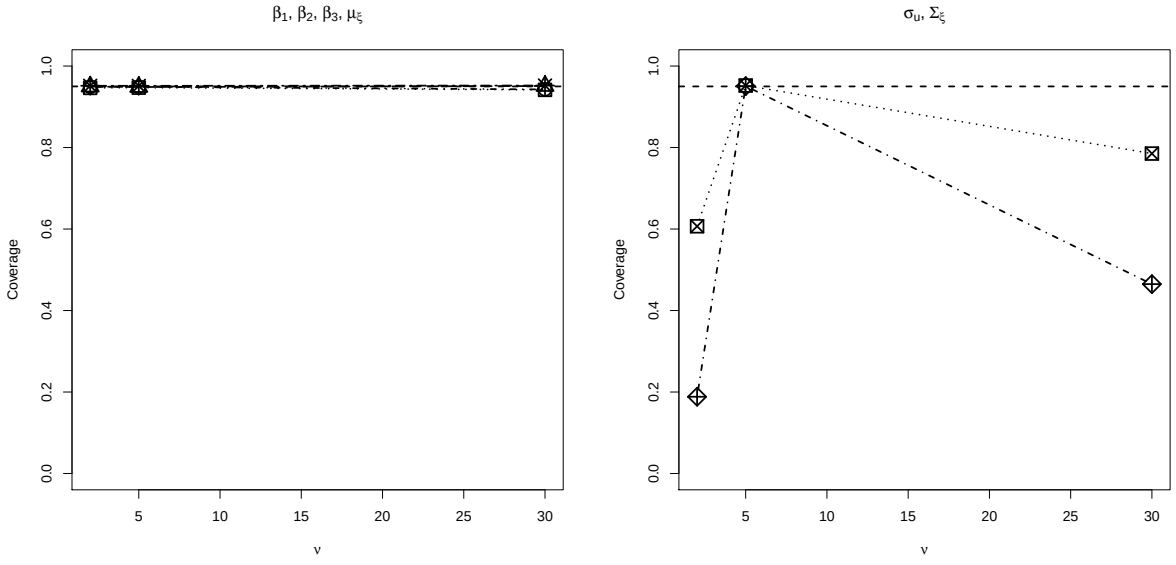

Figura 5.18: Cobertura dos intervalos com 95% de confiança para β_1 (quadrado), $\boldsymbol{\beta}_2$ (círculo), $\boldsymbol{\beta}_3$ (triângulo), $\boldsymbol{\mu}_\xi$ (asterisco) (gráfico à esquerda), σ_u (quadrado) e $\boldsymbol{\Sigma}_\xi$ (losango) (gráfico à direita) quando amostras são geradas assumindo $\nu = 5$ e no modelo assume-se ν iguais a $2,01$, 5 e 30.

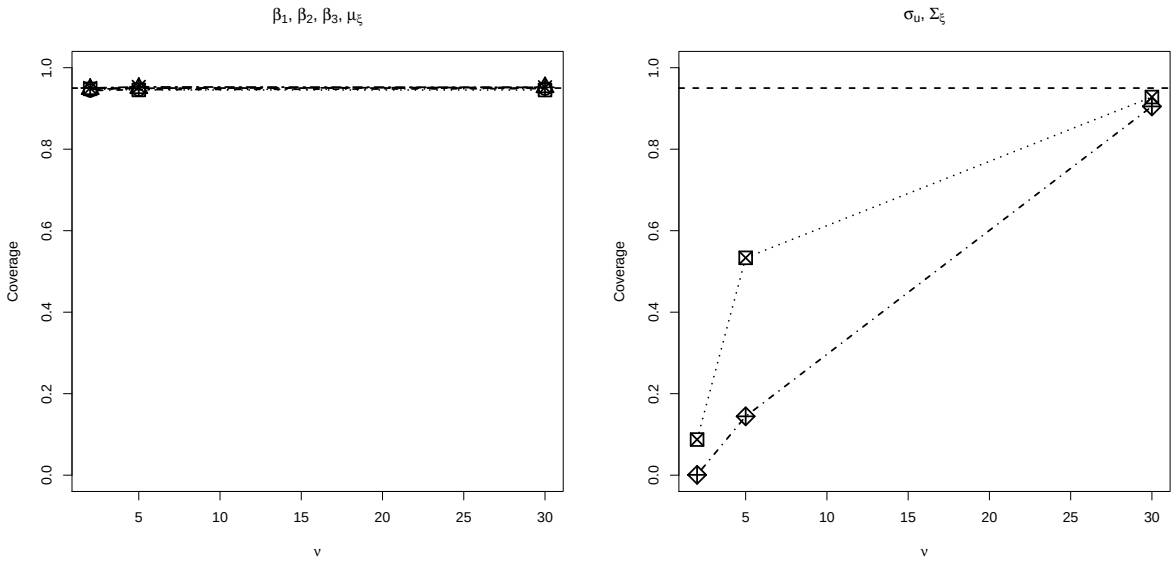

Figura 5.19: Cobertura dos intervalos com 95% de confiança para β_1 (quadrado), $\boldsymbol{\beta}_2$ (círculo), $\boldsymbol{\beta}_3$ (triângulo), $\boldsymbol{\mu}_\xi$ (asterisco) (gráfico à esquerda), σ_u (quadrado) e $\boldsymbol{\Sigma}_\xi$ (losango) (gráfico à direita) quando amostras são geradas assumindo $\nu = \infty$ e no modelo assume-se ν iguais a $2,01$, 5 e 30.

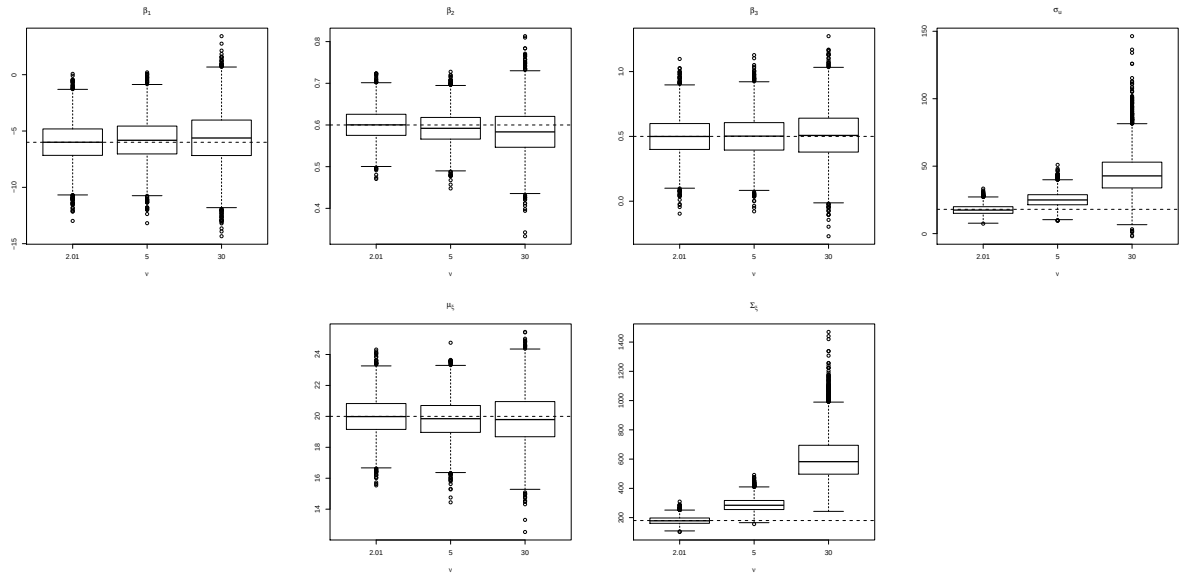

Figura 5.20: Distribuição dos estimadores de máxima verossimilhança de β_1, $\boldsymbol{\beta}_2$, $\boldsymbol{\beta}_3$, σ_u, $\boldsymbol{\mu}_\xi$ e $\boldsymbol{\Sigma}_\xi$ quando amostras são geradas assumindo $\nu = 2,01$ e no modelo assume-se ν iguais a $2,01$, 5 e 30.

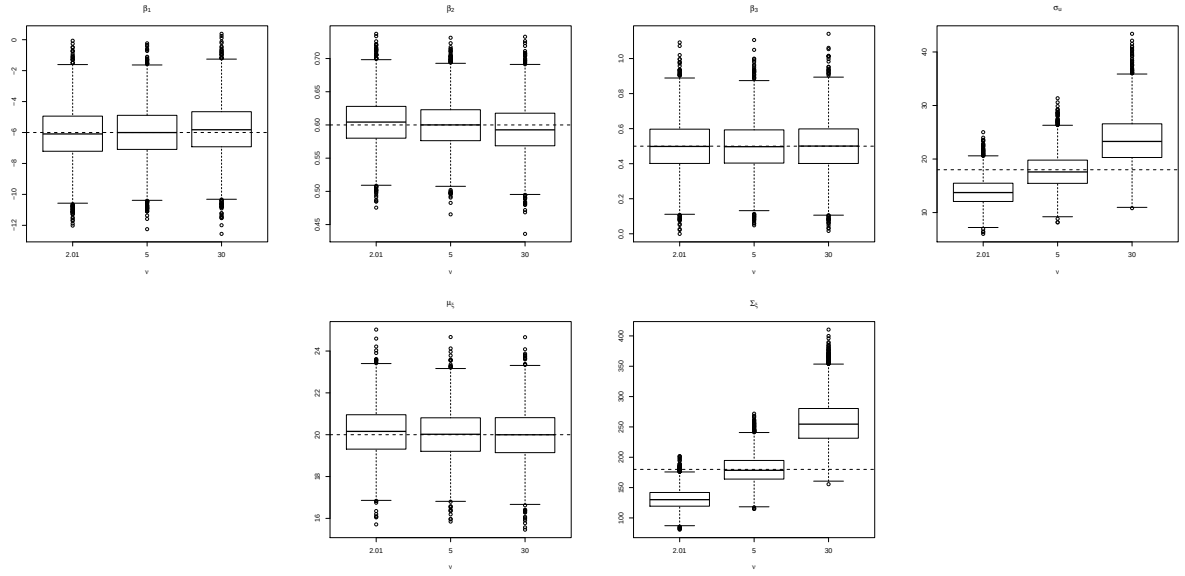

Figura 5.21: Distribuição dos estimadores de máxima verossimilhança de β_1, $\boldsymbol{\beta}_2$, $\boldsymbol{\beta}_3$, σ_u, $\boldsymbol{\mu}_\xi$ e $\boldsymbol{\Sigma}_\xi$ quando amostras são geradas assumindo $\nu = 5$ e no modelo assume-se ν iguais a $2,01$, 5 e 30.

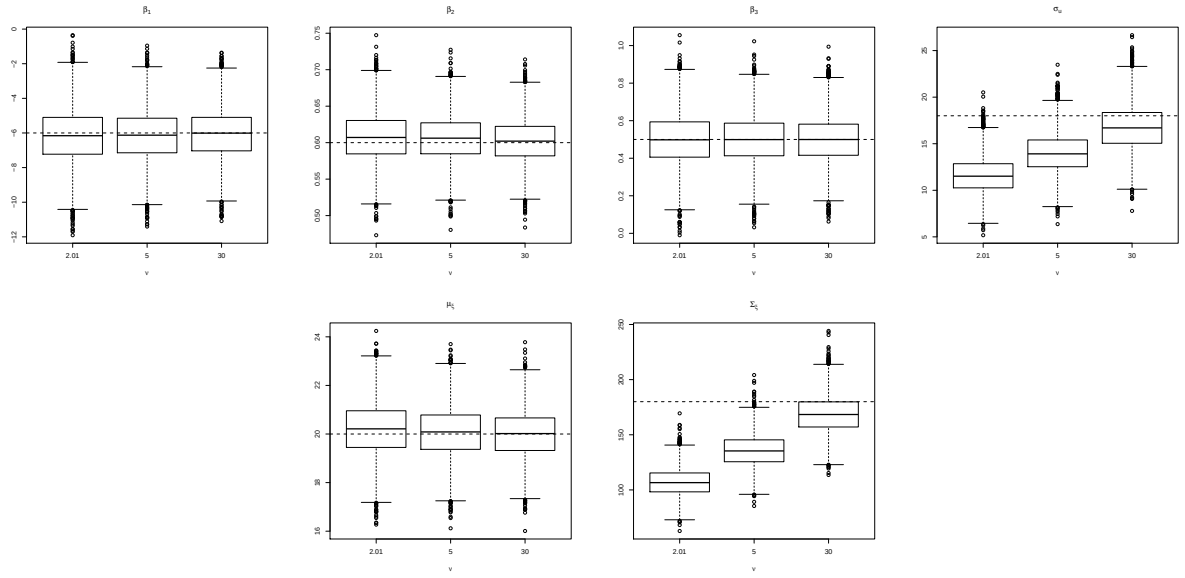

Figura 5.22: Distribuição dos estimadores de máxima verossimilhança de β_1, $\boldsymbol{\beta}_2$, $\boldsymbol{\beta}_3$, σ_u, $\boldsymbol{\mu}_\xi$ e $\boldsymbol{\Sigma}_\xi$ quando amostras são geradas assumindo $\nu = \infty$ e no modelo assume-se ν iguais a $2,01$, 5 e 30.

Capítulo 6

Estudo de caso: gastos ambulatoriais

Neste capítulo os dados reportados em Cameron e Trivedi (2010) disponíveis no conjunto de dados '*2001 Medical Expenditure Panel Survey*' são re-analisados utilizando o modelo proposto e, no caso bayesiano, também utilizando os modelos Tobit, t-Tobit e N-Tobit com erros nas covariáveis. A variável resposta corresponde a 3.328 observações de valores de gastos ambulatoriais (na escala logarítmica) em que 526 são iguais a zero, isto é, tem uma censura aproximada de $15,8\%$. As variáveis explicativas são idade (`age`, medida em dezenas de anos), gênero (`female`, indicadora de feminino), status educacional (`educ`, número de anos), status de seguro (`ins`), total de doenças crônicas (`totchr`), etnia (`blhisp`, indicadora de negro ou hispânico) e renda (`income`). Cameron e Trivedi (2010) analisaram o conjunto de dados usando o modelo Tobit e consideraram o método de máxima verossimilhança para estimar os parâmetros. A hipótese de normalidade é rejeitada por tais autores, que não incluíram a covariável renda dentre as variáveis explicativas. Trabalhos anteriores, inclusive Fuller (1987), argumentam que a variável renda é usualmente medida de forma imprecisa. Assim, diferente de Cameron e Trivedi (2010), propõe-se analisar estes mesmos dados considerando um modelo robusto que inclui a covariável renda medida com erro.

Para definir o valor de $\boldsymbol{\Delta}$ considera-se a sugestão dada em Fuller (1987) que estabelece que a variância do erro de medida $\boldsymbol{\Sigma}_v$ é aproximadamente 15% da variância total da renda, isto é, $\Lambda = (\boldsymbol{\Sigma}_\xi + \boldsymbol{\Sigma}_v)^{-1}\boldsymbol{\Sigma}_\xi = 0,85$. Consequentemente, segue que $\boldsymbol{\Delta} = 3/17$.

6.1 Estimadores bayesianos dos modelos tipo Tobit nos gastos ambulatoriais

Nesta seção inferência será feita sob o paradigma bayesiano. O modelo proposto será ajustado e comparado com os modelos N-Tobit com erros nas covariáveis, N-Tobit e t-Tobit. Como mencionado anteriormente, assume-se que a variável renda pode ser medida com erro. Como não há informações *a priori* sobre os parâmetros, distribuições *a priori* pouco informativas serão eliciadas para todos os parâmetros, isto é, assume-se $\gamma_1 \stackrel{d}{=} \boldsymbol{\gamma}_2 \stackrel{d}{=} \boldsymbol{\mu}_x \sim N_1(0, 10^6)$, $\boldsymbol{\beta}_3 \sim N_6(\mathbf{0}_6, 10^3[\mathbf{1}_6\mathbf{1}_6' + (10^3 - 1)\mathbf{I}_6])$, $\sigma_w \sim IG(2,0001, 10,001)$ e $\boldsymbol{\Sigma}_x \sim IW(6, (45 \times 10^5)^{1/2})$. Como consequência, as distribuições *a priori* para os parâmetros de regressão β_1 e $\boldsymbol{\beta}_2$ no modelo original são tais que $\beta_1 \mid \boldsymbol{\beta}_2 \sim N_1(0, 10^6[1+\boldsymbol{\Delta}(1+\boldsymbol{\Delta})^{-1}\boldsymbol{\beta}_2]^2)$ e $\boldsymbol{\beta}_2 \sim N_1(0, 10^6(1 + \boldsymbol{\Delta})^2)$. Para os modelos livre de erros assume-se $\beta_1 \sim N_1(0, 10^6)$, $\boldsymbol{\beta}_2 \sim N_7(\mathbf{0}_7, 10^3[\mathbf{1}_7\mathbf{1}_7'+(10^3-1)\mathbf{I}_7])$ e $\sigma_u \sim IG(2,0001, 10,001)$. Para os modelos t-Tobit e t-Tobit com erros nas covariáveis assume-se $\nu \sim TE(10^3; 2)$. Os parâmetros considerados para o MCMC são os mesmos assumidos no estudo Monte Carlo.

A Tabela 6.1 apresenta algumas medidas resumo *a posteriori* obtidas ajustando os modelos sob comparação.

Tabela 6.1: Estimativas *a posteriori* sob todos os modelos, conjunto de dados de gastos ambulatoriais.

		tce			nce			tse			nse		
Covariável	Parâmetro	Média	Mediana	D.P.	Média	Mediana	D.P.	Média	Mediana	D.P.	Média	Mediana	D.P.
	β_1	3,526	3,517	0,239	1,027	0,974	0,272	1,968	1,988	0,331	0,981	0,919	0,393
income	$\boldsymbol{\beta}_2$	0,004	0,003	0,003	0,004	0,003	0,002	0,002	0,002	0,002	0,003	0,003	0,002
age	$\boldsymbol{\beta}_{31}$	0,274	0,276	0,030	0,345	0,352	0,046	0,343	0,340	0,041	0,350	0,355	0,050
female	$\boldsymbol{\beta}_{32}$	0,746	0,748	0,065	1,361	1,360	0,096	1,127	1,118	0,108	1,350	1,348	0,098
educ	$\boldsymbol{\beta}_{33}$	0,061	0,062	0,015	0,126	0,126	0,017	0,112	0,112	0,018	0,130	0,133	0,023
blhisp	$\boldsymbol{\beta}_{34}$	-0,459	-0,457	0,087	-0,885	-0,896	0,108	-0,748	-0,736	0,102	-0,865	-0,858	0,112
totchr	$\boldsymbol{\beta}_{35}$	0,723	0,719	0,042	1,163	1,154	0,061	0,924	0,920	0,068	1,160	1,157	0,066
ins	$\boldsymbol{\beta}_{36}$	0,094	0,105	0,071	0,240	0,245	0,107	0,187	0,181	0,092	0,252	0,248	0,108
	σ_u	2,774	2,759	0,121	7,754	7,740	0,220	6,135	6,149	0,331	7,755	7,771	0,215
	$\boldsymbol{\mu}_\xi$	31,962	31,990	0,386	36,773	36,745	0,448						
	$\boldsymbol{\Sigma}_\xi$	216,367	216,891	7,798	605,718	605,242	14,521						
	ν	2,061	2,049	0,049				5,466	5,223	1,141			

Da Tabela 6.1 nota-se que as médias e medianas *a posteriori* para todos os parâmetros tendem a ser próximas entre si e que os efeitos de todas as covariáveis são positivos, exceto para etnia, para todos os modelos. Também nota-se que as estimativas *a posteriori* obtidas para os efeitos fixos ajustando os modelos N-Tobit com erros nas covariáveis, t-Tobit e Tobit são similares e, em geral, são maiores (em valores absolutos) do que as obtidas pelo modelo proposto. A exceção ocorre para β_1 em que as estimativas *a*

posteriori tendem a ser altas sob o modelo proposto e o t-Tobit. Observa-se também que as estimativas para todas as componentes da variância e para o grau de liberdade tendem a ser menores sob o modelo proposto.

A Tabela 6.2 mostra os intervalos de mais alta densidade *a posteriori* (HPD) para os parâmetros dos modelos ajustados.

Tabela 6.2: Os intervalos HPD de 95%, conjunto de dados de gastos ambulatoriais.

Covariável	Parâmetro	tce	nce	tse	nse
	β_1	[3,114 ; 3,934]	[0,595 ; 1,620]	[1,387 ; 2,614]	[0,321 ; 1,788]
income	$\boldsymbol{\beta}_2$	[0,000 ; 0,009]	[0,000 ; 0,008]	[-0,002 ; 0,005]	[-0,002 ; 0,006]
age	$\boldsymbol{\beta}_{31}$	[0,209 ; 0,322]	[0,261 ; 0,428]	[0,273 ; 0,423]	[0,238 ; 0,441]
female	$\boldsymbol{\beta}_{32}$	[0,630 ; 0,865]	[1,190 ; 1,558]	[0,955 ; 1,409]	[1,165 ; 1,524]
educ	$\boldsymbol{\beta}_{33}$	[0,034 ; 0,090]	[0,093 ; 0,158]	[0,076 ; 0,141]	[0,082 ; 0,166]
blhisp	$\boldsymbol{\beta}_{34}$	[-0,646 ; -0,285]	[-1,074 ; -0,662]	[-0,943 ; -0,542]	[-1,080 ; -0,671]
totchr	$\boldsymbol{\beta}_{35}$	[0,639 ; 0,798]	[1,062 ; 1,290]	[0,813 ; 1,053]	[1,044 ; 1,294]
ins	$\boldsymbol{\beta}_{36}$	[-0,035 ; 0,241]	[0,005 ; 0,410]	[0,007 ; 0,338]	[0,084 ; 0,495]
	σ_u	[2,573 ; 3,019]	[7,384 ; 8,180]	[5,585 ; 6,830]	[7,341 ; 8,138]
	$\boldsymbol{\mu}_\xi$	[31,289 ; 32,655]	[35,974 ; 37,707]		
	$\boldsymbol{\Sigma}_\xi$	[202,284 ; 230,961]	[571,823 ; 629,737]		
	ν	[2,000 ; 2,166]		[3,675 ; 8,050]	

A partir da Tabela 6.2 observa-se que ambos os modelos que consideram erros na covariável renda apontam que o efeito da renda é positivo com probabilidades $0,95$ ou mais, isto é, sob ambos os modelos a renda pode ser considerada significativa para explicar os gastos ambulatoriais (ver também a Figura 6.1). Mais ainda, apenas o modelo proposto mostra que o status do seguro não é relevante para explicar os gastos ambulatoriais.

Usualmente, perde-se robustez quando se estima o grau de liberdade do modelo. Apesar disso, conclui-se, a partir dos resultados exibidos na Tabela 6.2 e na Tabela 6.3, que a distribuição conjunta para os erros e a covariável latente possui cauda mais pesada que a distribuição normal. As estimativas *a posteriori* para ν sob ambos, modelo proposto e t-Tobit, revelam que o grau de liberdade é pequeno (média *a posteriori* igual a $2,061$ sob o modelo proposto e $5,466$ sob o modelo t-Tobit).

Mais ainda, similarmente ao que foi observado nos estudos de simulação, as estimativas *a posteriori* para σ_u são muito maiores ao comparar com o modelo proposto, enquanto a média é $2,774$ sob o modelo proposto, é $7,754$ sob o modelo N-Tobit com erros nas covariáveis. As estimativas *a posteriori* para $\boldsymbol{\Sigma}_\xi$ são também menores sob o modelo proposto. Considerando as médias *a posteriori* e as estimativas obtidas sob o modelo proposto segue, como consequência, que a média e a variância *a posteriori* de

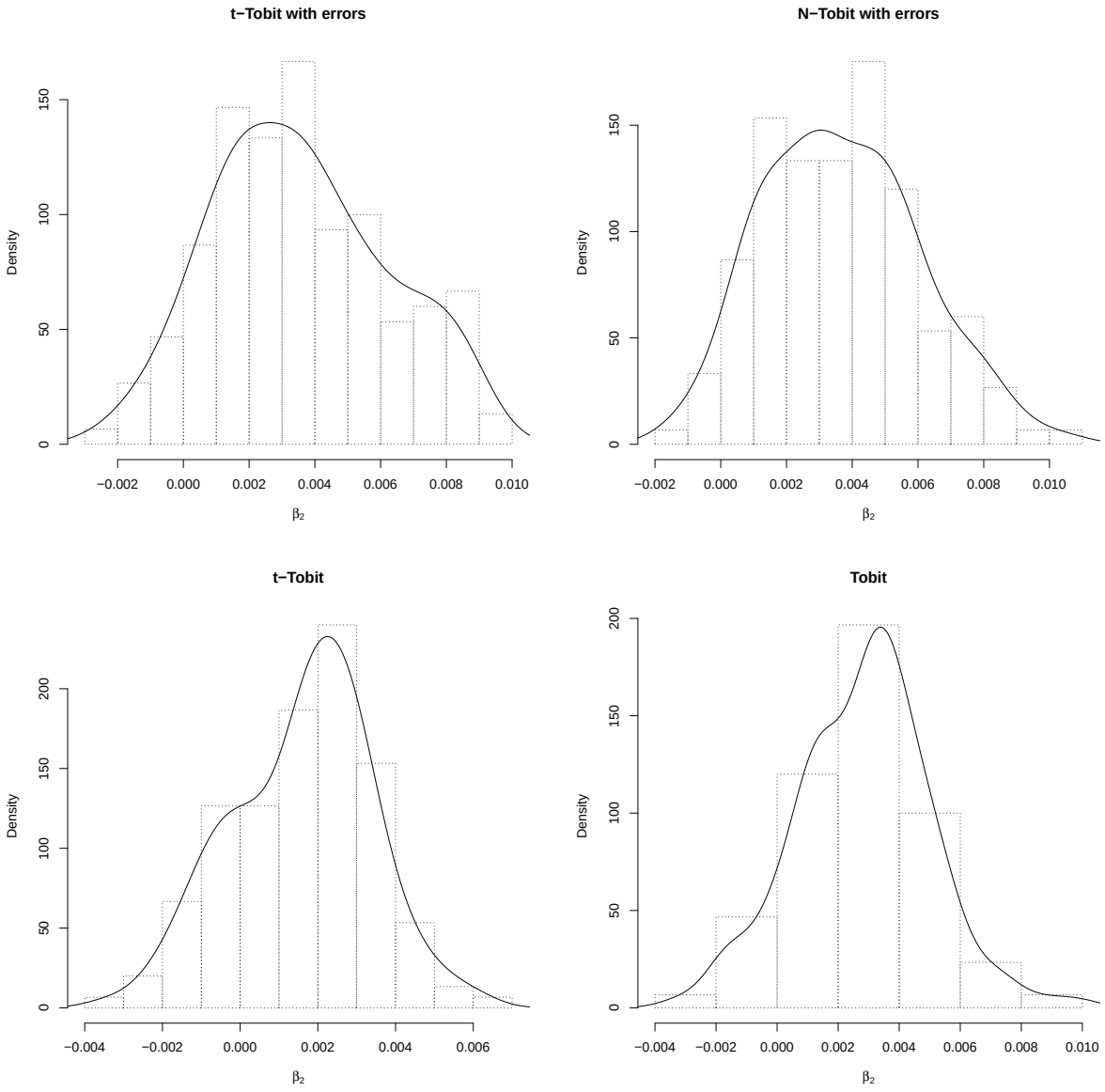

Figura 6.1: Distribuições *a posteriori* de $\boldsymbol{\beta}_2$.

$\boldsymbol{\Sigma}_v$ são, respectivamente, $38,18$ e $1,89$, dando evidência a favor do modelo com erro de classificação.

Em resumo, as distribuições *a posteriori* para ν e $\boldsymbol{\Sigma}_v$ revelam que um modelo com erro nas covariáveis robusto (modelo proposto) é mais apropriado para descrever o comportamento dos dados.

A Tabela 6.3 mostra as estatísticas LPML e DIC para cada modelo. Nota-se que, dentre cada classe de modelos ajustados (com erros e sem erros) ambas as estatísticas

trazem evidência de que o melhor modelo é o que considera a distribuição t-Student.

Tabela 6.3: LPML e DIC, conjunto de dados de gastos ambulatoriais.

Model	LPML	DIC
tce	-22528,00	45056,07
nce	-23158,83	46313,70
tse	-7367,83	14724,14
nse	-7502,82	15005,70

Para investigar se os modelos considerados são sensíveis às suas suposições considera-se, para cada modelo (Figura 6.2), a divergência de Kullback-Leibler (K-L) entre a distribuição *a posteriori* para $\boldsymbol{\theta}$ considerando todos os dados observados, $\pi = \pi(\boldsymbol{\theta} \mid \mathbf{D}_o)$, e a distribuição *a posteriori* para $\boldsymbol{\theta}$ sem a i-ésima observação, $\pi_{(-i)} = \pi(\boldsymbol{\theta} \mid \mathbf{D}_{o(i)})$, que é dada por $K(\pi, \pi_{(-i)}) = \int \pi(\boldsymbol{\theta} \mid \mathbf{D}_0) \log \left[\frac{\pi(\boldsymbol{\theta}|\mathbf{D}_0)}{\pi(\boldsymbol{\theta}|\mathbf{D}_{0(i)})} \right] d\boldsymbol{\theta}$. Mais detalhes sobre a divergência K-L e um procedimento computacional para aproximá-la pode ser encontrada em Lachos *et al.* (2011).

Da Figura 6.2 pode-se notar que há observações que são potencialmente influentes sob o modelo N-Tobit com erros nas covariáveis mas que não são sob os outros modelos. De fato, o modelo proposto parece ser menos afetado por observações atípicas do que os outros modelos uma vez que os valores de $K(\pi, \pi_{(-i)})$ são mais próximo de zero. Logo, tal modelo deve ser o preferido para ajustar aos dados.

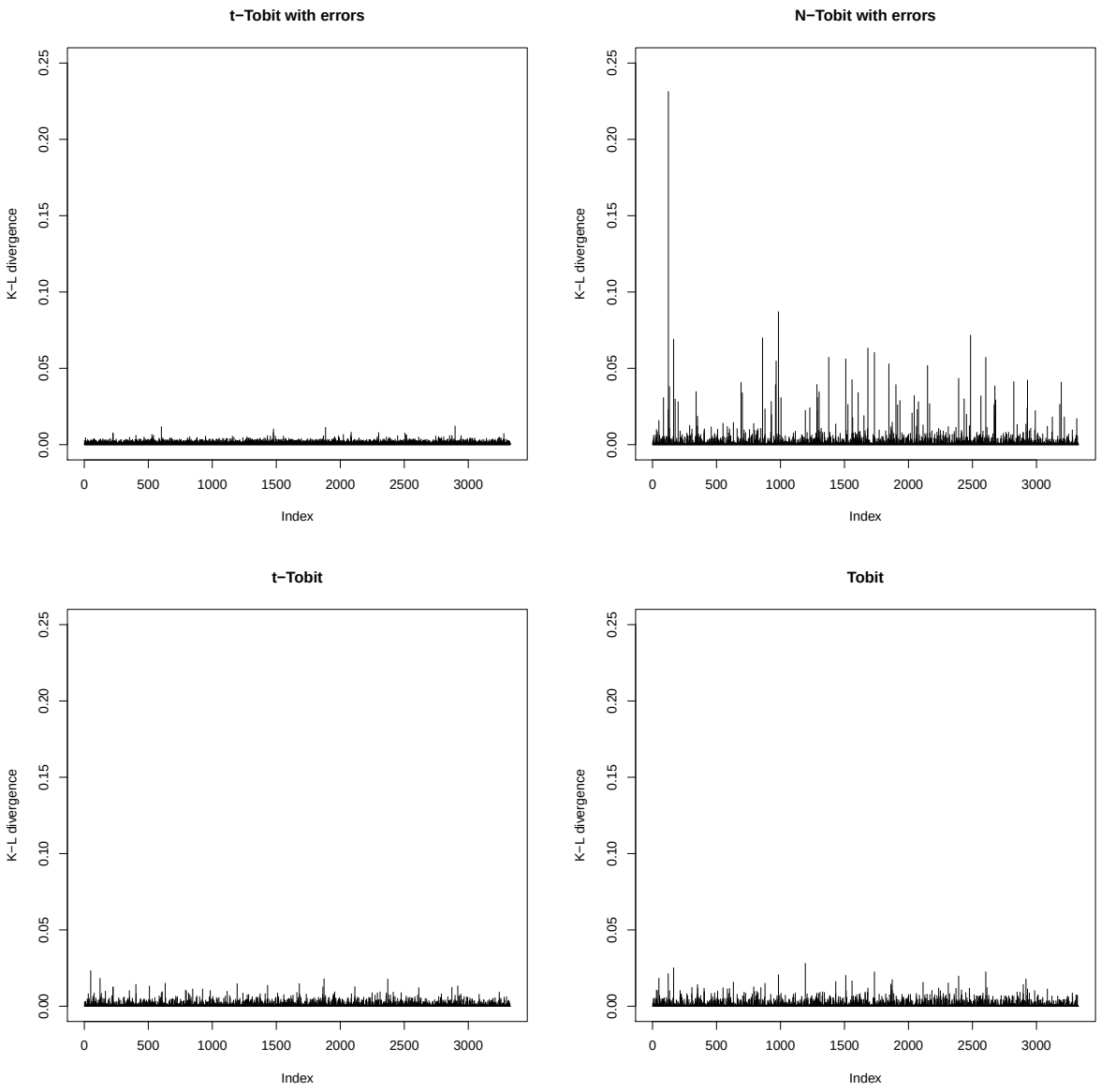

Figura 6.2: Divergência K-L para todos os modelos.

6.2 Estimadores de máxima verossimilhança do modelo proposto nos gastos ambulatoriais

Nesta seção os dados descritos no início do capítulo são analisados sob o paradigma da inferência clássica. O modelo proposto será ajustado aos dados considerando duas abordagens. Em uma delas o grau de liberdade ν é desconhecido (Modelo 1) e na outra considera-se $\nu = 171$ (Modelo 2). O valor 171 para ν foi o maior considerado para que não houvessem problemas computacionais envolvendo o cálculo da função gama no

R (R Core Team, 2014), software utilizado neste trabalho. O Modelo 2 fornece uma aproximação para o modelo N-Tobit com erros nas covariáveis. Para ambos os modelos serão considerados os estimadores de máxima verossimilhança aproximados pelo algoritmo ECM. O algoritmo ECM foi iniciado considerando $\widehat{\beta}_1^{(1)} = \widehat{\beta}_2^{(1)} = 0$, $\widehat{\boldsymbol{\beta}}_3^{(1)} = \mathbf{0}$, $\widehat{\sigma}_u^{(1)} = s_y^2 = 7,40$, $\widehat{\boldsymbol{\mu}}_\xi^{(1)} = \bar{\mathbf{x}} = 36,80$, $\widehat{\boldsymbol{\Sigma}}_\xi^{(1)} = \mathbf{S}_x = 712,95$ e, para o Modelo 1, $\widehat{\nu}^{(1)} = 5$. Foram necessárias 307 iterações até que $|\log f(\mathbf{D}_o \mid \widehat{\boldsymbol{\theta}}^{(t+1)})/\log f(\mathbf{D}_o \mid \widehat{\boldsymbol{\theta}}^{(t)}) - 1|$ fosse menor que 10^{-8} no Modelo 1 e 246 no Modelo 2.

A Tabela 6.4 mostra as estimativas de máxima verossimilhança bem como os erros padrão, os intervalos de 95% de confiança para todos os parâmetros e também os p-valores para testar as hipóteses nulas $H_0 : \beta_1 = 0$, $H_0 : \boldsymbol{\beta}_2 = 0$ e $H_0 : \boldsymbol{\beta}_{3j} = 0$, $j = 1, \ldots, 6$ sob os dois modelos. Para obter os intervalos de confiaça e os p-valores são consideradas as distribuições assintóticas dos EMV.

Tabela 6.4: Estimativas de máxima verossimilhança, desvio-padrão assintótico, intervalos de 95% de confiança e p-valores, conjunto de dados de gastos ambulatoriais.

Covariável	Parâmetro	EMV	Erro-padrão	95% Int. Conf.	p-valor
Modelo proposto com ν desconhecido (Modelo 1)					
	β_1	3,687	0,247	[3,204 ; 4,171]	0,000
income	$\boldsymbol{\beta}_2$	0,004	0,003	[-0,002 ; 0,009]	0,171
age	$\boldsymbol{\beta}_{31}$	0,263	0,032	[0,201 ; 0,326]	0,000
female	$\boldsymbol{\beta}_{32}$	0,722	0,070	[0,583 ; 0,860]	0,000
educ	$\boldsymbol{\beta}_{33}$	0,056	0,015	[0,027 ; 0,084]	0,000
blhisp	$\boldsymbol{\beta}_{34}$	-0,472	0,077	[-0,624 ; -0,321]	0,000
totchr	$\boldsymbol{\beta}_{35}$	0,714	0,042	[0,632 ; 0,796]	0,000
ins	$\boldsymbol{\beta}_{36}$	0,089	0,070	[-0,047 ; 0,226]	0,200
	σ_u	2,687	0,138	[2,415 ; 2,958]	
	$\boldsymbol{\mu}_\xi$	31,896	0,358	[31,194 ; 32,597]	
	$\boldsymbol{\Sigma}_\xi$	212,694	8,393	[196,244 ; 229,144]	
	ν	2,000	0,076	[1,850 ; 2,150]	
Modelo normal aproximado (Modelo 2)					
	β_1	1,042	0,330	[0,395 ; 1,690]	0,002
income	$\boldsymbol{\beta}_2$	0,004	0,002	[-0,001 ; 0,009]	0,107
age	$\boldsymbol{\beta}_{31}$	0,347	0,046	[0,258 ; 0,436]	0,000
female	$\boldsymbol{\beta}_{32}$	1,359	0,098	[1,166 ; 1,551]	0,000
educ	$\boldsymbol{\beta}_{33}$	0,125	0,021	[0,085 ; 0,165]	0,000
blhisp	$\boldsymbol{\beta}_{34}$	-0,866	0,109	[-1,079 ; -0,654]	0,000
totchr	$\boldsymbol{\beta}_{35}$	1,156	0,064	[1,030 ; 1,283]	0,000
ins	$\boldsymbol{\beta}_{36}$	0,249	0,102	[0,050 ; 0,448]	0,014
	σ_u	7,630	0,189	[7,259 ; 8,001]	
	$\boldsymbol{\mu}_\xi$	36,512	0,454	[35,622 ; 37,401]	
	$\boldsymbol{\Sigma}_\xi$	576,030	14,504	[547,603 ; 604,457]	

Da Tabela 6.4 pode-se concluir, sob ambos os modelos, que o efeito marginal de todas as covariáveis são positivas, exceto para etnia. Além da covariável gênero, o número total

de doenças crônicas e a etnia causam mais impacto nos gastos ambulatoriais (latente). Por exemplo, sob o Modelo 1, conclui-se que o gasto ambulatorial médio por mulher é $105,8\%$ mais alto que aquele por homem enquanto uma mudança de 10 anos na idade está associada com um aumento de $30,1\%$ no gasto ambulatorial médio.

A partir dos p-valores e dos itervalos de confiança também conclui-se, sob ambos os modelos, que a variável explicativa renda (medida com erro) não é significativa para explicar os gastos ambulatoriais se níveis de significância usuais são considerados. Conclusão similar pode ser tirada para o status de seguro sob o Modelo 1. Apesar disso, utilizando a propriedade da invariância dos estimadores de máxima verossimilhança e a estimativa para $\boldsymbol{\Sigma}_\xi$, uma vez que $\boldsymbol{\Delta} = 3/17$, tem-se que a estimativa de máxima verossimilhança para $\boldsymbol{\Sigma}_v$ é $\widehat{\boldsymbol{\Sigma}_v} = 37,53$ sob o Modelo 1 e $101,65$ sob o Modelo 2, ou seja, aqui também se conclui que não é razoável assumir que a variável renda seja livre de erros.

Comparando os modelos também nota-se que não é razoável assumir um modelo com caudas leves para esse conjunto de dados. Sob o Modelo 1, a estimativa de máxima verossimilhança para o grau de liberdade ν é $2,0$, que significa que a aproximação para o modelo normal não é apropriada nesse caso. Conclusões similares podem ser tiradas ao considerar as estatísticas AIC, BIC e a log-verossimilhança, dadas na Tabela 6.5. Além disso, outra questão relevante em modelos de regressão é avaliar a sensibilidade dos estimadores de máxima verossimilhança em relação a pequenas perturbações nas suposições do modelo na presença de observações atípicas. Como uma ferramenta de diagnóstico, a distância de Cook generalizada (GCD) (Figura 6.3) para o modelo Tobit (Barros *et al.*, 2010) será considerada. A distância de Cook é uma importante técnica de diagnóstico do método de influência global e permite estudar mudanças nas estimativas dos parâmetros se uma observação é descartada do conjunto de dados.

Tabela 6.5: AIC, BIC e log-verossimilhança, conjunto de dados de gastos ambulatoriais.

Modelo	AIC	BIC	Log-verossimilhança
Modelo 1	45056,03	45129,35	-22516,02
Modelo 2	46242,99	46310,20	-23110,49

Como pode ser notado da Figura 6.3, cada observação causa menos impacto nas estimativas de máxima verossimilhança sob o Modelo 1. Portanto, um modelo com cauda mais pesada fornece um melhor ajuste do que o modelo normal aproximado.

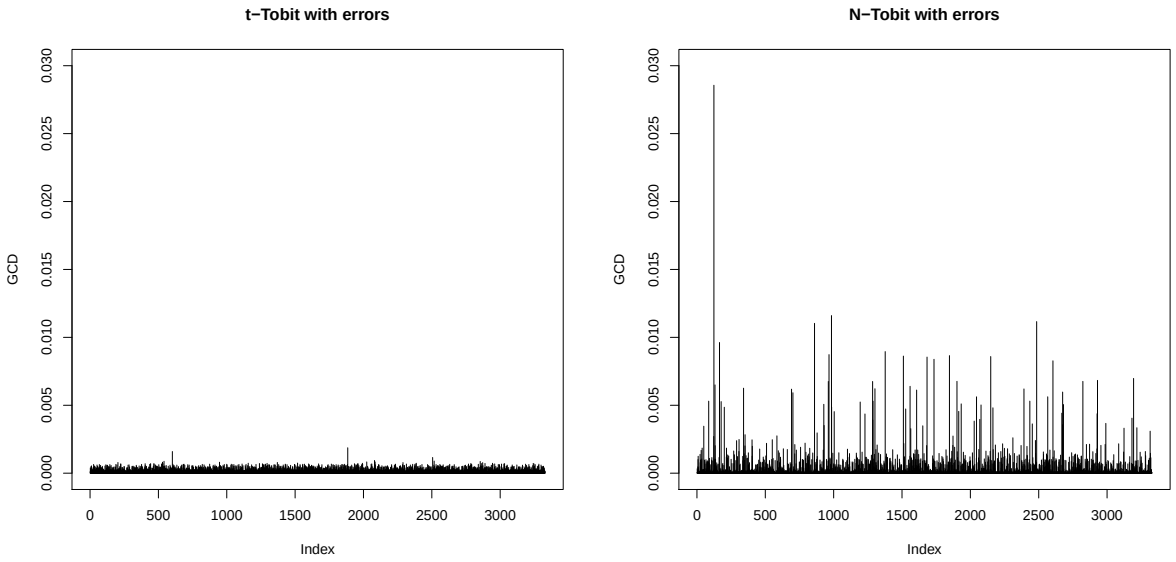

Figura 6.3: Distância de Cook generalizada para Modelo 1 (esquerda) e Modelo 2 (direita).

Capítulo 7

Conclusões

Neste trabalho foi introduzido um modelo tipo Tobit com erros nas covariáveis que assume uma distribuição t-Student multivariada para modelar conjuntamente o comportamento dos erros e das covariáveis latentes. Foram obtidas condições para identificar a verossimilhança do modelo proposto, as distribuições condicionais completas *a posteriori* dos parâmetros e das variáveis latentes do modelo proposto e foi introduzido um algoritmo ECM para aproximar os estimadores de máxima verossimilhança de todos os parâmetros, incluindo o grau de liberdade.

Foi realizado um estudo Monte Carlo para comparar o modelo proposto com alguns modelos previamente introduzidos na literatura no caso bayesiano e para avaliar a qualidade dos estimadores de máxima verossimilhança e a eficiência do algoritmo ECM proposto em diferentes cenários. Do ponto de vista baeysiano notou-se que, como esperado, estimativas tendem a ser melhores se a proporção de censura é pequena e também se o tamanho amostral é grande. Usualmente, as componentes da variância e o grau de liberdade são mal estimados se o modelo é mal especificado. Tais estimativas são ainda piores se assumem-se modelos livres de erros. Notou-se também que se os dados são gerados de uma distribuição com cauda pesada, obtem-se estimativas muito pobres para a variância das covariáveis latentes sempre que um modelo que assume distribuições com caudas mais leves para os erros e covariáveis latentes é ajustado mas o oposto não necessariamente ocorre. As estatísticas LPML e DIC mostram ser ferramentas ineficientes para comparar modelos com erros nas covariáveis com aqueles modelos livre de erros. Conclusões similares foram obtidas por Vidal e Iglesias (2008). Para esse propósito, estimativas *a posteriori* para as variâncias do modelo e das covariáveis latentes trabalharam como

uma ferramenta auxiliar na seleção de modelos. Do ponto de vista frequentista algumas das conclusões são que o EMV de $\boldsymbol{\Sigma}_\xi$ é usualmente viciado e que a falta de informação (pequenas amostras ou alta proporção de censura nas respostas) induz mais vício nesse estimador. Também conclui-se que há uma tendência de perda de robustez se o grau de liberdade é estimado mas os EMV são viciados se o valor de ν assumido no modelo não é o verdadeiro, principalmente para os parâmetros $\boldsymbol{\mu}_\xi$, $\boldsymbol{\Sigma}_\xi$, β_1 e σ_u.

Os dados sobre gastos ambulatoriais, reportados em Cameron e Trivedi (2010), foram analisados assumindo que a covariável renda é medida de forma imprecisa. Sob a ótica bayesiana conclui-se que o modelo proposto é o melhor para analisar tais dados. Em relação à inferência clássica, o p-valor para a renda foi aproximadamente 17% e a estimativa para $\boldsymbol{\Sigma}_v$ foi alta. Logo, conclui-se que é razoável assumir que tal covariável não pode ser considerada livre de erros. Também conclui-se que um modelo com cauda pesada fornece melhor ajuste que um modelo aproximadamente normal.

A principal crítica em relação ao modelo proposto é que ele assume o mesmo grau de liberdade para modelar conjuntamente o comportamento dos erros e das covariáveis latentes. Isso pode limitar o uso do modelo pois, de certa maneira, tal fato faz com que assumem-se distribuições de mesma cauda para modelar tais quantidades e isso pode não ser uma suposição apropriada em algumas situações práticas. No entanto, considerando o método de máxima verossimilhança, não seria uma tarefa fácil tratar um modelo com diferentes graus de liberdade para tais quantidades. Outra sugestão como trabalho futuro é considerar uma distribuição *a priori* para a condição de identificação. Isso resultaria em um modelo mais flexível.

Referências Bibliográficas

Amemiya, T. (1984). Tobit models: A survey. *Journal of Econometrics*, **24**, 3-61.

Amemiya, T. (1985). *Advanced econometrics*. Harvard University Press, Cambridge, MA.

Arellano-Valle, R. B., Bolfarine, H. (1995). On some characterizations of the t distribution. *Statistic and Probability Letters*, **25**, 179-85.

Arellano-Valle, R. B., Bolfarine, H. (1996). Elliptical Structural Models. *Communications in Statistics: Theory and Methods*, **25**, 2319-2341.

Arellano-Valle, R. B., Castro, L. M., González-Farias, G., Muñoz-Gajardo, K. A. (2012). Student-t censored regression model: properties and inference. *Statistical Methods and Applications*, **21**, 453-473.

Austin, P. C. (2002). Bayesian Extensions of the Tobit Model for Analyzing Measures of Health Status. *Medical Decision Making*, **22**, 152-162.

Barros, M., Galea, M., González, M., Leiva. V. (2010). Influence diagnostics in the tobit censored response model. *Statistical Methods & Applications*, **19**, (3), 379-397.

Bolfarine, H., Arellano-Valle, R. B. (1998). Weak Nondifferential Erros Models. *Statistics and Probability Letter*, **40**, 279-287.

Bolfarine, H., Arellano-Valle, R. B. (2005). Elliptical measurement error models: A Bayesian approach. Invited chapter (ch.22) in Bayesian Thinking: Modeling and Computation, edited by D. K. Dey and C. R. Rao, *Handbook of Statistic*, **25**, 669-688. Elsevier.

Cameron, A. C., Trivedi, P. K. (2010). *Microeconometrics Using Stata* (Revised ed.), College Station, TX: Stata Press.

Carlin, B., Louis, T. (2008). *Bayesian Methods for Data Analysis (Texts in Statistical Science)*. New York: Chapman and Hall/CRC.

Carriquiry, A. L., Gianola, D., Fernando, R. L. (1987). Mixed-model analysis of a censored normal distribution with reference to animal breeding. *Biometrics*, **43**, 929-939.

Chib, S. (1992). Bayes inference in the Tobit censored regression model. *Journal of Econometrics*, **51**, 79-99.

Cowles, M. K., Carlin, B. P., Connett, J. E. (1996). Bayesian Modeling of Longitudinal Ordinal Clinical Trial Compliance Data With Nonignorable Missingness. *Journal of the American Statistical Association*, **91**, 86-98.

Dellaportas, P., Stephens, D. A. (1995). Bayesian Analysis of Errors-in-Variables Regression Models. *Biometrics*, **51**, (3), 1085-1095.

Dempster, A. P., Laird, N. M., Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, **39** (1), 1-38.

Dey, D. K., Chen, M. H., Chang, H. (1997). Bayesian approach for the nonlinear random effects models. *Biometrics*, **53**, 1239-1252.

Fang, K. T., Kotz, S., Ng, K. W. (1990). *Symmetric Multivariate and Related Distributions*. Chapman and Hall, London and New York.

Fuller, W. A. (1987). *Measurement error models*. John Wiley & Sons. New York.

Gelfand, A. G. and Sahu, S.K. (1999). Identifiability, improper priors, and Gibbs sampling for generalized linear models. *Journal of the American Statistical Association*, **94** (445), 247–253.

Gilks, W. R., Best, N. G. e Tan, K. K. C. (1995). Adaptive rejection metropolis sampling within gibbs sampling, *Applied Statatistics*, **44**, 455-472.

Gilks, W. R. and Wild, P. (1992). Adaptive rejection sampling for gibbs sampling, *Applied Statatistics*, **41**, 337–348.

Gleser, L. J. (1992). The importance of assessing measurement reliability in multivariate regression, *Journal of the American Statistical Association*, **87**, 696-707.

Greene, W. H. (1990). *Econometric analysis*. Macmillan, New York, NY.

Hamilton, B. H. (1999). HMO selection and medicare costs: bayesian MCMC estimation of a robust panel data with survival. *Health Economics*, **8**, 403–414.

Ho, H. J., Lin, T., Chen, H., Wang, W. (2012) Some results on the truncated multivariate t distribution. *Journal of Statistical Planning and Inference*, **142**, 25-40.

Lachos, V. H., Bandyopadhyay, D., Dey, D. K. (2011). Linear and Nonlinear Mixed-Effects Models for Censored HIV Viral Loads Using Normal/Independent Distributions. *Biometrics*, **67**, 1594-1604.

Maddala, G. S. (1983). *Limited-dependent and qualitative variables in econometrics*. Cambridge University Press, New York, NY.

Marchenko, Y. V., and Genton, M. G. (2012). A Heckman selection-t model. *Journal of the American Statistical Association*, **107**, 304-317.

Matos, L. A., Prates, M. O., Chen, M. H., Lachos, V. H. (2013). Likelihood-based inference for mixed-effects models with censored response using the multivariate-t distribution. *Statistica Sinica*, **23**, 1323-1342.

Meng, X. L., Rubin, D. B. (1993). Maximum likelihood estimation via the ECM algorithm: A general framework. *Biometrika*, **80**, 267-278.

Massuia, M. B., Cabral, C. R. B., Matos, L. A., Lachos, V. H. (2014). Influence diagnostics for Student-t censored linear regression models. *Statistics: A Journal of Theoretical and Applied Statistics*, (a aparecer - disponível online).

Olsen, R.J. (1978). Note on uniqueness of the maximum likelihood estimator for the Tobit model. *Econometrica*, **46**. 1211-1215.

Polasek, W., Krause, A. (1993). Bayesian Regression Model with Simple Errors in Variables Structure. *Journal of the Royal Statistical Society. Series D (The Statistician)*, **42**, 571-580.

R Core Team. (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. http://www.R-project.org/

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B*, **64**, 583-639.

Swartz, T., Haitovsky, Y., Vexler, A., and Yang, T. (2004). Bayesian identifiability and misclassification in multinomial data. *The Canadian Journal of Statistics* **32,** (3), 1–18.

Sweeting, T. J. (1987). Approximate Bayesian analysis of censored survival data. *Biometrika*, **74**, 809-916.

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica*, **26**, 24-36.

van Dyk, D. A. and Meng, X.L. (2001). The art of data augmentation. *Journal of Computational and Graphical Statistics* (discussion paper), **10**(1), 1–50.

Vidal, I., Iglesias, P. (2008). Comparison between a measurement error model and a linear model without measurement error. *Computational Statistics and Data Analysis*, **53**, 92-102.

Wang, L. (1998). Estimation of censored linear error-in-variables models. *Journal of Econometrics*, **84**, 383-400.

Wang, L. (2002). A simple adjustement for measurement erros in some limited dependent variable. *Statistics & Probability Letters*, **58**, 427-433.

Wang, L. and Hsiao, C. (2007). Two-estage estimation of limited dependent variable models with error-in-variables. *Econometrics Journal*, **10**, 426-438.

Apêndice A

As distribuições condicionais completas *a posteriori* para o modelo t-Tobit com erros nas covariáveis

Neste apêndice são apresentados os cálculos para as distribuições condicionais completas *a posteriori* para todos os parâmetros e variáveis latentes. A dccp para a quantidade latente $\mathbf{g} = (g_1, \ldots, g_n)$ é dada por

$$
\begin{aligned}
f(\mathbf{g} \mid \eta, \mathbf{D}_o, \boldsymbol{\theta}) \;\propto\; & \prod_{i=1}^{n} \left(\frac{1}{\sqrt{g_i^{-1}}} \exp\left(-\frac{1}{2g_i^{-1}\sigma_w} [z_i - (\gamma_1 + \boldsymbol{\gamma}_2' \mathbf{x}_i + \boldsymbol{\beta}_3' \mathbf{b}_i)]^2 \right) \right)^{1-d_i} \\
\times \; & \left(\frac{1}{\sqrt{g_i^{-1}}} \exp\left(-\frac{1}{2g_i^{-1}\sigma_w} [y_i - (\gamma_1 + \boldsymbol{\gamma}_2' \mathbf{x}_i + \boldsymbol{\beta}_3' \mathbf{b}_i)]^2 \right) \right)^{d_i} \\
\times \; & (g_i^{-1})^{-k/2} \exp\left(-\frac{1}{2} (\mathbf{x}_i - \boldsymbol{\mu}_x)' (g_i^{-1})^{-1} \boldsymbol{\Sigma}_x^{-1} (\mathbf{x}_i - \boldsymbol{\mu}_x) \right) \\
\times \; & g_i^{\nu/2-1} \exp\left(-\frac{g_i}{2/\nu} \right) \\
\propto \; & \prod_{i=1}^{n} \left(g_i^{1/2} \exp\left(-\frac{g_i}{2\sigma_w / {a_1}_i^2} \right) \right)^{1-d_i} \left(g_i^{1/2} \exp\left(-\frac{g_i}{2\sigma_w / {a_2}_i^2} \right) \right)^{d_i} \\
\times \; & g_i^{k/2} \exp\left(-\frac{g_i}{2/q(\mathbf{x}_i)} \right) g_i^{\nu/2-1} \exp\left(-\frac{g_i}{2/\nu} \right) \\
\propto \; & \prod_{i=1}^{n} g_i^{(1-d_i)/2} \exp\left(-\frac{g_i}{2\sigma_w / [{a_1}_i^2 (1-d_i)]} \right) g_i^{d_i/2} \exp\left(-\frac{g_i}{2\sigma_w / ({a_2}_i^2 d_i)} \right) \\
\times \; & g_i^{k/2} \exp\left(-\frac{g_i}{2/q(\mathbf{x}_i)} \right) g_i^{\nu/2-1} \exp\left(-\frac{g_i}{2/\nu} \right)
\end{aligned}
$$

$$\propto \prod_{i=1}^{n} g_i^{\frac{1+k+\nu}{2}-1} \exp\left(-\frac{g_i}{2\sigma_w/[(1-d_i)a_{1t}^2 + d_i a_{2t}^2 + \sigma_w q(\mathbf{x}_i) + \sigma_w \nu]}\right),$$

em que $a_{1i} = z_i - (\gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x}_i + \boldsymbol{\beta}_3'\mathbf{b}_i)$ e $a_{2i} = y_i - (\gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x}_i + \boldsymbol{\beta}_3'\mathbf{b}_i)$. Logo, para todo $i = 1, \ldots, n$,

$$g_i \mid \eta_i, D_{oi}, \boldsymbol{\theta} \sim G\left(\frac{1+k+\nu}{2}, \frac{2\sigma_w}{(1-d_i)a_{1t}^2 + d_i a_{2t}^2 + \sigma_w q(\mathbf{x}_i) + \sigma_w \nu}\right).$$

Para γ_1 a dccp é

$$\begin{aligned}
\pi(\gamma_1 \mid \boldsymbol{\theta}_{-\gamma_1}, \mathbf{D}_c) &\propto \exp\left(-\frac{1}{2\sigma_1}(\gamma_1 - \mu_1)^2\right) \\
&\times \prod_{i=1}^{n}\left(\exp\left(-\frac{1}{2g_i^{-1}\sigma_w}[z_i - (\gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x}_i + \boldsymbol{\beta}_3'\mathbf{b}_i)]^2\right)\right)^{1-d_i} \\
&\times \left(\exp\left(-\frac{1}{2g_i^{-1}\sigma_w}[y_i - (\gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x}_i + \boldsymbol{\beta}_3'\mathbf{b}_i)]^2\right)\right)^{d_i} \\
&\propto \exp\left(-\frac{1}{2\sigma_1}(\gamma_1 - \mu_1)^2\right) \\
&\times \prod_{i=1}^{n}\exp\left(-\frac{1}{2\frac{g_i^{-1}\sigma_w}{1-d_i}}[\gamma_1 - (z_i - \boldsymbol{\gamma}_2'\mathbf{x}_i - \boldsymbol{\beta}_3'\mathbf{b}_i)]^2\right) \\
&\times \exp\left(-\frac{1}{2\frac{g_i^{-1}\sigma_w}{d_i}}[\gamma_1 - (y_i - \boldsymbol{\gamma}_2'\mathbf{x}_i - \boldsymbol{\beta}_3'\mathbf{b}_i)]^2\right) \\
&\propto \exp\left(-\frac{1}{2\sigma_1}(\gamma_1 - \mu_1)^2\right) \\
&\times \prod_{i=1}^{n}\exp\left(-\frac{1}{2g_i^{-1}\sigma_w}\{\gamma_1 - [(1-d_i)z_i + d_i y_i - \boldsymbol{\gamma}_2'\mathbf{x}_i - \boldsymbol{\beta}_3'\mathbf{b}_i]\}^2\right) \\
&\propto \exp\left(-\frac{1}{2\sigma_1}(\gamma_1 - \mu_1)^2\right)\exp\left(-\frac{1}{2\frac{\sigma_w}{\sum_{i=1}^n g_i}}\left(\gamma_1 - \frac{\sum_{i=1}^n a_{3t}}{\sum_{i=1}^n g_i}\right)^2\right) \\
&\propto \exp\left(-\frac{1}{2H_1}\left\{\gamma_1 - H_1[\sigma_1^{-1}\mu_1 + \sigma_w^{-1}\sum_{i=1}^{n} g_i((1-d_i)z_i + d_i y_i - \boldsymbol{\gamma}_2'\mathbf{x}_i - \boldsymbol{\beta}_3'\mathbf{b}_i)]\right\}^2\right),
\end{aligned}$$

em que $H_1^{-1} = \sigma_1^{-1} + \sigma_w^{-1}\sum_{i=1}^n g_i$. Logo,

$$\gamma_1 \mid \boldsymbol{\theta}_{-\gamma_1}, \mathbf{D}_c \sim N_1\left(H_1\left\{\sigma_1^{-1}\mu_1 + \sigma_w^{-1}\sum_{i=1}^{n} g_i\left[(1-d_i)z_i + d_i y_i - \boldsymbol{\gamma}_2'\mathbf{x}_i - \boldsymbol{\beta}_3'\mathbf{b}_i\right]\right\}, H_1\right).$$

Para $\boldsymbol{\gamma}_2$ segue que a dccp é

$$\begin{aligned}
\pi(\boldsymbol{\gamma}_2 \mid \boldsymbol{\theta}_{-\boldsymbol{\gamma}_2}, \mathbf{D}_c) &\propto \exp\left(-\frac{1}{2}(\boldsymbol{\gamma}_2 - \boldsymbol{\mu}_2)'{\boldsymbol{\Sigma}_2}^{-1}(\boldsymbol{\gamma}_2 - \boldsymbol{\mu}_2)\right) \\
&\times \prod_{i=1}^{n}\left(\exp\left(-\frac{1}{2g_i^{-1}\sigma_w}[z_i - (\gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x}_i + \boldsymbol{\beta}_3'\mathbf{b}_i)]^2\right)\right)^{1-d_i}
\end{aligned}$$

$$
\begin{aligned}
&\times \quad \left(\exp\left(-\frac{1}{2g_i^{-1}\sigma_w}[y_i-(\gamma_1+\boldsymbol{\gamma}_2'\mathbf{x}_i+\boldsymbol{\beta}_3'\mathbf{b}_i)]^2\right)\right)^{d_i}\\
&\propto \quad \exp\left(-\frac{1}{2}(\boldsymbol{\gamma}_2-\boldsymbol{\mu}_2)'{\boldsymbol{\Sigma}_2}^{-1}(\boldsymbol{\gamma}_2-\boldsymbol{\mu}_2)\right)\\
&\times \quad \prod_{i=1}^{n}\exp\left(-\frac{1}{2\frac{g_i^{-1}\sigma_w}{1-d_i}}[\mathbf{x}_i'\boldsymbol{\gamma}_2-(z_i-\gamma_1-\boldsymbol{\beta}_3'\mathbf{b}_i)]^2\right)\\
&\times \quad \exp\left(-\frac{1}{2\frac{g_i^{-1}\sigma_w}{d_i}}[\mathbf{x}_i'\boldsymbol{\gamma}_2-(y_i-\gamma_1-\boldsymbol{\beta}_3'\mathbf{b}_i)]^2\right)\\
&\propto \quad \exp\left(-\frac{1}{2}(\boldsymbol{\gamma}_2-\boldsymbol{\mu}_2)'{\boldsymbol{\Sigma}_2}^{-1}(\boldsymbol{\gamma}_2-\boldsymbol{\mu}_2)\right)\\
&\times \quad \prod_{i=1}^{n}\exp\left(-\frac{1}{2g_i^{-1}\sigma_w}\{\mathbf{x}_i'\boldsymbol{\gamma}_2-[(1-d_i)z_i+d_iy_i-\gamma_1-\boldsymbol{\beta}_3'\mathbf{b}_i]\}^2\right)\\
&\propto \quad \exp\left(-\frac{1}{2}(\boldsymbol{\gamma}_2-\boldsymbol{\mu}_2)'{\boldsymbol{\Sigma}_2}^{-1}(\boldsymbol{\gamma}_2-\boldsymbol{\mu}_2)\right)\\
&\times \quad \exp\left(-\frac{1}{2\sigma_w}\sum_{i=1}^{n}g_i(\mathbf{x}_i'\boldsymbol{\gamma}_2-[(1-d_i)z_i+d_iy_i-\gamma_1-\boldsymbol{\beta}_3'\mathbf{b}_i])^2\right)\\
&\propto \quad \exp\left(-\frac{1}{2}(\boldsymbol{\gamma}_2-\boldsymbol{\mu}_2)'{\boldsymbol{\Sigma}_2}^{-1}(\boldsymbol{\gamma}_2-\boldsymbol{\mu}_2)\right)\\
&\times \quad \exp\left(-\frac{1}{2}\left[\boldsymbol{\gamma}_2-\left(\sum_{i=1}^{n}g_i\mathbf{x}_i\mathbf{x}_i'\right)^{-1}\left(\sum_{i=1}^{n}g_i[(1-d_i)z_i+d_iy_i-\gamma_1-\boldsymbol{\beta}_3'\mathbf{b}_i]\mathbf{x}_i\right)\right]'\right.\\
&\times \quad \left.\left[\sigma_w\left(\sum_{i=1}^{n}g_i\mathbf{x}_i\mathbf{x}_i'\right)^{-1}\right]^{-1}\right.\\
&\times \quad \left.\left[\boldsymbol{\gamma}_2-\left(\sum_{i=1}^{n}g_i\mathbf{x}_i\mathbf{x}_i'\right)^{-1}\left(\sum_{i=1}^{n}g_i[(1-d_i)z_i+d_iy_i-\gamma_1-\boldsymbol{\beta}_3'\mathbf{b}_i]\mathbf{x}_i\right)\right]\right)\\
&\propto \quad \exp\left(-\frac{1}{2}\left\{\boldsymbol{\gamma}_2-\mathbf{H}_2\left[\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2+\sigma_w^{-1}\sum_{i=1}^{n}g_i\left((1-d_i)z_i+d_iy_i-\gamma_1-\boldsymbol{\beta}_3'\mathbf{b}_i\right)\mathbf{x}_i\right]\right\}'\right.\\
&\times \quad \left.{\mathbf{H}_2}^{-1}\left\{\boldsymbol{\gamma}_2-\mathbf{H}_2\left[\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2+\sigma_w^{-1}\sum_{i=1}^{n}g_i\left((1-d_i)z_i+d_iy_i-\gamma_1-\boldsymbol{\beta}_3'\mathbf{b}_i\right)\mathbf{x}_i\right]\right\}\right),
\end{aligned}
$$

onde $\mathbf{H}_2^{-1}=\boldsymbol{\Sigma}_2^{-1}+\sigma_w^{-1}\sum_{i=1}^{n}g_i\mathbf{x}_i\mathbf{x}_i'$. Logo,

$$
\boldsymbol{\gamma}_2 \mid \boldsymbol{\theta}_{-\boldsymbol{\gamma}_2},\mathbf{D}_c \sim N_k\left(\mathbf{H}_2\left\{\boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2+\sigma_w^{-1}\sum_{i=1}^{n}g_i\left[(1-d_i)z_i+d_iy_i-\gamma_1-\boldsymbol{\beta}_3'\mathbf{b}_i\right]\mathbf{x}_i\right\},\mathbf{H}_2\right).
$$

Similarmente, obtem-se a dccp de $\boldsymbol{\beta}_3$ que é a seguinte distribuição normal

$$\boldsymbol{\beta}_3 \mid \boldsymbol{\theta}_{-\boldsymbol{\beta}_3}, \mathbf{D}_c \sim N_{k_1}\left(\mathbf{H}_3\left\{\boldsymbol{\Sigma}_3^{-1}\boldsymbol{\mu}_3 + \sigma_w^{-1}\sum_{i=1}^{n} g_i\left[(1-d_i)z_i + d_i y_i - \gamma_1 - \boldsymbol{\gamma}_2'\mathbf{x}_i\right]\mathbf{b}_i\right\}, \mathbf{H}_3\right),$$

onde $\mathbf{H}_3^{-1} = \boldsymbol{\Sigma}_3^{-1} + \sigma_w^{-1}\sum_{i=1}^{n} g_i \mathbf{b}_i\mathbf{b}_i'$.

A dccp de σ_w também possui forma fechada como provada no que segue

$$\begin{aligned}
\pi(\sigma_w \mid \boldsymbol{\theta}_{-\sigma_w}, \mathbf{D}_c) &\propto \sigma_w^{-\alpha_w - 1}\exp\left(-\frac{\beta_w}{\sigma_w}\right) \\
&\times \prod_{i=1}^{n}\left(\sigma_w^{-1/2}\exp\left(-\frac{1}{2g_i^{-1}\sigma_w}[z_i - (\gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x}_i + \boldsymbol{\beta}_3'\mathbf{b}_i)]^2\right)\right)^{1-d_i} \\
&\times \left(\sigma_w^{-1/2}\exp\left(-\frac{1}{2g_i^{-1}\sigma_w}[y_i - (\gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x}_i + \boldsymbol{\beta}_3'\mathbf{b}_i)]^2\right)\right)^{d_i} \\
&\propto \sigma_w^{-\alpha_w - 1}\exp\left(-\frac{\beta_w}{\sigma_w}\right) \\
&\times \sigma_w^{-\frac{1}{2}\sum_{i=1}^{n}(1-d_i)}\exp\left(-\frac{1}{2\sigma_w}\sum_{i=1}^{n} g_i[z_i - (\gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x}_i + \boldsymbol{\beta}_3'\mathbf{b}_i)]^2(1-d_i)\right) \\
&\times \sigma_w^{-\frac{1}{2}\sum_{i=1}^{n} d_i}\exp\left(-\frac{1}{2\sigma_w}\sum_{i=1}^{n} g_i[y_i - (\gamma_1 + \boldsymbol{\gamma}_2'\mathbf{x}_i + \boldsymbol{\beta}_3'\mathbf{b}_i)]^2 d_i\right) \\
&\propto \sigma_w^{-\left(\alpha_w + \frac{n}{2}\right) - 1}\exp\left(-\frac{1}{\sigma_w}\left\{\beta_w + \frac{1}{2}\sum_{i=1}^{n} g_i[(1-d_i)a_{1t}^2 + d_i a_{2t}^2]\right\}\right).
\end{aligned}$$

Logo,

$$\sigma_w \mid \boldsymbol{\theta}_{-\sigma_w}, \mathbf{D}_c \sim IG\left(\alpha_w + \frac{n}{2}, \beta_w + \frac{1}{2}\sum_{i=1}^{n} g_i\left[(1-d_i)a_{1t}^2 + d_i a_{2t}^2\right]\right).$$

Para $\boldsymbol{\mu}_x$ pode-se provar que sua dccp é

$$\begin{aligned}
\pi(\boldsymbol{\mu}_x \mid \boldsymbol{\theta}_{-\boldsymbol{\mu}_x}, \mathbf{D}_c) &\propto \exp\left(-\frac{1}{2}(\boldsymbol{\mu}_x - \boldsymbol{\mu}_4)'\boldsymbol{\Sigma}_4^{-1}(\boldsymbol{\mu}_x - \boldsymbol{\mu}_4)\right) \\
&\times \prod_{i=1}^{n}\exp\left(-\frac{1}{2}(\mathbf{x}_i - \boldsymbol{\mu}_x)' g_i \boldsymbol{\Sigma}_x^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_x)\right) \\
&\propto \exp\left(-\frac{1}{2}(\mu_x - \boldsymbol{\mu}_4)'\boldsymbol{\Sigma}_4^{-1}(\boldsymbol{\mu}_x - \boldsymbol{\mu}_4)\right) \\
&\times \exp\left(-\frac{1}{2}\left(\boldsymbol{\mu}_x - \frac{1}{\sum_{i=1}^{n} g_i}\sum_{i=1}^{n} g_i\mathbf{x}_i\right)'\sum_{i=1}^{n} g_i\Sigma_x^{-1}\left(\mu_x - \frac{1}{\sum_{i=1}^{n} g_i}\sum_{i=1}^{n} g_i\mathbf{x}_i\right)\right) \\
&\propto \exp\left(-\frac{1}{2}\left[\boldsymbol{\mu}_x - \mathbf{H}_4\left(\boldsymbol{\Sigma}_4^{-1}\boldsymbol{\mu}_4 + \boldsymbol{\Sigma}_x^{-1}\sum_{i=1}^{n} g_i\mathbf{x}_i\right)\right]'\right.
\end{aligned}$$

$$\times \quad \mathbf{H}_4{}^{-1}\left[\boldsymbol{\mu}_x - \mathbf{H}_4\left(\boldsymbol{\Sigma}_4^{-1}\boldsymbol{\mu}_4 + \boldsymbol{\Sigma}_x^{-1}\sum_{i=1}^{n} g_i\mathbf{x}_i\right)\right]\Bigg),$$

onde $\mathbf{H}_4^{-1} = \boldsymbol{\Sigma}_4^{-1} + \boldsymbol{\Sigma}_x^{-1}\sum_{i=1}^{n} g_i$. Logo,

$$\boldsymbol{\mu}_x \mid \boldsymbol{\theta}_{-\boldsymbol{\mu}_x}, \mathbf{D}_c \sim N_k\left(\mathbf{H}_4\left(\boldsymbol{\Sigma}_4^{-1}\boldsymbol{\mu}_4 + \boldsymbol{\Sigma}_x^{-1}\sum_{i=1}^{n} g_i\mathbf{x}_i\right), \mathbf{H}_4\right).$$

Para a componente de variância $\boldsymbol{\Sigma}_x$ tem-se

$$\begin{aligned}
\pi(\boldsymbol{\Sigma}_x \mid \boldsymbol{\theta}_{-\boldsymbol{\Sigma}_x}, \mathbf{D}_c) &\propto (\det(\boldsymbol{\Sigma}_x))^{-\frac{m_1+k+1}{2}} \exp\left(-\frac{1}{2}\operatorname{tr}(\boldsymbol{\Psi}_1\boldsymbol{\Sigma}_x^{-1})\right) \\
&\times \prod_{i=1}^{n} (\det(\boldsymbol{\Sigma}_x))^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\operatorname{tr}((\mathbf{x}_i - \boldsymbol{\mu}_x)' g_i \boldsymbol{\Sigma}_x^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_x))\right) \\
&\propto (\det(\boldsymbol{\Sigma}_x))^{-\frac{(m_1+n)+k+1}{2}} \exp\left(-\frac{1}{2}\operatorname{tr}(\boldsymbol{\Psi}_1\boldsymbol{\Sigma}_x^{-1})\right) \\
&\times \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\operatorname{tr}((\mathbf{x}_i - \boldsymbol{\mu}_x)' g_i \boldsymbol{\Sigma}_x^{-1}(\mathbf{x}_i - \boldsymbol{\mu}_x))\right) \\
&\propto (\det(\boldsymbol{\Sigma}_x))^{-\frac{(m_1+n)+k+1}{2}} \\
&\times \exp\left(-\frac{1}{2}\operatorname{tr}\left(\boldsymbol{\Psi}_1\boldsymbol{\Sigma}_x^{-1} + \sum_{i=1}^{n} g_i(\mathbf{x}_i - \boldsymbol{\mu}_x)(\mathbf{x}_i - \boldsymbol{\mu}_x)'\boldsymbol{\Sigma}_x^{-1}\right)\right).
\end{aligned}$$

Logo, segue que

$$\boldsymbol{\Sigma}_x \mid \boldsymbol{\theta}_{-\boldsymbol{\Sigma}_x}, \mathbf{D}_c \sim IW\left(m_1 + n, \boldsymbol{\Psi}_1 + \sum_{i=1}^{n} g_i(\mathbf{x}_i - \boldsymbol{\mu}_x)(\mathbf{x}_i - \boldsymbol{\mu}_x)'\right).$$

Finalmente, para o grau de liberdade tem-se que sua dccp é

$$\begin{aligned}
\pi(\nu \mid \boldsymbol{\theta}_{-\nu}, \mathbf{D}_c) &\propto \exp\left(-\frac{\nu}{\lambda}\right)\prod_{i=1}^{n}\frac{1}{\Gamma(\nu/2)(2/\nu)^{\nu/2}} g_i^{\nu/2-1}\exp\left(-\frac{g_i}{2/\nu}\right), \quad \nu > 2 \\
&\propto \left[\frac{1}{\Gamma(\nu/2)(2/\nu)^{\nu/2}}\right]^n \exp\left(-\nu\left[\frac{1}{2}\left(\sum_{i=1}^{n} g_i - \log(g_i)\right) + \frac{1}{\lambda}\right]\right), \quad \nu > 2,
\end{aligned}$$

que não pertence à alguma família conhecida de distribuições.

Apêndice B

Momentos da distribuição t-Student truncada

Neste apêndice encontram-se alguns resultados e provas relacionados aos momentos da distribuição t-Student truncada que são necessários na construção do algoritmo ECM e no cálculo da matriz de informação observada.

Corolário 1. *Se $Z \sim t_1(0, 1, \nu)$, então os momentos ímpares e pares de Z são*

$$\begin{aligned} E(Z^{2m+1} \mid Z \leq a) &= \frac{\nu^{m+1/2}}{\pi^{1/2} T_1(a \mid \nu)} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \sum_{j=0}^{m} \binom{m}{j} \frac{(-1)^{m-j+1}}{\nu - 2j - 1} \left(1 + \frac{a^2}{\nu}\right)^{-\frac{\nu-2j-1}{2}}, \\ E(Z^{2m} \mid Z \leq a) &= \frac{\nu^m}{T_1(a \mid \nu)} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \sum_{j=0}^{m} \binom{m}{j} (-1)^{m-j} T_1\left(a\left(\frac{\nu - 2j}{\nu}\right)^{1/2} \mid \nu - 2j\right) \frac{\Gamma\left(\frac{\nu}{2} - j\right)}{\Gamma\left(\frac{\nu+1}{2} - j\right)}. \end{aligned}$$

A prova do Corolário 1 segue direto do Teorema 1 de Ho *et al.* (2012).

Corolário 2. *Se $X \sim t_1(\mu, \sigma, \nu)$ então*

$$E\left(m_{(X,1,\nu)}^r X^m \mid X \leq a\right) = \left(\frac{\nu+1}{\nu}\right)^r \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{\Gamma\left(\frac{\nu+2r}{2}\right)}{\Gamma\left(\frac{\nu+1+2r}{2}\right)} \frac{T_1(a \mid \mu, \sigma_r, \nu + 2r)}{T_1(a \mid \mu, \sigma, \nu)} E(W^m \mid W \leq a),$$

onde $W \sim t_1(\mu, \sigma_r, \nu + 2r)$, $\sigma_r = \frac{\nu}{\nu+2r}\sigma$ e $\nu + 2r > 0$.

A prova do Corolário 2 segue direto da Proposição 2 de Matos *et al.* (2013).

Lema 1. *Suponha que $\eta \mid \mathbf{x}, g \sim N_1(\mu(\mathbf{x}), g^{-1}\sigma_\eta)$, $\mathbf{x} \mid g \sim N_k(\boldsymbol{\mu}_x, g^{-1}\boldsymbol{\Sigma}_x)$ e $g \sim G(\nu/2, 2/\nu)$. Então tem-se:*

i. A distribuição condicional de g dado $\mathbf{x}$ é

$$g \mid \mathbf{x} \sim G\left(\frac{\nu + k}{2}, \frac{2}{\nu + q(\mathbf{x})}\right).$$

ii. A distribuição condicional de g dado $\mathbf{x}$ e η é

$$g \mid \mathbf{x}, \eta \sim G\left(\frac{\nu+k+1}{2}, \frac{2}{\nu+q(\mathbf{x})+q_*(\eta)}\right),$$

onde $q_*(\eta) = (\eta - \mu(\mathbf{x}))^2/\sigma_\eta$ e $q(\mathbf{x}) = (\mathbf{x}-\boldsymbol{\mu}_x)\boldsymbol{\Sigma}_x^{-1}(\mathbf{x}-\boldsymbol{\mu}_x)'$.

iii. Condicional em $\mathbf{x}$ e $\eta \leq a$, a distribuição de g tem a seguinte função de densidade de probabilidade:

$$f_{g|\mathbf{x},\eta\leq a}(u) = \frac{1}{T_1\left(a; \mu(\mathbf{x}), \sigma_\eta\left(\frac{\nu+q(\mathbf{x})}{\nu+k}\right)\right)} f_{g|\mathbf{x}}(u)\Phi_1\left(a; \mu(\mathbf{x}), u^{-1}\sigma_\eta\right), \quad u>0,$$

onde $f_{g|\mathbf{x}}(\cdot)$ é a fdp da distribuição em (i.).

A prova do Lema 1 é direta e por isso será omitida.

Lema 2. *Suponha que $\eta \mid \mathbf{x}, g \sim N_1(\mu(\mathbf{x}), g^{-1}\sigma_\eta)$, $\mathbf{x} \mid g \sim N_k(\boldsymbol{\mu}_x, g^{-1}\boldsymbol{\Sigma}_x)$ e $g \sim G(\nu/2, 2/\nu)$. Então, as esperanças condicionais de $g\eta^m$, $m \geq 0$, e $\log g$ dados $\mathbf{x}$ e $\eta \leq a$ são dadas, respectivamente, por*

$$\begin{aligned}
E(g\eta^m \mid \eta \leq a, \mathbf{x}) &= \left(\frac{\nu+k}{\nu+q(\mathbf{x})}\right) E\left(\eta^m\left(m_{(\eta,1,\nu+k)}\right) \mid \eta \leq a, \mathbf{x}\right), \\
E(\log g \mid \eta \leq a, \mathbf{x}) &= \psi\left(\frac{\nu+k+1}{2}\right) + \log(2) + \log\left(\frac{\nu+k}{\nu+q(\mathbf{x})}\right) \\
&\quad -E(\log(\nu+k+q(\eta)) \mid \eta \leq a, \mathbf{x}).
\end{aligned}$$

Prova: Assumindo as suposições dadas e considerando as propriedades de esperança condicional segue que

$$\begin{aligned}
E(g\eta^m \mid \eta \leq a, \mathbf{x}) &= E(E(g \mid \eta, \mathbf{x})\eta^m \mid \eta \leq a, \mathbf{x}) \\
&= E\left(\left(\frac{\nu+k+1}{\nu+q(\mathbf{x})+q^*(\eta)}\right)\eta^m \mid \eta \leq a, \mathbf{x}\right) \\
&= \left(\frac{\nu+k}{\nu+q(\mathbf{x})}\right) E\left(\left(m_{(\eta,1,\nu+k)}\right)\eta^m \mid \eta \leq a, \mathbf{x}\right);
\end{aligned}$$

$$\begin{aligned}
E(\log g \mid \eta \leq a, \mathbf{x}) &= E(E(\log g \mid \eta, \mathbf{x}) \mid \eta \leq a, \mathbf{x}) \\
&= \psi\left(\frac{\nu+k+1}{2}\right) + \log(2) - E(\log(\nu+q(\mathbf{x})+q^*(\eta)) \mid \eta \leq a, \mathbf{x}) \\
&= \psi\left(\frac{\nu+k+1}{2}\right) + \log(2) + \log\left(\frac{\nu+k}{\nu+q(\mathbf{x})}\right) - E(\log(\nu+k+q(\eta)) \mid \eta \leq a, \mathbf{x}).
\end{aligned}$$

□

Para o cálculo de $E\left(\left(m_{(\eta,1,\nu+k)}\right)\eta^m \mid \eta \leq a, \mathbf{x}\right)$ utiliza-se o resultado apresentado no Corolário 2, lembrando que, se $X \sim t_1(\mu, \sigma, \nu)$, então

$$E(X^m \mid X \leq a) = \sum_{j=0}^{m} \binom{m}{j} \mu^{m-j} \sigma^{j/2} E\left(Z^j \mid Z \leq \frac{a-\mu}{\sqrt{\sigma}}\right),$$

onde $Z \sim t_1(0, 1, \nu)$. Por sua vez, $E\left(Z^j \mid Z \leq \frac{a-\mu}{\sqrt{\sigma}}\right)$ pode ser obtido através dos resultados apresentados no Corolário 1. As esperanças que envolvem a função $\log(\nu + k + q(\eta))$ são obtidas de forma numérica, através do comando `integrate` do software R (R Core Team, 2014).

Apêndice C

Maximizando a função de log-verossimilhança completa do modelo t-Tobit com erros nas covariáveis

Outro resultado que é necessário para a construção do algoritmo ECM é o primeiro diferencial da função $Q(\boldsymbol{\theta} \mid \hat{\boldsymbol{\theta}}^{(t)})$ dada em (4.4). Esse diferencial é apresentado no seguinte lema.

Lema 3. *O primeiro diferencial da função $Q(\boldsymbol{\theta} \mid \hat{\boldsymbol{\theta}}^{(t)})$, dada em (4.4), é*

$$
\begin{aligned}
\mathrm{d}\, Q(\boldsymbol{\theta} \mid \widehat{\boldsymbol{\theta}}^{(t)}) &= \left(-\frac{n}{2\sigma_w} + \frac{1}{2\sigma_w^2}\sum_{i=1}^{n}(\widehat{g_i S_{\eta_i}})^{(t)}\right)(\mathrm{d}\,\sigma_w) + \left(\frac{1}{\sigma_w}\sum_{i=1}^{n}(\widehat{g_i\eta_{0i}})^{(t)}\right)(\mathrm{d}\,\gamma_1) \\
&+ \left(\frac{1}{\sigma_w}\sum_{i=1}^{n}(\widehat{g_i\eta_{0i}})^{(t)}\mathbf{x}_i'\right)(\mathrm{d}\,\boldsymbol{\gamma}_2) + \left(\frac{1}{\sigma_w}\sum_{i=1}^{n}(\widehat{g_i\eta_{0i}})^{(t)}\mathbf{b}_i'\right)(\mathrm{d}\,\boldsymbol{\gamma}_3) \\
&+ \left(\frac{1}{2}\sum_{i=1}^{n}(\widehat{g}_i)^{(t)}\,\mathrm{vec}(\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_{\mathbf{x}_i}\boldsymbol{\Sigma}_x^{-1})'\mathbf{D} - \frac{n}{2}\,\mathrm{vec}(\boldsymbol{\Sigma}_x^{-1})'\mathbf{D}\right)(\mathrm{d}\,v(\boldsymbol{\Sigma}_x)) \\
&+ \left(\sum_{i=1}^{n}(\widehat{g}_i)^{(t)}\mathbf{x}_{0i}'\boldsymbol{\Sigma}_x^{-1}\right)(\mathrm{d}\,\boldsymbol{\mu}_x) \\
&+ \left(\frac{n}{2}\log\left(\frac{\nu}{2}\right) - \frac{n}{2}\psi\left(\frac{\nu}{2}\right) + \frac{n}{2} + \frac{1}{2}\sum_{i=1}^{n}(\widehat{\log(g_i)})^{(t)} - \frac{1}{2}\sum_{i=1}^{n}(\widehat{g}_i)^{(t)}\right)(\mathrm{d}\,\nu),
\end{aligned}
$$

onde $(\widehat{g_i\eta_{0i}})^{(t)} = (\widehat{g_i\eta_i})^{(t)} - \mu(\mathbf{x}_i)(\widehat{g}_i)^{(t)}$, $i = 1, \ldots, n$.

O primeiro diferencial de $Q(\boldsymbol{\theta} \mid \widehat{\boldsymbol{\theta}}^{(t)})$ resulta em uma combinação linear de outros diferenciais. Para encontrar o valor de γ_1 que maximiza $Q(\boldsymbol{\theta} \mid \widehat{\boldsymbol{\theta}}^{(t)})$ condicionalmente nos

dados observados e em $\widehat{\boldsymbol{\theta}}^{(t)}$ deve-se igualar a zero o termo que multiplica o diferencial de γ_1, $(\mathrm{d}\,\gamma_1)$, considerando que $\boldsymbol{\theta}_{-\gamma_1}$ seja conhecido, e resolver tal equação em função de γ_1. Para os outros parâmetros repete-se o mesmo procedimento.

Apêndice D

Detalhes sobre o cálculo da matriz de informação observada para os parâmetros do modelo t-Tobit com erros nas covariáveis

Neste apêndice seguem alguns detalhes sobre a obtenção da matriz de informação observada para os parâmetros do modelo proposto. Inicialmente, deve-se obter o segundo diferencial da função de log-verossimilhança do modelo proposto.

De forma geral, se $\mathbf{x} \sim t_k(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$, então o primeiro diferencial de $\log t_k(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$ é dado por

$$\begin{aligned}(\mathrm{d}\log t_k(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)) &= -\frac{1}{2}\operatorname{tr}(\boldsymbol{\Sigma}^{-1}(\mathrm{d}\,\boldsymbol{\Sigma})) + \left[\frac{1}{2}\psi\left(\frac{\nu+k}{2}\right) - \frac{1}{2}\psi\left(\frac{\nu}{2}\right) + \frac{1}{2}\log\nu + \frac{1}{2}\right. \\ &\quad \left. -\frac{1}{2}\log(\nu + q(\mathbf{x})) - \frac{1}{2}m_{(\mathbf{x},k,\nu)}\right](\mathrm{d}\,\nu) \\ &\quad -\frac{1}{2}m_{(\mathbf{x},k,\nu)}(\mathrm{d}\,q(\mathbf{x})),\end{aligned}$$

onde $\operatorname{tr}(\mathbf{A})$ denota o traço da matriz $\mathbf{A}$,

$$(\mathrm{d}\,q(\mathbf{x})) = 2(\mathrm{d}\,\mathbf{x}_0)'\boldsymbol{\Sigma}^{-1}\mathbf{x}_0 - (\mathrm{d}\,v(\boldsymbol{\Sigma}))'\mathbf{D}'(\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{\Sigma}^{-1})\operatorname{vec}(\mathbf{S}_x), \tag{D.1}$$

$\mathbf{S}_x = \mathbf{x}_0\mathbf{x}_0'$ e, se $\mathbf{A}$ é uma matriz simétrica de dimensão $k \times k$, então $\mathbf{D}$ é a matriz de duplicação com dimensão $k^2 \times k(k+1)/2$ tal que $\operatorname{vec}(\mathbf{A}) = \mathbf{D}v(\mathbf{A})$. O segundo diferencial de $\log t_k(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$ é dado por

$$(\mathrm{d}^2\log t_k(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)) = \frac{1}{2}(\mathrm{d}\,v(\boldsymbol{\Sigma}))'\mathbf{D}'(\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{\Sigma}^{-1})\mathbf{D}(\mathrm{d}\,v(\boldsymbol{\Sigma})) - \frac{1}{2}\operatorname{vec}(\boldsymbol{\Sigma}^{-1})'\mathbf{D}(\mathrm{d}^2\,v(\boldsymbol{\Sigma}))$$

$$
\begin{aligned}
&+\left[\frac{1}{4}\psi'\left(\frac{\nu+k}{2}\right)-\frac{1}{4}\psi'\left(\frac{\nu}{2}\right)+\frac{1}{2\nu}-h_{(\nu,k)}m_{(\mathbf{x},k,\nu)}\right.\\
&\left.+\frac{1}{2}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}\right](\mathrm{d}\,\nu)^2+\left[\frac{1}{2}\psi\left(\frac{\nu+k}{2}\right)-\frac{1}{2}\psi\left(\frac{\nu}{2}\right)\right.\\
&\left.+\frac{1}{2}\log\nu+\frac{1}{2}-\frac{1}{2}\log(\nu+q(\mathbf{x}))-\frac{1}{2}m_{(\mathbf{x},k,\nu)}\right](\mathrm{d}^2\,\nu) \qquad \text{(D.2)}\\
&+\left[h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}-h_{(\nu,k)}m_{(\mathbf{x},k,\nu)}\right](\mathrm{d}\,q(\mathbf{x}))(\mathrm{d}\,\nu)\\
&+\frac{1}{2}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}(\mathrm{d}\,q(\mathbf{x}))^2-\frac{1}{2}m_{(\mathbf{x},k,\nu)}(\mathrm{d}^2\,q(\mathbf{x})),
\end{aligned}
$$

onde $\psi'(x)=\frac{\mathrm{d}}{\mathrm{d}x}\psi(x)$ é a função trigama e

$$
\begin{aligned}
(\mathrm{d}^2\,q(\mathbf{x})) &= -4(\mathrm{d}\,\mathbf{x}_0)'(\mathbf{x}_0'\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{\Sigma}^{-1})\mathbf{D}(\mathrm{d}\,v(\boldsymbol{\Sigma}))+2(\mathrm{d}^2\,\mathbf{x}_0)'\boldsymbol{\Sigma}^{-1}\mathbf{x}_0+2(\mathrm{d}\,\mathbf{x}_0)'\boldsymbol{\Sigma}^{-1}(\mathrm{d}\,\mathbf{x}_0)\\
&\quad +2(\mathrm{d}\,v(\boldsymbol{\Sigma}))'\mathbf{D}'(\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{\Sigma}^{-1}\mathbf{S}_x\boldsymbol{\Sigma}^{-1})\mathbf{D}(\mathrm{d}\,v(\boldsymbol{\Sigma}))-(\mathrm{d}^2\,v(\boldsymbol{\Sigma}))'\mathbf{D}'(\boldsymbol{\Sigma}^{-1}\otimes\boldsymbol{\Sigma}^{-1})\,\mathrm{vec}(\mathbf{S}_x).
\end{aligned}
$$

Portanto, a partir de (D.2) obtém-se o segundo diferencial de $\log t_k(\mathbf{x};\boldsymbol{\mu}_x,\boldsymbol{\Sigma}_x,\nu)$ e de $\log t_1\left(y;\mu(\mathbf{x}),\left(\frac{\nu+q(\mathbf{x})}{\nu+k}\right)\sigma_w,\nu+k\right)$ fazendo as devidas substituições.

Se $y\sim t_1(\mu,\sigma,\nu)$ então o primeiro diferencial de $\log T_1(a;\mu,\sigma,\nu)$, onde $a\in\mathbb{R}$, pode ser obtido por

$$
(\mathrm{d}\log T_1(a;\mu,\sigma,\nu))=\frac{1}{T_1(a;\mu,\sigma,\nu)}(\mathrm{d}\,T_1(a;\mu,\sigma,\nu)),
$$

onde

$$
\begin{aligned}
(\mathrm{d}\,T_1(a;\mu,\sigma,\nu)) &= -\frac{1}{2\sigma}T_1(a;\mu,\sigma,\nu)(\mathrm{d}\,\sigma)+\left[\frac{1}{2}\psi\left(\frac{\nu+1}{2}\right)-\frac{1}{2}\psi\left(\frac{\nu}{2}\right)+\frac{1}{2}\log\nu+\frac{1}{2}\right.\\
&\quad \left.-\frac{1}{2}E(\log(\nu+q(y))\mid y\le a)-\frac{1}{2}E\left(m_{(y,1,\nu)}\mid y\le a\right)\right]T_1(a;\mu,\sigma,\nu)(\mathrm{d}\,\nu)\\
&\quad -\frac{1}{2}E\left((\mathrm{d}\,q(y))m_{(y,1,\nu)}\mid y\le a\right)T_1(a;\mu,\sigma,\nu).
\end{aligned}
$$

Logo,

$$
\begin{aligned}
(\mathrm{d}\log T_1(a;\mu,\sigma,\nu)) &= -\frac{1}{2\sigma}(\mathrm{d}\,\sigma)+\left(\frac{1}{2}\psi\left(\frac{\nu+1}{2}\right)-\frac{1}{2}\psi\left(\frac{\nu}{2}\right)+\frac{1}{2}\log\nu+\frac{1}{2}\right.\\
&\quad \left.-\frac{1}{2}E(\log(\nu+q(y))\mid y\le a)-\frac{1}{2}E\left(m_{(y,1,\nu)}\mid y\le a\right)\right)(\mathrm{d}\,\nu)\\
&\quad -\frac{1}{2}E\left((\mathrm{d}\,q(y))m_{(y,1,\nu)}\mid y\le a\right),
\end{aligned}
$$

onde $(\mathrm{d}\,q(y))$ é obtido a partir de (D.1). Portanto, o segundo diferencial de $\log T_1(a;\mu,\sigma,\nu)$ é dado por

$$
\begin{aligned}
(\mathrm{d}^2 \log T_1(a;\mu,\sigma,\nu)) &= \frac{1}{2\sigma^2}(\mathrm{d}\,\sigma)^2 - \frac{1}{2\sigma}(\mathrm{d}^2\,\sigma) + \left(\frac{1}{4}\psi'\left(\frac{\nu+1}{2}\right) - \frac{1}{4}\psi'\left(\frac{\nu}{2}\right) + \frac{1}{2\nu}\right)(\mathrm{d}\,\nu)^2 \\
&+ \left(\frac{1}{2}\psi\left(\frac{\nu+1}{2}\right) - \frac{1}{2}\psi\left(\frac{\nu}{2}\right) + \frac{1}{2}\log\nu + \frac{1}{2}\right. \\
&\left. -\frac{1}{2}E(\log(\nu+q(y)) \mid y \le a) - \frac{1}{2}E\left(m_{(y,1,\nu)} \mid y \le a\right)\right)(\mathrm{d}^2\,\nu) \\
&-\frac{1}{2}\left(\mathrm{d}\,E\left(\log(\nu+q(y)) \mid y \le a\right)\right)(\mathrm{d}\,\nu) \\
&-\frac{1}{2}\left(\mathrm{d}\,E\left(m_{(y,1,\nu)} \mid y \le a\right)\right)(\mathrm{d}\,\nu) \\
&-\frac{1}{2}\left(\mathrm{d}\,E\left((\mathrm{d}\,q(y))m_{(y,1,\nu)} \mid y \le a\right)\right), \qquad \text{(D.3)}
\end{aligned}
$$

onde

$$
\begin{aligned}
\left(\mathrm{d}\,E\left(\log(\nu+q(y)) \mid y \le a\right)\right) &= -\frac{1}{T_1(a;\mu,\sigma,\nu)}E\left(\log(\nu+q(y)) \mid y \le a\right)(\mathrm{d}\,T_1(a;\mu,\sigma,\nu)) \\
&+ \left(\frac{1}{\nu+1}\right)E\left((\mathrm{d}\,q(y))m_{(y,1,\nu)} \mid y \le a\right) \\
&-\frac{1}{2}E\left((\mathrm{d}\,q(y))m_{(y,1,\nu)}\log(\nu+q(y)) \mid y \le a\right) \\
&-\frac{1}{2\sigma}E\left(\log(\nu+q(y)) \mid y \le a\right)(\mathrm{d}\,\sigma) \\
&+ \left[\left(\frac{1}{\nu+1}\right)E\left(m_{(y,1,\nu)} \mid y \le a\right)\right. \\
&+\frac{1}{2}\psi\left(\frac{\nu+1}{2}\right)E\left(\log(\nu+q(y)) \mid y \le a\right) \\
&-\frac{1}{2}\psi\left(\frac{\nu}{2}\right)E\left(\log(\nu+q(y)) \mid y \le a\right) \\
&+\frac{1}{2}\log\nu E\left(\log(\nu+q(y)) \mid y \le a\right) \\
&+\frac{1}{2}E\left(\log(\nu+q(y)) \mid y \le a\right) \\
&-\frac{1}{2}E\left(\left(\log(\nu+q(y))\right)^2 \mid y \le a\right) \\
&\left.-\frac{1}{2}E\left(m_{(y,1,\nu)}\log(\nu+q(y)) \mid y \le a\right)\right](\mathrm{d}\,\nu),
\end{aligned}
$$

$$
\begin{aligned}
\left(\mathrm{d}\,E\left(m_{(y,1,\nu)} \mid y \le a\right)\right) &= -\frac{1}{T_1(a;\mu,\sigma,\nu)}E\left(m_{(y,1,\nu)} \mid y \le a\right)(\mathrm{d}\,T_1(a;\mu,\sigma,\nu)) \\
&-\frac{1}{2}\left(\frac{\nu+3}{\nu+1}\right)E\left((\mathrm{d}\,q(y))m^2_{(y,1,\nu)} \mid y \le a\right) \\
&-\frac{1}{2\sigma}E\left(m_{(y,1,\nu)} \mid y \le a\right)(\mathrm{d}\,\sigma)
\end{aligned}
$$

$$
\begin{aligned}
&+\left[\left(\frac{1}{\nu+1}\right)E\left(m_{(y,1,\nu)} \mid y \le a\right)\right.\\
&-\frac{1}{2}\left(\frac{\nu+3}{\nu+1}\right)E\left(m^2_{(y,1,\nu)} \mid y \le a\right)\\
&+\frac{1}{2}\psi\left(\frac{\nu+1}{2}\right)E\left(m_{(y,1,\nu)} \mid y \le a\right)\\
&-\frac{1}{2}\psi\left(\frac{\nu}{2}\right)E\left(m_{(y,1,\nu)} \mid y \le a\right)\\
&+\frac{1}{2}\log\nu E\left(m_{(y,1,\nu)} \mid y \le a\right)\\
&+\frac{1}{2}E\left(m_{(y,1,\nu)} \mid y \le a\right)\\
&\left.-\frac{1}{2}E\left(\log(\nu+q(y))m_{(y,1,\nu)} \mid y \le a\right)\right](\mathrm{d}\,\nu) \quad \text{e}
\end{aligned}
$$

$$
\begin{aligned}
\left(\mathrm{d}\,E\left((\mathrm{d}\,q(y))m_{(y,1,\nu)} \mid y \le a\right)\right) \;=\; & -\frac{1}{T_1(a;\mu,\sigma,\nu)}E\left((\mathrm{d}\,q(y))m_{(y,1,\nu)} \mid y \le a\right)(\mathrm{d}\,T_1(a;\mu,\sigma,\nu))\\
&+E\left((\mathrm{d}^2\,q(y))m_{(y,1,\nu)} \mid y \le a\right)\\
&-\frac{1}{2}\left(\frac{\nu+3}{\nu+1}\right)E\left((\mathrm{d}\,q(y))^2m^2_{(y,1,\nu)} \mid y \le a\right)\\
&-\frac{1}{2\sigma}E\left((\mathrm{d}\,q(y))m_{(y,1,\nu)} \mid y \le a\right)(\mathrm{d}\,\sigma)\\
&+\left[\left(\frac{\nu+3}{\nu+1}\right)E\left((\mathrm{d}\,q(y))m_{(y,1,\nu)} \mid y \le a\right)\right.\\
&-\left(\frac{\nu+3}{\nu+1}\right)E\left((\mathrm{d}\,q(y))m^2_{(y,1,\nu)} \mid y \le a\right)\\
&+\frac{1}{2}\psi\left(\frac{\nu+1}{2}\right)E\left((\mathrm{d}\,q(y))m_{(y,1,\nu)} \mid y \le a\right)\\
&-\frac{1}{2}\psi\left(\frac{\nu}{2}\right)E\left((\mathrm{d}\,q(y))m_{(y,1,\nu)} \mid y \le a\right)\\
&+\frac{1}{2}\log\nu E\left((\mathrm{d}\,q(y))m_{(y,1,\nu)} \mid y \le a\right)\\
&\left.-\frac{1}{2}E\left((\mathrm{d}\,q(y))\log(\nu+q(y))m_{(y,1,\nu)} \mid y \le a\right)\right](\mathrm{d}\,\nu)
\end{aligned}
$$

Logo, $\mathrm{d}^2\log T_1\left(0;\mu(\mathbf{x}),\left(\frac{\nu+q(\mathbf{x})}{\nu+k}\right)\sigma_w,\nu+k\right)$ pode ser obtido de (D.3).

A matriz de informação observada de $(y_0,\sigma_w,\mathbf{x}_0,v(\boldsymbol{\Sigma}_x),\nu)$ é dada por

$$
\boldsymbol{I}^o((y_0,\sigma_w,\mathbf{x}_0,v(\boldsymbol{\Sigma}_x),\nu)) = \boldsymbol{I}^x + d\boldsymbol{I}^y + (1-d)\boldsymbol{I}^0,
$$

onde a matriz $\boldsymbol{I}^x$ é dada por

$$\boldsymbol{I}^x = -\begin{pmatrix} 0 & 0 & \mathbf{0}' & \mathbf{0}' & 0 \\ 0 & 0 & \mathbf{0}' & \mathbf{0}' & 0 \\ \mathbf{0} & \mathbf{0} & \boldsymbol{I}^x_{\mathbf{x}_0\mathbf{x}_0} & \boldsymbol{I}^x_{\mathbf{x}_0 v(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^x_{\mathbf{x}_0\nu} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{I}^{x'}_{\mathbf{x}_0 v(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^x_{v(\boldsymbol{\Sigma}_x)v(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^x_{v(\boldsymbol{\Sigma}_x)\nu} \\ 0 & 0 & \boldsymbol{I}^{x'}_{\mathbf{x}_0\nu} & \boldsymbol{I}^{x'}_{v(\boldsymbol{\Sigma}_x)\nu} & I^x_{\nu\nu} \end{pmatrix},$$

em que

$$\boldsymbol{I}^x_{\mathbf{x}_0\mathbf{x}_0} = 2h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1} - m_{(\mathbf{x},k,\nu)}\boldsymbol{\Sigma}_x^{-1},$$

$$\begin{aligned}\boldsymbol{I}^x_{\mathbf{x}_0 v(\boldsymbol{\Sigma}_x)} &= -2h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0\,\mathrm{vec}(\mathbf{S}_x)'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\mathbf{D} \\ &\quad +2m_{(\mathbf{x},k,\nu)}(\mathbf{x}_0'\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\mathbf{D},\end{aligned}$$

$$\boldsymbol{I}^x_{\mathbf{x}_0\nu} = -2h_{(\nu,k)}m_{(\mathbf{x},k,\nu)}\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0 + 2h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0,$$

$$\begin{aligned}\boldsymbol{I}^x_{v(\boldsymbol{\Sigma}_x)v(\boldsymbol{\Sigma}_x)} &= \frac{1}{2}\mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\mathbf{D} + \frac{1}{2}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}\mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1})\mathbf{D} \\ &\quad -m_{(\mathbf{x},k,\nu)}\mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1})\mathbf{D},\end{aligned}$$

$$\begin{aligned}\boldsymbol{I}^x_{v(\boldsymbol{\Sigma}_x)\nu} &= h_{(\nu,k)}m_{(\mathbf{x},k,\nu)}\mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\,\mathrm{vec}(\mathbf{S}_x) \\ &\quad -h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}\mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\,\mathrm{vec}(\mathbf{S}_x),\end{aligned}$$

$$\boldsymbol{I}^x_{\nu\nu} = \frac{1}{4}\psi'\left(\frac{\nu+k}{2}\right) - \frac{1}{4}\psi'\left(\frac{\nu}{2}\right) + \frac{1}{2\nu} - h_{(\nu,k)}m_{(\mathbf{x},k,\nu)} + \frac{1}{2}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}.$$

A matriz $\boldsymbol{I}^y$ é

$$\boldsymbol{I}^y = -\begin{pmatrix} I^y_{y_0y_0} & I^y_{y_0\sigma_w} & \boldsymbol{I}^{y'}_{y_0\mathbf{x}_0} & \boldsymbol{I}^{y'}_{y_0 v(\boldsymbol{\Sigma}_x)} & I^y_{y_0\nu} \\ I^y_{y_0\sigma_w} & I^y_{\sigma_w\sigma_w} & \boldsymbol{I}^{y'}_{\sigma_w\mathbf{x}_0} & \boldsymbol{I}^{y'}_{\sigma_w v(\boldsymbol{\Sigma}_x)} & I^y_{\sigma_w\nu} \\ \boldsymbol{I}^y_{y_0\mathbf{x}_0} & \boldsymbol{I}^y_{\sigma_w\mathbf{x}_0} & \boldsymbol{I}^y_{\mathbf{x}_0\mathbf{x}_0} & \boldsymbol{I}^y_{\mathbf{x}_0 v(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^y_{\mathbf{x}_0\nu} \\ \boldsymbol{I}^y_{y_0 v(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^y_{\sigma_w v(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^{y'}_{\mathbf{x}_0 v(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^y_{v(\boldsymbol{\Sigma}_x)v(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^y_{v(\boldsymbol{\Sigma}_x)\nu} \\ I^y_{y_0\nu} & I^y_{\sigma_w\nu} & \boldsymbol{I}^{y'}_{\mathbf{x}_0\nu} & \boldsymbol{I}^{y'}_{v(\boldsymbol{\Sigma}_x)\nu} & I^y_{\nu\nu} \end{pmatrix}$$

e suas entradas são dadas por

$$I^y_{y_0y_0} = \frac{2}{\sigma_w^2}h_{(\nu+k,1)}m^2_{(y,1,\nu+k)}m^2_{(\mathbf{x},k,\nu)}y_0^2 - \frac{1}{\sigma_w}m_{(y,1,\nu+k)}m_{(\mathbf{x},k,\nu)},$$

$$
\begin{aligned}
I^{y}_{y_0\sigma_w} &= -\frac{2}{\sigma_w^3} h_{(\nu+k,1)} m^2_{(y,1,\nu+k)} m^2_{(\mathbf{x},k,\nu)} y_0^3 \\
&\quad + \frac{2}{\sigma_w^2} m_{(y,1,\nu+k)} m_{(\mathbf{x},k,\nu)} y_0,
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{I}^{y'}_{y_0\mathbf{x}_0} &= -\frac{4}{\sigma_w^2} h_{(\nu+k,1)} m^2_{(y,1,\nu+k)} h_{(\nu,k)} m^3_{(\mathbf{x},k,\nu)} y_0^3 \mathbf{x}_0' \boldsymbol{\Sigma}_x^{-1} \\
&\quad + \frac{4}{\sigma_w} m_{(y,1,\nu+k)} h_{(\nu,k)} m^2_{(\mathbf{x},k,\nu)} y_0 \mathbf{x}_0' \boldsymbol{\Sigma}_x^{-1},
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{I}^{y'}_{y_0 v(\boldsymbol{\Sigma}_x)} &= \frac{2}{\sigma_w^2} h_{(\nu+k,1)} m^2_{(y,1,\nu+k)} h_{(\nu,k)} m^3_{(\mathbf{x},k,\nu)} y_0^3 \operatorname{vec}(\mathbf{S}_x)' (\boldsymbol{\Sigma}_x^{-1} \otimes \boldsymbol{\Sigma}_x^{-1}) \mathbf{D} \\
&\quad - \frac{2}{\sigma_w} m_{(y,1,\nu+k)} h_{(\nu,k)} m^2_{(\mathbf{x},k,\nu)} y_0 \operatorname{vec}(\mathbf{S}_x)' (\boldsymbol{\Sigma}_x^{-1} \otimes \boldsymbol{\Sigma}_x^{-1}) \mathbf{D},
\end{aligned}
$$

$$
\begin{aligned}
I^{y}_{y_0\nu} &= -\frac{2}{\sigma_w} h_{(\nu+k,1)} m_{(y,1,\nu+k)} m_{(\mathbf{x},k,\nu)} y_0 \\
&\quad + \frac{2}{\sigma_w} h_{(\nu+k,1)} m^2_{(y,1,\nu+k)} m_{(\mathbf{x},k,\nu)} y_0 \\
&\quad - \frac{2}{\sigma_w^2} h_{(\nu+k,1)} m^2_{(y,1,\nu+k)} h_{(\nu,k)} m^3_{(\mathbf{x},k,\nu)} y_0^3 \\
&\quad + \frac{2}{\sigma_w^2} h_{(\nu+k,1)} m^2_{(y,1,\nu+k)} h_{(\nu,k)} m^2_{(\mathbf{x},k,\nu)} y_0^3 \\
&\quad + \frac{2}{\sigma_w} m_{(y,1,\nu+k)} h_{(\nu,k)} m^2_{(\mathbf{x},k,\nu)} y_0 \\
&\quad - \frac{2}{\sigma_w} m_{(y,1,\nu+k)} h_{(\nu,k)} m_{(\mathbf{x},k,\nu)} y_0,
\end{aligned}
$$

$$
\begin{aligned}
I^{y}_{\sigma_w\sigma_w} &= \frac{1}{2\sigma_w^2} + \frac{1}{2\sigma_w^4} h_{(\nu+k,1)} m^2_{(y,1,\nu+k)} m^2_{(\mathbf{x},k,\nu)} y_0^4 \\
&\quad - \frac{1}{2\sigma_w^4} m_{(y,1,\nu+k)} m^2_{(\mathbf{x},k,\nu)} y_0^4 \\
&\quad - \frac{1}{\sigma_w^3} h_{(\nu+k,1)} m_{(\mathbf{x},k,\nu)},
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{I}^{y'}_{\sigma_w\mathbf{x}_0} &= \frac{2}{\sigma_w^3} h_{(\nu+k,1)} m^2_{(y,1,\nu+k)} h_{(\nu,k)} m^2_{(\mathbf{x},k,\nu)} y_0^4 \mathbf{x}_0' \boldsymbol{\Sigma}_x^{-1} \\
&\quad - \frac{2}{\sigma_w^2} m_{(y,1,\nu+k)} h_{(\nu,k)} m^2_{(\mathbf{x},k,\nu)} y_0^2 \mathbf{x}_0' \boldsymbol{\Sigma}_x^{-1},
\end{aligned}
$$

$$
\boldsymbol{I}^{y'}_{\sigma_w v(\boldsymbol{\Sigma}_x)} = -\frac{1}{\sigma_w^3} h_{(\nu+k,1)} m^2_{(y,1,\nu+k)} h_{(\nu,k)} m^2_{(\mathbf{x},k,\nu)} y_0^4 \operatorname{vec}(\mathbf{S}_x)' (\boldsymbol{\Sigma}_x^{-1} \otimes \boldsymbol{\Sigma}_x^{-1}) \mathbf{D}
$$

$$+\frac{1}{\sigma_w^2}m_{(y,1,\nu+k)}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}y_0^2\,\mathrm{vec}(\mathbf{S}_x)'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\mathbf{D},$$

$$\begin{aligned}
I^y_{\sigma_w\nu} &= \frac{1}{\sigma_w^2}h_{(\nu+k,1)}m_{(y,1,\nu+k)}m_{(\mathbf{x},k,\nu)}y_0^2 \\
&\quad -\frac{1}{\sigma_w^2}h_{(\nu+k,1)}m^2_{(y,1,\nu+k)}m_{(\mathbf{x},k,\nu)}y_0^2 \\
&\quad +\frac{1}{\sigma_w^3}h_{(\nu+k,1)}m^2_{(y,1,\nu+k)}h_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}y_0^4 \\
&\quad -\frac{1}{\sigma_w^3}h_{(\nu+k,1)}m^2_{(y,1,\nu+k)}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}y_0^4 \\
&\quad -\frac{1}{\sigma_w^2}m_{(y,1,\nu+k)}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}y_0^2 \\
&\quad +\frac{1}{\sigma_w^2}m_{(y,1,\nu+k)}h_{(\nu,k)}m_{(\mathbf{x},k,\nu)}y_0^2,
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{I}^y_{\mathbf{x}_0\mathbf{x}_0} &= 2h^2_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1} - h_{(\nu,k)}m_{(\mathbf{x},k,\nu)}\boldsymbol{\Sigma}_x^{-1} \\
&\quad +\frac{2}{\sigma_w^2}h_{(\nu+k,1)}m^2_{(y,1,\nu+k)}h^2_{(\nu,k)}m^4_{(\mathbf{x},k,\nu)}y_0^4\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1} \\
&\quad -\frac{4}{\sigma_w}m_{(y,1,\nu+k)}h^2_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}y_0^2\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1} \\
&\quad +\frac{1}{\sigma_w}m_{(y,1,\nu+k)}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}y_0^2\boldsymbol{\Sigma}_x^{-1},
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{I}^y_{\mathbf{x}_0 v(\boldsymbol{\Sigma}_x)} &= -2h^2_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0\,\mathrm{vec}(\mathbf{S}_x)'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\mathbf{D} \\
&\quad +2h_{(\nu,k)}m_{(\mathbf{x},k,\nu)}(\mathbf{x}_0'\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\mathbf{D} \\
&\quad -\frac{2}{\sigma_w^2}h_{(\nu+k,1)}m^2_{(y,1,\nu+k)}h^2_{(\nu,k)}m^4_{(\mathbf{x},k,\nu)}y_0^4\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0\,\mathrm{vec}(\mathbf{S}_x)'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\mathbf{D} \\
&\quad +\frac{4}{\sigma_w}m_{(y,1,\nu+k)}h^2_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}y_0^2\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0\,\mathrm{vec}(\mathbf{S}_x)'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\mathbf{D} \\
&\quad -\frac{2}{\sigma_w}m_{(y,1,\nu+k)}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}y_0^2(\mathbf{x}_0'\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\mathbf{D},
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{I}^y_{\mathbf{x}_0\nu} &= 2h^2_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0 \\
&\quad +\frac{2}{\sigma_w}h_{(\nu+k,1)}m_{(y,1,\nu+k)}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}y_0^2\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0 \\
&\quad -\frac{2}{\sigma_w}h_{(\nu+k,1)}m^2_{(y,1,\nu+k)}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}y_0^2\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0 \\
&\quad +\frac{2}{\sigma_w^2}h_{(\nu+k,1)}m^2_{(y,1,\nu+k)}h^2_{(\nu,k)}m^4_{(\mathbf{x},k,\nu)}y_0^4\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0 \\
&\quad -\frac{2}{\sigma_w^2}h_{(\nu+k,1)}m^2_{(y,1,\nu+k)}h^2_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}y_0^4\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0
\end{aligned}$$

$$
\begin{aligned}
&-\frac{4}{\sigma_w} m_{(y,1,\nu+k)} h^2_{(\nu,k)} m^3_{(\mathbf{x},k,\nu)} y_0^2 \boldsymbol{\Sigma}_x^{-1} \mathbf{x}_0 \\
&+\frac{2}{\sigma_w} m_{(y,1,\nu+k)} h^2_{(\nu,k)} m^2_{(\mathbf{x},k,\nu)} y_0^2 \boldsymbol{\Sigma}_x^{-1} \mathbf{x}_0,
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{I}^y_{v(\boldsymbol{\Sigma}_x)v(\boldsymbol{\Sigma}_x)} &= \frac{1}{2} h^2_{(\nu,k)} m^2_{(\mathbf{x},k,\nu)} \mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1} \otimes \boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1})\mathbf{D} \\
&\quad - h_{(\nu,k)} m_{(\mathbf{x},k,\nu)} \mathbf{D}'(\boldsymbol{\Sigma}_x^{-1} \otimes \boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1})\mathbf{D} \\
&\quad + \frac{1}{2\sigma_w^2} h_{(\nu+k,1)} m^2_{(y,1,\nu+k)} h^2_{(\nu,k)} m^4_{(\mathbf{x},k,\nu)} y_0^4 \mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1} \otimes \boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1})\mathbf{D} \\
&\quad - \frac{1}{\sigma_w} m_{(y,1,\nu+k)} h^2_{(\nu,k)} m^3_{(\mathbf{x},k,\nu)} y_0^2 \mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1} \otimes \boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1})\mathbf{D} \\
&\quad + \frac{1}{\sigma_w} m_{(y,1,\nu+k)} h_{(\nu,k)} m^2_{(\mathbf{x},k,\nu)} y_0^2 \mathbf{D}'(\boldsymbol{\Sigma}_x^{-1} \otimes \boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1})\mathbf{D},
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{I}^y_{v(\boldsymbol{\Sigma}_x)\nu} &= -h^2_{(\nu,k)} m^2_{(\mathbf{x},k,\nu)} (\boldsymbol{\Sigma}_x^{-1} \otimes \boldsymbol{\Sigma}_x^{-1}) \operatorname{vec}(\mathbf{S}_x) \\
&\quad - \frac{1}{\sigma_w} h_{(\nu+k,1)} m_{(y,1,\nu+k)} h_{(\nu,k)} m^2_{(\mathbf{x},k,\nu)} y_0^2 (\boldsymbol{\Sigma}_x^{-1} \otimes \boldsymbol{\Sigma}_x^{-1}) \operatorname{vec}(\mathbf{S}_x) \\
&\quad + \frac{1}{\sigma_w} h_{(\nu+k,1)} m^2_{(y,1,\nu+k)} h_{(\nu,k)} m^2_{(\mathbf{x},k,\nu)} y_0^2 (\boldsymbol{\Sigma}_x^{-1} \otimes \boldsymbol{\Sigma}_x^{-1}) \operatorname{vec}(\mathbf{S}_x) \\
&\quad - \frac{1}{\sigma_w^2} h_{(\nu+k,1)} m^2_{(y,1,\nu+k)} h^2_{(\nu,k)} m^4_{(\mathbf{x},k,\nu)} y_0^4 (\boldsymbol{\Sigma}_x^{-1} \otimes \boldsymbol{\Sigma}_x^{-1}) \operatorname{vec}(\mathbf{S}_x) \\
&\quad + \frac{1}{\sigma_w^2} h_{(\nu+k,1)} m^2_{(y,1,\nu+k)} h^2_{(\nu,k)} m^3_{(\mathbf{x},k,\nu)} y_0^4 (\boldsymbol{\Sigma}_x^{-1} \otimes \boldsymbol{\Sigma}_x^{-1}) \operatorname{vec}(\mathbf{S}_x) \\
&\quad + \frac{2}{\sigma_w} m_{(y,1,\nu+k)} h^2_{(\nu,k)} m^3_{(\mathbf{x},k,\nu)} y_0^2 (\boldsymbol{\Sigma}_x^{-1} \otimes \boldsymbol{\Sigma}_x^{-1}) \operatorname{vec}(\mathbf{S}_x) \\
&\quad - \frac{1}{\sigma_w} m_{(y,1,\nu+k)} h^2_{(\nu,k)} m^2_{(\mathbf{x},k,\nu)} y_0^2 (\boldsymbol{\Sigma}_x^{-1} \otimes \boldsymbol{\Sigma}_x^{-1}) \operatorname{vec}(\mathbf{S}_x),
\end{aligned}
$$

$$
\begin{aligned}
I^y_{\nu\nu} &= \frac{1}{2\sigma_w^2} h^2_{(\nu,k)} m^2_{(\mathbf{x},k,\nu)} - \frac{1}{\sigma_w^2} h^2_{(\nu,k)} m_{(\mathbf{x},k,\nu)} + \frac{1}{2\sigma_w^2} h^2_{(\nu,k)} \\
&\quad - h^2_{(\nu,k)} + h^2_{(\nu,k)} m_{(\mathbf{x},k,\nu)} + \frac{1}{4}\psi'\left(\frac{\nu+k+1}{2}\right) - \frac{1}{4}\psi'\left(\frac{\nu+k}{2}\right) \\
&\quad + \frac{1}{2} h_{(\nu,k)} - h_{(\nu+k,1)} m_{(y,1,\nu+k)} + \frac{1}{2} h_{(\nu+k,1)} m^2_{(y,1,\nu+k)} \\
&\quad + \frac{1}{\sigma_w} h_{(\nu+k,1)} m_{(y,1,\nu+k)} h_{(\nu,k)} m^2_{(\mathbf{x},k,\nu)} y_0^2 \\
&\quad - \frac{1}{\sigma_w} h_{(\nu+k,1)} m_{(y,1,\nu+k)} h_{(\nu,k)} m_{(\mathbf{x},k,\nu)} y_0^2 \\
&\quad - \frac{1}{\sigma_w} h_{(\nu+k,1)} m^2_{(y,1,\nu+k)} h_{(\nu,k)} m^2_{(\mathbf{x},k,\nu)} y_0^2 \\
&\quad + \frac{1}{\sigma_w} h_{(\nu+k,1)} m^2_{(y,1,\nu+k)} h_{(\nu,k)} m_{(\mathbf{x},k,\nu)} y_0^2
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{1}{2\sigma_w^2}h_{(\nu+k,1)}m^2_{(y,1,\nu+k)}h^2_{(\nu,k)}m^4_{(\mathbf{x},k,\nu)}y_0^4\\
&-\frac{1}{\sigma_w^2}h_{(\nu+k,1)}m^2_{(y,1,\nu+k)}h^2_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}y_0^4\\
&+\frac{1}{2\sigma_w^2}h_{(\nu+k,1)}m^2_{(y,1,\nu+k)}h^2_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}y_0^4\\
&-\frac{1}{\sigma_w}h_{(\nu+k,1)}h^2_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}y_0^2\\
&+\frac{1}{\sigma_w}m_{(y,1,\nu+k)}h^2_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}y_0^2.
\end{aligned}
$$

Finalmente, a matriz $\boldsymbol{I}^0$ é da forma

$$
\boldsymbol{I}^0 = -\begin{pmatrix}
I^0_{y_0y_0} & I^y_{y_0\sigma_w} & \boldsymbol{I}^{0'}_{y_0\mathbf{x}_0} & \boldsymbol{I}^{0'}_{y_0v(\boldsymbol{\Sigma}_x)} & I^0_{y_0\nu}\\
I^0_{y_0\sigma_w} & I^0_{\sigma_w\sigma_w} & \boldsymbol{I}^{0'}_{\sigma_w\mathbf{x}_0} & \boldsymbol{I}^{0'}_{\sigma_wv(\boldsymbol{\Sigma}_x)} & I^0_{\sigma_w\nu}\\
\boldsymbol{I}^0_{y_0\mathbf{x}_0} & \boldsymbol{I}^0_{\sigma_w\mathbf{x}_0} & \boldsymbol{I}^0_{\mathbf{x}_0\mathbf{x}_0} & \boldsymbol{I}^0_{\mathbf{x}_0v(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^0_{\mathbf{x}_0\nu}\\
\boldsymbol{I}^0_{y_0v(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^0_{\sigma_wv(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^{0'}_{\mathbf{x}_0v(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^0_{v(\boldsymbol{\Sigma}_x)v(\boldsymbol{\Sigma}_x)} & \boldsymbol{I}^0_{v(\boldsymbol{\Sigma}_x)\nu}\\
I^0_{y_0\nu} & I^0_{\sigma_w\nu} & \boldsymbol{I}^{0'}_{\mathbf{x}_0\nu} & \boldsymbol{I}^{0'}_{v(\boldsymbol{\Sigma}_x)\nu} & I^0_{\nu\nu}
\end{pmatrix},
$$

em que

$$
\begin{aligned}
I^0_{y_0y_0} &= -\frac{1}{\sigma_w^2}m^2_{(\mathbf{x},k,\nu)}\left(E\left(\eta_0 m_{(\eta,1,\nu+k)} \mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\right)^2\\
&\quad -\frac{1}{\sigma_w}m_{(\mathbf{x},k,\nu)}E\left(m_{(\eta,1,\nu+k)} \mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\\
&\quad +\frac{2}{\sigma_w^2}m^2_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}E\left(\eta_0^2 m^2_{(\eta,1,\nu+k)} \mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\\
&\quad +\frac{1}{\sigma_w^2}m^2_{(\mathbf{x},k,\nu)}E\left(\eta_0^2 m^2_{(\eta,1,\nu+k)} \mid \eta\le 0,\mathbf{b},\mathbf{x}\right),
\end{aligned}
$$

$$
\begin{aligned}
I^0_{y_0\sigma_w} &= \frac{1}{\sigma_w^3}m^2_{(\mathbf{x},k,\nu)}E\left(\eta_0 m_{(\eta,1,\nu+k)} \mid \eta\le 0,\mathbf{b},\mathbf{x}\right)E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\\
&\quad +\frac{2}{\sigma_w^2}m_{(\mathbf{x},k,\nu)}E\left(\eta_0 m_{(\eta,1,\nu+k)} \mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\\
&\quad -\frac{2}{\sigma_w^3}m^2_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}E\left(\eta_0^3 m^2_{(\eta,1,\nu+k)} \mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\\
&\quad -\frac{1}{\sigma_w^3}m^2_{(\mathbf{x},k,\nu)}E\left(\eta_0^3 m^2_{(\eta,1,\nu+k)} \mid \eta\le 0,\mathbf{b},\mathbf{x}\right),
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{I}^{0'}_{y_0\mathbf{x}_0} &= \frac{2}{\sigma_w^2}h_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}\\
&\quad \times E\left(\eta_0 m_{(\eta,1,\nu+k)} \mid \eta\le 0,\mathbf{b},\mathbf{x}\right)E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\mathbf{x}_0'\boldsymbol{\Sigma}_x^{-1}\\
&\quad +\frac{4}{\sigma_w}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}E\left(\eta_0 m_{(\eta,1,\nu+k)} \mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\mathbf{x}_0'\boldsymbol{\Sigma}_x^{-1}
\end{aligned}
$$

$$
\begin{aligned}
&-\frac{4}{\sigma_w^2}h_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}E\left(\eta_0^3 m^2_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\mathbf{x}_0'\boldsymbol{\Sigma}_x^{-1}\\
&-\frac{2}{\sigma_w^2}h_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}E\left(\eta_0^3 m^2_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\mathbf{x}_0'\boldsymbol{\Sigma}_x^{-1},
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{I}^{0'}_{y_0 v(\boldsymbol{\Sigma}_x)} &= -\frac{1}{\sigma_w^2}h_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}E\left(\eta_0 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\\
&\quad\times E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\mathrm{vec}(\mathbf{S}_x)'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\mathbf{D}'\\
&\quad-\frac{2}{\sigma_w}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}\\
&\quad\times E\left(\eta_0 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\mathrm{vec}(\mathbf{S}_x)'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\mathbf{D}'\\
&\quad+\frac{1}{\sigma_w^2}h_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}(\nu+k+3)\\
&\quad\times E\left(\eta_0^3 m^2_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\mathrm{vec}(\mathbf{S}_x)'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\mathbf{D}',
\end{aligned}
$$

$$
\begin{aligned}
I^0_{y_0\nu} &= -\frac{1}{\sigma_w}m_{(\mathbf{x},k,\nu)}E\left(\log(\nu+k+q(\eta)) \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)E\left(\eta_0 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\\
&\quad-\frac{2}{\sigma_w}m_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}E\left(\eta_0 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\\
&\quad+\frac{1}{\sigma_w}m_{(\mathbf{x},k,\nu)}E\left(\log(\nu+k+q(\eta))\eta_0 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\\
&\quad-\frac{1}{\sigma_w}m_{(\mathbf{x},k,\nu)}E\left(m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)E\left(\eta_0 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\\
&\quad+\frac{1}{\sigma_w}m_{(\mathbf{x},k,\nu)}E\left(\eta_0 m^2_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\\
&\quad+\frac{1}{\sigma_w^2}h_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}E\left(\eta_0 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\\
&\quad\times E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\\
&\quad-\frac{1}{\sigma_w^2}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}E\left(\eta_0 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\\
&\quad\times E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\\
&\quad+\frac{2}{\sigma_w}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}E\left(\eta_0 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\\
&\quad-\frac{2}{\sigma_w}h_{(\nu,k)}m_{(\mathbf{x},k,\nu)}E\left(\eta_0 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\\
&\quad-\frac{1}{\sigma_w^2}h_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}(\nu+k+3)E\left(\eta_0^3 m^2_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\\
&\quad+\frac{1}{\sigma_w^2}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}(\nu+k+3)E\left(\eta_0^3 m^2_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right),
\end{aligned}
$$

$$
\boldsymbol{I}^{0'}_{\sigma_w\sigma_w} = \frac{1}{2\sigma_w^2}-\frac{1}{4\sigma_w^4}m^2_{(\mathbf{x},k,\nu)}\left(E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\right)^2
$$

$$
\begin{aligned}
&-\frac{1}{\sigma_w^3} m_{(\mathbf{x},k,\nu)} E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta \leq 0, \mathbf{b}, \mathbf{x}\right) \\
&+\frac{1}{4\sigma_w^4} m_{(\mathbf{x},k,\nu)}^2 h_{(\nu+k,1)}(\nu+k+3) E\left(\eta_0^4 m_{(\eta,1,\nu+k)}^2 \mid \eta \leq 0, \mathbf{b}, \mathbf{x}\right),
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{I}^{0'}_{\sigma_w \mathbf{x}_0} &= -\frac{1}{\sigma_w^3} h_{(\nu,k)} m_{(\mathbf{x},k,\nu)}^3 \left(E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta \leq 0, \mathbf{b}, \mathbf{x}\right)\right)^2 \mathbf{x}_0' \boldsymbol{\Sigma}_x^{-1} \\
&\quad -\frac{2}{\sigma_w^2} h_{(\nu,k)} m_{(\mathbf{x},k,\nu)}^2 E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta \leq 0, \mathbf{b}, \mathbf{x}\right) \mathbf{x}_0' \boldsymbol{\Sigma}_x^{-1} \\
&\quad +\frac{1}{\sigma_w^3} h_{(\nu,k)} m_{(\mathbf{x},k,\nu)}^3 h_{(\nu+k,1)}(\nu+k+3) E\left(\eta_0^4 m_{(\eta,1,\nu+k)}^2 \mid \eta \leq 0, \mathbf{b}, \mathbf{x}\right) \mathbf{x}_0' \boldsymbol{\Sigma}_x^{-1},
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{I}^{0'}_{\sigma_w v(\boldsymbol{\Sigma}_x)} &= \frac{1}{2\sigma_w^3} h_{(\nu,k)} m_{(\mathbf{x},k,\nu)}^3 \\
&\quad \times \left(E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta \leq 0, \mathbf{b}, \mathbf{x}\right)\right)^2 \operatorname{vec}(\mathbf{S}_x)' (\boldsymbol{\Sigma}_x^{-1} \otimes \boldsymbol{\Sigma}_x^{-1}) \mathbf{D}' \\
&\quad +\frac{1}{\sigma_w^2} h_{(\nu,k)} m_{(\mathbf{x},k,\nu)}^2 \\
&\quad \times E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta \leq 0, \mathbf{b}, \mathbf{x}\right) \operatorname{vec}(\mathbf{S}_x)' (\boldsymbol{\Sigma}_x^{-1} \otimes \boldsymbol{\Sigma}_x^{-1}) \mathbf{D}' \\
&\quad -\frac{1}{2\sigma_w^3} h_{(\nu,k)} m_{(\mathbf{x},k,\nu)}^3 h_{(\nu+k,1)}(\nu+k+3) \\
&\quad \times E\left(\eta_0^4 m_{(\eta,1,\nu+k)}^2 \mid \eta \leq 0, \mathbf{b}, \mathbf{x}\right) \operatorname{vec}(\mathbf{S}_x)' (\boldsymbol{\Sigma}_x^{-1} \otimes \boldsymbol{\Sigma}_x^{-1}) \mathbf{D}',
\end{aligned}
$$

$$
\begin{aligned}
I^0_{\sigma_w \nu} &= \frac{1}{2\sigma_w^2} m_{(\mathbf{x},k,\nu)} E\left(\log(\nu+k+q(\eta)) \mid \eta \leq 0, \mathbf{b}, \mathbf{x}\right) E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta \leq 0, \mathbf{b}, \mathbf{x}\right) \\
&\quad +\frac{1}{\sigma_w^2} m_{(\mathbf{x},k,\nu)} h_{(\nu+k,1)} E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta \leq 0, \mathbf{b}, \mathbf{x}\right) \\
&\quad -\frac{1}{2\sigma_w^2} m_{(\mathbf{x},k,\nu)} E\left(\log(\nu+k+q(\eta)) \eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta \leq 0, \mathbf{b}, \mathbf{x}\right) \\
&\quad +\frac{1}{2\sigma_w^2} m_{(\mathbf{x},k,\nu)} E\left(m_{(\eta,1,\nu+k)} \mid \eta \leq 0, \mathbf{b}, \mathbf{x}\right) E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta \leq 0, \mathbf{b}, \mathbf{x}\right) \\
&\quad -\frac{1}{2\sigma_w^2} m_{(\mathbf{x},k,\nu)} h_{(\nu+k,1)}(\nu+k+3) E\left(\eta_0^2 m_{(\eta,1,\nu+k)}^2 \mid \eta \leq 0, \mathbf{b}, \mathbf{x}\right) \\
&\quad -\frac{1}{2\sigma_w^3} h_{(\nu,k)} m_{(\mathbf{x},k,\nu)}^3 \left(E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta \leq 0, \mathbf{b}, \mathbf{x}\right)\right)^2 \\
&\quad +\frac{1}{2\sigma_w^3} h_{(\nu,k)} m_{(\mathbf{x},k,\nu)}^2 \left(E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta \leq 0, \mathbf{b}, \mathbf{x}\right)\right)^2 \\
&\quad -\frac{1}{\sigma_w^2} h_{(\nu,k)} m_{(\mathbf{x},k,\nu)}^2 E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta \leq 0, \mathbf{b}, \mathbf{x}\right) \\
&\quad +\frac{1}{\sigma_w^2} h_{(\nu,k)} m_{(\mathbf{x},k,\nu)} E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta \leq 0, \mathbf{b}, \mathbf{x}\right) \\
&\quad +\frac{1}{2\sigma_w^3} h_{(\nu,k)} m_{(\mathbf{x},k,\nu)}^3 h_{(\nu+k,1)}(\nu+k+3) E\left(\eta_0^4 m_{(\eta,1,\nu+k)}^2 \mid \eta \leq 0, \mathbf{b}, \mathbf{x}\right)
\end{aligned}
$$

$$-\frac{1}{2\sigma_w^3}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}(\nu+k+3)E\left(\eta_0^4 m^2_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right),$$

$$\begin{aligned}
\boldsymbol{I}^0_{\mathbf{x}_0\mathbf{x}_0} &= 2h^2_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1}\\
&\quad -h_{(\nu,k)}m_{(\mathbf{x},k,\nu)}\boldsymbol{\Sigma}_x^{-1}\\
&\quad -\frac{1}{\sigma_w^2}h^2_{(\nu,k)}m^4_{(\mathbf{x},k,\nu)}\left(E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\right)^2\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1}\\
&\quad -\frac{4}{\sigma_w}h^2_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1}\\
&\quad +\frac{1}{\sigma_w}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\boldsymbol{\Sigma}_x^{-1}\\
&\quad +\frac{1}{\sigma_w^2}h^2_{(\nu,k)}m^4_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}(\nu+k+3)E\left(\eta_0^4 m^2_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1},
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{I}^0_{\mathbf{x}_0 v(\boldsymbol{\Sigma}_x)} &= -2h^2_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0\,\mathrm{vec}(\mathbf{S}_x)'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\mathbf{D}\\
&\quad +2h_{(\nu,k)}m_{(\mathbf{x},k,\nu)}(\mathbf{x}_0'\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\mathbf{D}\\
&\quad +\frac{1}{\sigma_w^2}h^2_{(\nu,k)}m^4_{(\mathbf{x},k,\nu)}\left(E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\right)^2\\
&\quad \times\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0\,\mathrm{vec}(\mathbf{S}_x)'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\mathbf{D}\\
&\quad +\frac{4}{\sigma_w}h^2_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0\,\mathrm{vec}(\mathbf{S}_x)'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\mathbf{D}\\
&\quad -\frac{2}{\sigma_w}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)(\mathbf{x}_0'\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\mathbf{D}\\
&\quad -\frac{1}{\sigma_w^2}h^2_{(\nu,k)}m^4_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}(\nu+k+3)E\left(\eta_0^4 m^2_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\\
&\quad \times\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0\,\mathrm{vec}(\mathbf{S}_x)'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\mathbf{D},
\end{aligned}$$

$$\begin{aligned}
\boldsymbol{I}^0_{\mathbf{x}_0\nu} &= 2h^2_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0\\
&\quad +\frac{1}{\sigma_w}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}E\left(\log(\nu+k+q(\eta)) \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0\\
&\quad +\frac{2}{\sigma_w}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0\\
&\quad -\frac{1}{\sigma_w}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}E\left(\log(\nu+k+q(\eta))\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0\\
&\quad +\frac{1}{\sigma_w}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}E\left(m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0\\
&\quad -\frac{1}{\sigma_w}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}(\nu+k+3)E\left(\eta_0^2 m^2_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0\\
&\quad -\frac{1}{\sigma_w^2}h^2_{(\nu,k)}m^4_{(\mathbf{x},k,\nu)}\left(E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\right)^2\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0\\
&\quad +\frac{1}{\sigma_w^2}h^2_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}\left(E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\right)^2\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0\\
&\quad -\frac{4}{\sigma_w}h^2_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0
\end{aligned}$$

$$
\begin{aligned}
&+\frac{2}{\sigma_w}h^2_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0\\
&+\frac{1}{\sigma_w^2}h^2_{(\nu,k)}m^4_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}(\nu+k+3)E\left(\eta_0^4 m^2_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0\\
&-\frac{1}{\sigma_w^2}h^2_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}(\nu+k+3)E\left(\eta_0^4 m^2_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_0,
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{I}^0_{v(\boldsymbol{\Sigma}_x)v(\boldsymbol{\Sigma}_x)} \;=\;& \frac{1}{2}h^2_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}\mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1})\mathbf{D}\\
&-h_{(\nu,k)}m_{(\mathbf{x},k,\nu)}\mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1})\mathbf{D}\\
&-\frac{1}{4\sigma_w^2}h^2_{(\nu,k)}m^4_{(\mathbf{x},k,\nu)}\left(E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\right)^2\mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1})\mathbf{D}\\
&-\frac{1}{\sigma_w}h^2_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1})\mathbf{D}\\
&+\frac{1}{\sigma_w}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1})\mathbf{D}\\
&+\frac{1}{4\sigma_w^2}h^2_{(\nu,k)}m^4_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}(\nu+k+3)E\left(\eta_0^4 m^2_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\\
&\times\mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1}\mathbf{S}_x\boldsymbol{\Sigma}_x^{-1})\mathbf{D},
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{I}^0_{v(\boldsymbol{\Sigma}_x)\nu} \;=\;& -h^2_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}\mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\,\mathrm{vec}(\mathbf{S}_x)\\
&-\frac{1}{2\sigma_w}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}E\left(\log(\nu+k+q(\eta)) \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\\
&\times\mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\,\mathrm{vec}(\mathbf{S}_x)\\
&-\frac{1}{\sigma_w}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\\
&\times\mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\,\mathrm{vec}(\mathbf{S}_x)\\
&+\frac{1}{2\sigma_w}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}E\left(\log(\nu+k+q(\eta))\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\,\mathrm{vec}(\mathbf{S}_x)\\
&-\frac{1}{2\sigma_w}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}E\left(m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\\
&\times\mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\,\mathrm{vec}(\mathbf{S}_x)\\
&+\frac{1}{2\sigma_w}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}(\nu+k+3)E\left(\eta_0^2 m^2_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\\
&\times\mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\,\mathrm{vec}(\mathbf{S}_x)\\
&+\frac{1}{2\sigma_w^2}h^2_{(\nu,k)}m^4_{(\mathbf{x},k,\nu)}\left(E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\right)^2\mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\,\mathrm{vec}(\mathbf{S}_x)\\
&-\frac{1}{2\sigma_w^2}h^2_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}\left(E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\right)^2\mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\,\mathrm{vec}(\mathbf{S}_x)\\
&+\frac{2}{\sigma_w}h^2_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\,\mathrm{vec}(\mathbf{S}_x)\\
&-\frac{1}{\sigma_w}h^2_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}E\left(\eta_0^2 m_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\,\mathrm{vec}(\mathbf{S}_x)\\
&-\frac{1}{2\sigma_w^2}h^2_{(\nu,k)}m^4_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}(\nu+k+3)E\left(\eta_0^4 m^2_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\\
&\times\mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\,\mathrm{vec}(\mathbf{S}_x)\\
&+\frac{1}{2\sigma_w^2}h^2_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}(\nu+k+3)E\left(\eta_0^4 m^2_{(\eta,1,\nu+k)} \mid \eta\leq 0,\mathbf{b},\mathbf{x}\right)\\
&\times\mathbf{D}'(\boldsymbol{\Sigma}_x^{-1}\otimes\boldsymbol{\Sigma}_x^{-1})\,\mathrm{vec}(\mathbf{S}_x),
\end{aligned}
$$

$$
I^0_{\nu\nu} \;=\; \frac{1}{2}h^2_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}-\frac{1}{2}h^2_{(\nu,k)}+\frac{1}{4}\psi'\left(\frac{\nu+k+1}{2}\right)-\frac{1}{4}\psi'\left(\frac{\nu+k}{2}\right)+\frac{1}{2}h_{(\nu,k)}
$$

$$
\begin{aligned}
&-\frac{1}{4}\left(E\left(\log(\nu+k+q(\eta))\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\right)^2\\
&-\frac{1}{2}E\left(m_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)E\left(\log(\nu+k+q(\eta))\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\\
&+\frac{1}{2\sigma_w}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}E\left(\log(\nu+k+q(\eta))\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)E\left(\eta_0^2 m_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\\
&-\frac{1}{2\sigma_w}h_{(\nu,k)}m_{(\mathbf{x},k,\nu)}E\left(\log(\nu+k+q(\eta))\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)E\left(\eta_0^2 m_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\\
&-h_{(\nu+k,1)}E\left(m_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)+\frac{1}{\sigma_w}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}E\left(\eta_0^2 m_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\\
&-\frac{1}{\sigma_w}h_{(\nu,k)}m_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}E\left(\eta_0^2 m_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\\
&+\frac{1}{4}E\left(\left(\log(\nu+k+q(\eta))\right)^2\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)+\frac{1}{2}E\left(\log(\nu+k+q(\eta))m_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\\
&-\frac{1}{2\sigma_w}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}E\left(\log(\nu+k+q(\eta))\eta_0^2 m_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\\
&+\frac{1}{2\sigma_w}h_{(\nu,k)}m_{(\mathbf{x},k,\nu)}E\left(\log(\nu+k+q(\eta))\eta_0^2 m_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\\
&-\frac{1}{4}\left(E\left(m_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\right)^2\\
&+\frac{1}{2\sigma_w}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}E\left(m_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)E\left(\eta_0^2 m_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\\
&-\frac{1}{2\sigma_w}h_{(\nu,k)}m_{(\mathbf{x},k,\nu)}E\left(m_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)E\left(\eta_0^2 m_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\\
&+\frac{1}{4}h_{(\nu+k,1)}(\nu+k+3)E\left(m^2_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\\
&-\frac{1}{2\sigma_w}h_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}(\nu+k+3)E\left(\eta_0^2 m^2_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\\
&+\frac{1}{2\sigma_w}h_{(\nu,k)}m_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}(\nu+k+3)E\left(\eta_0^2 m^2_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\\
&-\frac{1}{4\sigma_w^2}h^2_{(\nu,k)}m^4_{(\mathbf{x},k,\nu)}\left(E\left(\eta_0^2 m_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\right)^2\\
&+\frac{1}{2\sigma_w^2}h^2_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}\left(E\left(\eta_0^2 m_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\right)^2\\
&-\frac{1}{4\sigma_w^2}h^2_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}\left(E\left(\eta_0^2 m_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\right)^2\\
&-\frac{1}{\sigma_w}h^2_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}E\left(\eta_0^2 m_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)+\frac{1}{\sigma_w}h^2_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}E\left(\eta_0^2 m_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\\
&+\frac{1}{4\sigma_w^2}h^2_{(\nu,k)}m^4_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}(\nu+k+3)E\left(\eta_0^4 m^2_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\\
&-\frac{1}{2\sigma_w^2}h^2_{(\nu,k)}m^3_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}(\nu+k+3)E\left(\eta_0^4 m^2_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right)\\
&+\frac{1}{4\sigma_w^2}h^2_{(\nu,k)}m^2_{(\mathbf{x},k,\nu)}h_{(\nu+k,1)}(\nu+k+3)E\left(\eta_0^4 m^2_{(\eta,1,\nu+k)}\mid \eta\le 0,\mathbf{b},\mathbf{x}\right).
\end{aligned}
$$