

Inferência em Modelos Heterocedásticos*

Francisco Cribari-Neto**

Ana Cristina Nunes Soares***

Sumário: 1. Introdução; 2. O modelo e estimadores; 3. Um novo estimador; 4. Testes *bootstrap*; 5. Avaliação numérica; 6. Discussão e conclusões.

Palavras-chave: *bootstrap*; heterocedasticidade; pontos de alavancagem; regressão; testes quase-*t*.

Códigos JEL: C12; C13; C15.

Este artigo analisa o desempenho em amostras finitas do estimador consistente de matrizes de covariâncias proposto por Halbert White. O comportamento deste estimador é estudado tanto sob homocedasticidade quanto sob heterocedasticidade usando métodos de simulação de Monte Carlo. Outros dois estimadores consistentes são também analisados, a saber: o estimador HC3, que constitui uma boa aproximação do estimador jackknife, e o estimador de *bootstrap* ponderado. O método de *bootstrap* é ainda usado para obter valores críticos para testes quase-*t*. É feita também uma análise do desempenho do *bootstrap* duplo, no qual um segundo nível de *bootstrap* é realizado dentro de cada réplica de *bootstrap* de primeiro nível. Por fim, com base no HC3, um novo estimador é proposto para levar em consideração o efeito de pontos de alavancagem sobre a inferência resultante, a partir de testes quase-*t* associados.

This paper focuses on the finite-sample behavior of heteroskedasticity-consistent covariance matrix estimators and associated quasi-*t* tests. The estimator most commonly used is that proposed by Halbert White. Its finite-sample behavior under both homoskedasticity and heteroskedasticity is analyzed using Monte Carlo methods. The paper considers two other consistent estimators, namely: the HC3 estimator, which is an approximation to the jackknife estimator, and the weighted *bootstrap* estimator. Additionally, it evaluates the finite-sample behavior of two *bootstrap* quasi-*t* tests: the test based on a single bootstrapping scheme and the test based on a

* Artigo recebido em ago. 2001 e aprovado em jun. 2002. Os autores agradecem ao CNPq e à Finep o apoio financeiro parcial.

**Departamento de Estatística da Universidade Federal de Pernambuco. E-mail:cribari@ufpe.br.

***Centro de Informática da Universidade Federal de Pernambuco. E-mail: acns@cin.ufpe.br.

double, nested bootstrapping scheme. The latter is very computer-intensive, but proves to work well in small samples. Finally, the paper proposes a new estimator, called HC4, tailored to take into account the effect of leverage points in the design matrix on associated quasi- t tests.

1. Introdução

Regressões envolvendo dados de corte transversal tipicamente apresentam comportamento heterocedástico, ou seja, apresentam variâncias do termo de erro do modelo que não são constantes para todas as observações. Na presença de heterocedasticidade o estimador de mínimos quadrados ordinários (EMQO) dos parâmetros lineares da estrutura de regressão permanece não-viesado (isto é, em média se iguala ao parâmetro verdadeiro) e consistente (ou seja, converge em probabilidade para o parâmetro verdadeiro à medida que o número de observações aumenta). Contudo, o estimador usual da matriz de covariâncias do EMQO dos parâmetros de regressão é viesado e inconsistente quando há heterocedasticidade. Uma prática comum nestes casos é a utilização de estimativas de mínimos quadrados desses parâmetros em conjunção com alguma estimativa de sua matriz de covariâncias que seja consistente tanto sob homocedasticidade quanto sob heterocedasticidade de forma desconhecida. Neste sentido, Wooldridge (2000:249) escreve que nas últimas décadas os econométricos aprenderam a ajustar erros-padrão e estatísticas de teste para que sejam válidos na presença de heterocedasticidade de forma desconhecida, o que é conveniente porque podemos reportar estatísticas que são confiáveis independentemente do tipo de heterocedasticidade presente nos dados.

O estimador consistente da matriz de covariâncias proposto por White (1980) é o mais utilizado em aplicações práticas. De fato, em meados de 2001 o Institute for Scientific Information listava mais de 2600 citações ao artigo onde o estimador de White foi proposto, o que é indicativo de sua ampla repercussão entre economistas, estatísticos e investigadores das mais diversas áreas. Este artigo utiliza métodos de simulação de Monte Carlo para avaliar o desempenho do estimador de White em amostras de tamanho típico, assim como o desempenho de testes quase- t associados. Avalia, ainda, o desempenho de dois outros estimadores consistentes, a saber: o estimador HC3 sugerido por Davidson e MacKinnon (1993) e o estimador de *bootstrap* ponderado proposto por Wu (1986). Esses estimadores são igualmente utilizados para se construir testes quase- t , cujo desempenho é também avaliado. Uma estratégia alternativa é a utilização do método de *bootstrap* para

a obtenção de valores críticos ou valores p para o teste quase- t que utiliza o estimador de White. A eficácia dessa estratégia é implementada a partir do método de *bootstrap* não-paramétrico tradicional, bem como a partir de um esquema de *bootstrap* duplo, que se torna computacionalmente muito mais intensivo.

Os resultados revelam que o estimador de White pode ser muito viesado em amostras finitas, conduzindo a testes quase- t liberais. Mostram ainda, que a existência de pontos de alta alavancagem tem grande influência sobre o desempenho dos estimadores consistentes e testes associados. Para contornar este problema, propomos um novo estimador, construído a partir do estimador HC3, que utiliza ponderações para as diferentes observações que são calculadas a partir de seus respectivos graus de alavancagem. Os resultados numéricos revelam que este estimador conduz a testes quase- t que não são marcadamente liberais e, conseqüentemente, a inferências mais confiáveis. Mostra-se ainda que a realização de inferência via testes de *bootstrap* duplo também é confiável. De fato, os dois enfoques apresentam desempenho similar em amostras de tamanho típico. A vantagem da utilização do estimador aqui proposto e denominado HC4 sobre esquemas de *bootstrap* duplo reside em sua simplicidade prática e computacional.

2. O Modelo e Estimadores

O modelo considerado é o linear de regressão da forma

$$y = X\beta + u$$

onde:

y é um vetor $n \times 1$ de observações da variável dependente;

X é uma matriz fixa de posto completo de dimensão $n \times p$ ($p < n$) contendo observações sobre as variáveis explicativas;

$\beta = (\beta_1, \dots, \beta_p)'$ é um vetor $p \times 1$ de parâmetros desconhecidos;

u é um vetor $n \times 1$ de distúrbios aleatórios (erros) com média zero e matriz de covariâncias $\Omega = \text{diag}(\sigma_1^2, \dots, \sigma_n^2)$.

Quando os erros são homocedásticos, então $\sigma_i^2 = \sigma^2 > 0$, ou seja, $\Omega = \sigma^2 I_n$, onde I_n é a matriz identidade de ordem n . O estimador de mínimos quadrados ordinários de β é dado por $\hat{\beta} = (X'X)^{-1}X'y$, cuja média é β (isto é, ele é não-viesado) e cuja variância é dada por

$$\Psi = (X'X)^{-1}X'\Omega(X'X)^{-1}$$

Sob homocedasticidade, ou seja, $\Omega = \sigma^2 I_n$, esta expressão se simplifica a $\sigma^2 (X'X)^{-1}$, podendo ser facilmente estimada como $\hat{\sigma}^2 (X'X)^{-1}$, onde $\hat{\sigma}^2 = \hat{u}'\hat{u}/(n-p)$. Aqui, $\hat{u} = (I_n - X(X'X)^{-1}X')y = My$ representa o vetor $n \times 1$ de resíduos de mínimos quadrados.

O estimador consistente para a matrix Ψ mais comumente utilizado em aplicações empíricas é o estimador de White (1980):

$$\hat{\Psi} = (X'X)^{-1}X'\hat{\Omega}X(X'X)^{-1}$$

onde $\hat{\Omega} = \text{diag}(\hat{u}_1^2, \dots, \hat{u}_n^2)$. Ou seja, $\hat{\Omega}$ é uma matriz diagonal formada a partir do vetor contendo os quadrados dos resíduos de mínimos quadrados. Este estimador é consistente quando os erros são homocedásticos e quando há heterocedasticidade de forma desconhecida (White, 1980). Entretanto, o estimador de White pode apresentar viés substancial em amostras de tamanho finito (Cribari-Neto, 1999, Cribari-Neto e Zarkos, 2001); (MacKinnon e White, 1985).

Um estimador alternativo que geralmente possui melhor desempenho em pequenas amostras é construído a partir do estimador de White, mas incorporando a ele termos de correção. A idéia é usar

$$\hat{\Omega} = \text{diag} \left\{ \hat{u}_1^2 / (1 - h_1)^2, \dots, \hat{u}_n^2 / (1 - h_n)^2 \right\}$$

onde h_i é o i -ésimo elemento diagonal da “matriz chapéu” $H = X(X'X)^{-1}X'$, $i = 1, \dots, n$. O estimador resultante é conhecido como HC3 e é uma aproximação do estimador jackknife considerado por MacKinnon e White (1985).¹

Uma alternativa computacionalmente intensiva reside na utilização do método *bootstrap*, proposto por Bradley Efron em um influente artigo no *Annals of Statistics* (Efron, 1979). Este método objetiva a obtenção de medidas estatísticas de precisão baseadas em simulações computacionais e geralmente fornece uma aproximação para a estatística de interesse que é mais precisa do que aquela obtida a partir de sua aproximação assintótica de primeira ordem.² Em sua forma mais simples, o algoritmo de *bootstrap* pode ser descrito da seguinte forma:

- a) obtenha uma amostra aleatória u_1^*, \dots, u_n^* de \hat{u} com reposição;
- b) forme uma amostra de *bootstrap* $y^* = X\hat{\beta} + u^*$, onde $u^* = (u_1^*, \dots, u_n^*)'$;
- c) obtenha a estimativa de MQO de β : $\beta^* = (X'X)^{-1}X'y^*$;

¹Ver Davidson e MacKinnon (1993, §16.3).

²Para mais detalhes, ver, entre outros, Davison e Hinkley (1997) e Efron e Tibshirani (1993).

- d) repita os passos anteriores um grande número de vezes (digamos, B);
- e) calcule a variância dos $B + 1$ vetores de estimativas obtidas usando os passos acima (os B vetores obtidos do esquema de *bootstrap* e o vetor inicial).

Contudo, esse esquema de *bootstrap* não leva em consideração o fato que as variâncias das observações são diferentes quando há heterocedasticidade. De fato, estimativas de *bootstrap* obtidas como descrito não são nem consistentes nem assintoticamente não-viesadas quando os dados provêm de um mecanismo gerador heterocedástico (Wu, 1986). Um estimador de *bootstrap* robusto à presença de heterocedasticidade foi proposto por Wu (1986) e pode ser descrito da seguinte forma:

- a1) para cada i , $i = 1, \dots, n$, obtenha aleatoriamente t_i^* de uma distribuição com média zero e variância um.
- b1) Forme a amostra de *bootstrap* (y^*, X) , onde $y_i^* = X_i\hat{\beta} + t_i^*\hat{u}_i/(1 - h_i)$; aqui, X_i denota a i -ésima linha da matriz X .

Os passos c até e permanecem inalterados. Note que a variância (no esquema de *bootstrap*) de $t_i^*\hat{u}_i$ não é constante quando os erros originais não são homocedásticos. Desta forma, este esquema de “*bootstrap* ponderado” leva em consideração a possível não-constância das variâncias dos erros. Note também que o passo b1 foi alterado, uma vez que a proposta original de Wu era a de dividir cada resíduo por $\sqrt{1 - h_i}$, e não por $1 - h_i$.

3. Um Novo Estimador

Vários estudos têm revelado que a introdução de pontos de alavancagem na matriz de regressores X deteriora o desempenho em amostras finitas de estimadores consistentes da matriz de covariâncias do EMQO, conduzindo a testes quase- t associados liberais (Cribari-Neto e Gois, 2002) e (Cribari-Neto e Zarkos, 2001). Uma forma de identificar observações potencialmente influentes é analisar os correspondentes elementos diagonais (h_i) da “matriz-chapéu” $H = X(X'X)^{-1}X'$. Pode ser mostrado que $0 < h_i < 1$ para todo i e $\sum_{i=1}^p h_i = p$, onde p denota o posto da matriz X . Dessa forma, os h_i têm valor médio p/n . Uma regra prática usada por econométricos e estatísticos é que valores de h_i superiores a duas ou três vezes a média (isto é, $2p/n$ e $3p/n$) são tomados como influentes e merecedores de análise mais detalhada (Judge et alii, 1988:893). O estimador HC3 inclui um

termo de correção que considera o grau de alavancagem de cada observação medido pelo correspondente elemento diagonal da matriz H pois usa-se, como visto,

$$\widehat{\Omega} = \text{diag} \left\{ \widehat{u}_1^2 / (1 - h_1)^2, \dots, \widehat{u}_n^2 / (1 - h_n)^2 \right\}$$

Outro estimador adotado em algumas aplicações é o HC2 (MacKinnon e White, 1985), que utiliza

$$\widehat{\Omega} = \text{diag} \left\{ \widehat{u}_1^2 / (1 - h_1), \dots, \widehat{u}_n^2 / (1 - h_n) \right\}$$

Assim, o estimador HC3 incorpora uma correção mais acentuada para atenuar o fato de que os resíduos tendem a flutuar menos que os erros verdadeiros, uma vez que eleva os termos de correção pelos quais os resíduos ao quadrado são divididos por quantidades menores.

O estimador que propomos, que denominamos HC4, utiliza

$$\widehat{\Omega} = \text{diag} \left\{ \widehat{u}_1^2 / (1 - h_1)^{\delta_1}, \dots, \widehat{u}_n^2 / (1 - h_n)^{\delta_n} \right\}$$

onde $\delta_i = n h_i / \sum_{j=1}^n h_j = n h_i / p$. A última igualdade vem do fato que a soma dos h_i é igual ao posto da matriz X , ou seja, igual a p .³ Aqui, o expoente que controla o grau de desconto para a observação i é dado pela razão entre o valor de h_i e a média dos h_i .

Como observado por Chesher e Jewitt (1987:1219), a possibilidade de viés negativo severo no estimador de White ocorre quando há valores elevados de h_i porque os resíduos de mínimos quadrados associados possuem magnitude reduzida em média e o estimador interpreta resíduos pequenos como indicação de variâncias pequenas. A divisão dos resíduos ao quadrado por $(1 - h_i)^2$ no estimador HC3 procura corrigir essa tendência, uma vez que o valor deste termo será menor para observações que possuem h_i elevado. Nossa proposta reside em acentuar a magnitude desse ajustamento elevando o termo de ajuste, $1 - h_i$, à potência δ_i , onde $\delta_i = h_i / \bar{h}$. Aqui, \bar{h} denota o valor médio dos h_i , isto é, $\bar{h} = n^{-1} \sum_{i=1}^n h_i = p/n$. Como $0 < 1 - h_i < 1$ e $\delta_i > 0$, temos que $0 < (1 - h_i)^{\delta_i} < 1$. Assim, os resíduos ao quadrado serão mais “inflacionados” quando as observações correspondentes tiverem maior grau de alavancagem.

³Note que a matriz H é simétrica e idempotente e, assim, seu traço é igual a seu posto, que é, por sua vez, igual a p .

4. Testes Bootstrap

Um procedimento alternativo é construir a estatística de teste quase- t usando o estimador de White e realizar o teste baseado nessa estatística, em conjunção com valores críticos obtidos a partir de um esquema de *bootstrap* em que se calcula a estatística em cada réplica desse esquema e utilizam-se as pseudo-estatísticas de teste para estimar a distribuição nula da estatística quase- t . Uma característica desejável deste procedimento é que a quantidade avaliada no processo de *bootstrap* é assintoticamente pivotal, ou seja, possui uma distribuição limite livre de parâmetros desconhecidos.

O procedimento é o seguinte. Inicialmente, calcule a estatística quase- t , digamos τ . Execute, então, os seguintes passos:

- a) para cada i , $i = 1, \dots, n$, obtenha um número aleatório t_i^* de uma população com média zero e variância unitária.
- b) construa a amostra de *bootstrap* (y^*, X) , onde $y_i^* = X_i \tilde{\beta} + t_i^* \tilde{u}_i / (1 - h_i)$; aqui, $\tilde{\beta}$ e \tilde{u} são o vetor de estimativas restritas e o vetor de resíduos restritos associados da regressão de y em X ;
- c) obtenha a estimativa de MQO de β , $\beta^* = (X'X)^{-1}X'y^*$ e, calcule a estatística quase- t , τ^* ;
- d) repita os passos acima um grande número de vezes (digamos, B);
- e) calcule o quantil de interesse da distribuição empírica das $B + 1$ realizações da estatística quase- t (a realização inicial e as B realizações de *bootstrap*) obtidas usando os passos a até d.
- f) realize o teste utilizando a estatística quase- t calculada inicialmente (τ) junto com o valor crítico de *bootstrap* obtido no passo e.

Note que no teste *bootstrap* não utilizamos valores críticos assintóticos da distribuição normal padrão, mas sim valores críticos estimados em um esquema de reamostragem de *bootstrap*. A regra de decisão pode ser mais convenientemente expressa em termos do valor p do teste e de sua estimativa de *bootstrap*. O valor p aproximado via esquema de *bootstrap*, para um teste bicaudal, é dado por

$$p = \frac{1 + \#\{|\tau_b^*| \geq |\tau|\}}{B + 1}$$

a partir dos resultados $\tau_1^*, \dots, \tau_B^*$ das B réplicas de *bootstrap*. A hipótese nula é rejeitada se este valor p for inferior à probabilidade de erro do tipo I selecionada (tamanho nominal).

É possível obter um valor p de *bootstrap* mais preciso usando um esquema de *bootstrap* duplo, que se torna, contudo, computacionalmente mais intensivo. A idéia básica é realizar um segundo nível de *bootstrap* dentro de cada réplica do esquema principal (Davison e Hinkley, 1997, §4.5.) Sejam $\tau_1^*, \dots, \tau_B^*$ as B realizações de *bootstrap* da estatística de teste. Temos, assim, o seguinte esquema de *bootstrap* duplo, onde C denota o número de réplicas no segundo nível de *bootstrap* e $b = 1, \dots, B$ indexa o primeiro nível:

- a) para cada i , $i = 1, \dots, n$, obtenha um número aleatório t_i^{**} de uma população com média zero e variância unitária;
- b) construa a amostra de *bootstrap* (y^{**}, X) , onde $y_i^{**} = X_i \tilde{\beta}^\dagger + t_i^{**} \tilde{u}_i^\dagger / (1 - h_i)$; aqui, $\tilde{\beta}^\dagger$ e \tilde{u}^\dagger são o vetor de estimativas restritas e o vetor de resíduos restritos associados da regressão de y^* em X ;
- c) obtenha a estimativa de MQO de β , $\beta^{**} = (X'X)^{-1} X' y^{**}$, e calcule a estatística quase- t , τ^{**} ;
- d) calcule p_b^* usando a equação (1), a seguir;
- e) use as realizações dos dois níveis de *bootstrap* para obter um valor p ajustado para realizar o teste (ver a seguir).

Os passos a até d devem ser realizados para cada réplica de *bootstrap* ($b = 1, \dots, B$). O valor p de *bootstrap* ajustado é dado por

$$p_{adj} = \frac{1 + \#\{p_b^* \leq p\}}{B + 1}$$

onde, para cada b ,

$$p_b^* = \frac{1 + \#\{|\tau_{bc}^{**}| \geq |\tau_b^*|\}}{C + 1} \quad (1)$$

Rejeita-se a hipótese nula, novamente testada contra uma hipótese alternativa bicaudal, se $p_{adj} \leq \alpha$, onde α é o tamanho nominal do teste. Note que o número total de réplicas de *bootstrap* passa a ser $B \times C$, implicando, assim, maior custo computacional.

5. Avaliação Numérica

Os resultados numéricos apresentados nesta seção correspondem ao modelo $y_i = \beta_1 + \beta_2 x_i + \sigma_i u_i$, $i = 1, \dots, n$. Os tamanhos amostrais considerados foram $n = 50, 100, 150$. No caso onde $n = 50$, os valores de x_i foram obtidos como números aleatórios de uma distribuição lognormal através da exponenciação de números aleatórios independentes obtidos de uma distribuição normal padrão. Para os demais tamanhos amostrais os valores da covariável foram replicados, ou seja, cada valor de x_i foi replicado duas vezes quando $n = 100$ e três vezes quando $n = 150$. Esta replicação de covariáveis garante que o grau de heterocedasticidade permaneça constante à medida que o número de observações aumenta e é comumente utilizada na literatura (Cribari-Neto, 1999, Cribari-Neto e Zarkos, 2001, MacKinnon e White, 1985). Os erros, u_i , são independente e identicamente distribuídos seguindo a distribuição $\mathcal{N}(0, 1)$. Os dados foram gerados utilizando $\beta_1 = 1$ e $\beta_2 = 0$. (Esta escolha de valores para os parâmetros lineares não afeta significativamente os resultados obtidos.) Quando há homocedasticidade, $\sigma_i = 1$ para todos os valores de i . Por outro lado, a geração de dados heterocedásticos deu-se usando

$$\sigma_i^2 = \exp \{ \gamma x_i + \gamma x_i^2 \}$$

O grau de heterocedasticidade pode ser medido usando $\lambda = (\max \sigma_i^2) / (\min \sigma_i^2)$; ou seja, sob homocedasticidade $\lambda = 1$ e sob heterocedasticidade $\lambda > 1$. As simulações foram realizadas utilizando 5 mil réplicas de Monte Carlo, 999 réplicas de *bootstrap* de primeiro nível e 249 réplicas de *bootstrap* de segundo nível. Cada experimento requer, portanto, um total de aproximadamente 1,25 bilhão de réplicas, indicando, assim, que as simulações realizadas são extremamente intensivas computacionalmente. Todos os experimentos foram programados utilizando a linguagem de programação C (Cribari-Neto, 1999) e compilados com o compilador gcc (Stallman, 1999) no sistema operacional Linux (MacKinnon, 1999).

As tabelas 1 e 2 apresentam, respectivamente, os vieses relativos totais e a raiz quadrada do erro quadrático médio total ($\times \sqrt{5000}$). Os seguintes estimadores foram analisados: de mínimos quadrados ordinários (“MQO”); de White

(‘white’), HC3; o estimador aqui proposto e denominado HC4; o estimador obtido do esquema de *bootstrap* ponderado. Definimos o viés relativo total como a soma dos valores absolutos dos vieses relativos individuais dos elementos diagonais (variâncias) da matriz de covariâncias das estimativas de mínimos quadrados. O viés relativo individual de um estimador, por sua vez, é definido como a média das estimativas menos o valor verdadeiro do parâmetro, sendo esta diferença dividida pelo valor verdadeiro do parâmetro. O viés relativo total fornece, assim, uma medida agregada do viés das estimativas das variâncias. Estes resultados estão na tabela 1. A tabela 2 contém as raízes quadradas das somas dos erros quadráticos médios individuais para os diferentes estimadores (REQM). Essas quantidades medem, assim, tanto o viés quanto a variabilidade dos estimadores, uma vez que o erro quadrático médio de um estimador é dado pela soma de seu viés ao quadrado e de sua variância.

Tabela 1
Vieses relativos totais

n	γ	λ	MQO	White	HC3	HC4	Boot
50	0.00	1.00	0.0042	0.1927	0.2149	1.2686	0.2132
	0.04	4.57	0.6338	0.3059	0.2618	2.0833	0.2597
	0.08	20.86	1.0613	0.4481	0.2841	2.9204	0.2805
	0.12	95.27	1.3205	0.5493	0.2905	3.4955	0.2909
100	0.00	1.00	0.0002	0.0969	0.0982	0.4264	0.0954
	0.04	4.57	0.6243	0.1808	0.0877	0.6528	0.0871
	0.08	20.86	1.0215	0.2358	0.1223	0.9753	0.1141
	0.12	95.27	1.2711	0.2848	0.1328	1.1906	0.1364
150	0.00	1.00	0.0007	0.0586	0.0703	0.2607	0.0691
	0.04	4.57	0.6119	0.1072	0.0736	0.4108	0.0736
	0.08	20.86	1.0111	0.1638	0.0737	0.5750	0.0704
	0.12	95.27	1.2533	0.1957	0.0814	0.7021	0.0779

No que se refere à análise dos vieses dos estimadores de matrizes de covariância (tabela 1), notamos inicialmente que o estimador de mínimos quadrados ordinários é não-viesado sob homocedasticidade, mas fortemente viesado quando esta propriedade não se verifica, o que era esperado. Entre os estimadores consistentes, os que apresentam menor viés são os estimadores HC3 e de *bootstrap* ponderado, ao passo que o estimador HC4 se revela o mais viesado. Considere, por exemplo, o caso onde $n = 100$ e $\gamma = 0.12$ (o que resulta em $\lambda = 95.27$). Nesta situação os vieses relativos totais dos estimadores de White, HC3, HC4 e de *bootstrap* ponderado são, respectivamente, 28.48%, 13.28%, 119.06% e 13.64%. Assim, no que diz respeito ao viés, os estimadores com melhor comportamento em amostras finitas são o HC3 e o de *bootstrap* ponderado.

Tabela 2
REQM totais ($\times\sqrt{5000}$)

n	γ	λ	MQO	White	HC3	HC4	Boot
50	0.00	1.00	0.6754	0.8885	1.1228	2.3936	1.1354
	0.04	4.57	2.0667	2.3119	3.4669	10.0024	3.4750
	0.08	20.86	9.2075	8.8181	13.6044	41.3884	13.6726
	0.12	95.27	41.3048	38.8268	60.2640	184.3878	60.7768
100	0.00	1.00	0.2364	0.3298	0.3714	0.5147	0.3793
	0.04	4.57	0.9212	0.8536	1.0181	1.7352	1.0314
	0.08	20.86	4.3146	3.4525	4.2550	7.5932	4.2391
	0.12	95.27	19.6086	15.4746	19.1685	34.2992	19.7064
150	0.00	1.00	0.1288	0.1797	0.1943	0.2380	0.2014
	0.04	4.57	0.5754	0.4834	0.5490	0.7922	0.5571
	0.08	20.86	2.8005	1.9395	2.2214	3.3051	2.2526
	0.12	95.27	12.7449	8.5950	9.8799	14.7574	10.1502

As raízes quadradas dos erros quadráticos médios totais são apresentadas na tabela 2 ($\times\sqrt{5000}$). Nota-se que entre os estimadores consistentes o que apresenta menor erro quadrático médio total é o de White, sendo o estimador HC4 o que tem desempenho mais pobre mais uma vez. Os estimadores HC3 e de *bootstrap* ponderado novamente apresentam desempenho semelhante.

A tabela 3 apresenta os tamanhos estimados dos testes quase- t associados aos estimadores considerados; as variâncias estimadas a partir desses estimadores são utilizadas no denominador da estatística de teste. O interesse reside em testar a hipótese nula $\mathcal{H}_0 : \beta_j = \beta_j^{(0)}$, $j = 1, \dots, p$, onde $\beta_j^{(0)}$ é uma dada constante, contra uma hipótese alternativa bicaudal. A estatística de teste pode ser escrita como

$$\frac{\widehat{\beta}_j - \beta_j^{(0)}}{\sqrt{\widehat{\text{var}}(\widehat{\beta}_j)}}$$

onde $\widehat{\text{var}}(\widehat{\beta}_j)$ denota a variância estimada de $\widehat{\beta}_j$ obtida a partir de um dos estimadores consistentes considerados.

Sob a hipótese nula, esta estatística de teste possui distribuição-limite $\mathcal{N}(0, 1)$. Dessa forma, o teste é realizado comparando o valor (absoluto) da estatística ao valor crítico (assintótico) do teste, que é dado pelo quantil de nível $(1 - \alpha)/2$ da distribuição normal padrão. No experimento de simulação testa-se $\mathcal{H}_0 : \beta_2 = 0$, ou seja, considera-se o teste de exclusão de uma variável independente irrelevante. Além dos testes quase- t construídos a partir dos diferentes estimadores consistentes (realizados utilizando valores críticos assintóticos), são considerados também dois testes de *bootstrap*, baseados na estatística quase- t que utiliza o estimador de

White em sua construção. O primeiro teste de *bootstrap* baseia-se na construção de pseudo-estatísticas de teste com base em $B = 999$ réplicas de *bootstrap* e em sua utilização para obtenção de uma estimativa da distribuição nula da estatística quase- t . O segundo teste de *bootstrap* se baseia no mesmo princípio, mas utiliza um esquema de *bootstrap* duplo com $B = 999$ réplicas de *bootstrap* de primeiro nível e $C = 249$ réplicas de segundo nível. As taxas de rejeição (expressas como percentagens) estimadas dos diferentes testes correspondentes ao nível nominal de $\alpha = 5\%$ encontram-se listadas na tabela 3.

Tabela 3
Taxas de rejeição de testes quase- t ao nível nominal de $\alpha = 5\%$

n	γ	λ	MQO	White	HC3	HC4	Boot	Boot teste	Boot duplo
50	0.00	1.00	5.94	9.68	6.86	3.38	6.94	6.48	5.04
	0.04	4.57	20.40	13.74	8.88	3.94	8.84	8.54	6.26
	0.08	20.86	41.70	16.58	10.16	4.08	9.94	9.42	7.58
	0.12	95.27	55.46	18.26	10.70	3.34	10.82	8.48	6.98
100	0.00	1.00	5.40	7.54	5.76	4.02	5.90	5.88	4.56
	0.04	4.57	19.84	9.46	6.94	4.12	7.04	6.80	5.14
	0.08	20.86	39.42	10.18	7.36	3.88	7.56	7.26	5.80
	0.12	95.27	50.00	11.18	7.96	3.84	8.04	7.88	7.14
150	0.00	1.00	4.90	6.56	5.38	4.02	5.42	5.32	4.78
	0.04	4.57	19.24	7.90	6.18	4.24	6.34	6.12	4.68
	0.08	20.86	37.92	7.80	6.18	3.90	6.10	6.24	5.74
	0.12	95.27	49.06	8.28	6.52	4.10	6.46	6.74	6.54

Os resultados apresentados na tabela 3 conduzem a importantes conclusões. Em primeiro lugar, o teste que utiliza o estimador de mínimos quadrados ordinários é amplamente liberal quando há desvios de homocedasticidade. Em segundo lugar, o teste que utiliza o estimador de White se mostra liberal, sendo tanto mais liberal quanto mais forte é o grau de heterocedasticidade (medido por λ). Por exemplo, quando $n = 100$ e $\gamma = 0.12$ o teste quase- t baseado no estimador de White rejeita a hipótese nula (incorretamente) mais de 11% das vezes, ou seja, mais que duas vezes o tamanho nominal especificado para o teste (de 5%). Os testes quase- t que utilizam os estimadores HC3 e de *bootstrap* ponderado também se revelam liberais, rejeitando a hipótese nula (incorretamente) com frequência superior ao esperado com base no tamanho nominal do teste. Estes testes são, contudo, menos liberais do que o que utiliza o estimador de White; por exemplo, quando $n = 100$ e $\gamma = 0.12$, as taxas de rejeição dos testes que utilizam esses estimadores encontram-se em torno de 8%. No que se refere aos testes de *bootstrap*, o teste que utiliza o esquema simples de reamostragem de *bootstrap* também é liberal, apresentando desempenho ligeiramente superior ao teste que utiliza o estimador de *bootstrap* ponderado na construção da estatística de teste. O teste baseado no esquema de

bootstrap duplo em geral apresenta bom desempenho. Por fim, o estimador HC4, que havia apresentado desempenho pobre quando julgado de acordo com viés e erro quadrático médio, apresenta aqui desempenho muito bom, conduzindo a testes associados cujas taxas de rejeição encontram-se próximas àquelas obtidas pelo teste de *bootstrap* duplo, sem, contudo, requerer o custo computacional necessário para a realização de tal teste. Em suma, os dois testes que apresentam desempenho satisfatório em amostras finitas são o que utiliza o estimador HC4 proposto na seção 3 e o teste baseado no esquema de reamostragem de *bootstrap* duplo descrito na seção 4.

A matriz de regressores X utilizada no experimento descrito acima possui pontos de alta alavancagem. De fato, três das 50 observações possuem h_i correspondentes que ultrapassam $3p/n = 0.12$. Cribari-Neto e Zarkos (2001) argumentam que a existência de tais pontos possui maior influência sobre o desempenho dos diferentes estimadores consistentes e testes associados do que o grau de heterocedasticidade em si. A existência de pontos de alta alavancagem na matriz X tende a conduzir a testes liberais.⁴ Os resultados apresentados acima revelam que este problema pode ser evitado utilizando o estimador HC4 na construção de estatísticas de teste ou empregando um esquema de *bootstrap* duplo.

A fim de examinar o efeito de pontos de alavancagem sobre as diferentes estratégias de inferência consideradas, as três observações cujas medidas associadas de alavancagem ultrapassavam $3p/n = 0.12$ foram removidas e substituídas por três novas observações para x geradas da mesma distribuição. Checou-se, então, se esta nova amostra de 50 valores para a covariável possuía pontos de alavancagem, os quais, quando detectados, foram removidos e substituídos por outros. O processo só foi encerrado quando a amostra x_1, \dots, x_{50} não apresentava indícios de possuir pontos de alavancagem. As simulações foram, então, realizadas novamente, desta vez usando o novo conjunto de valores para x . Os valores de γ foram alterados a fim de fornecerem graus de heterocedasticidade semelhantes aos do experimento anterior. Os resultados correspondentes aos apresentados nas tabelas 1, 2 e 3 encontram-se respectivamente nas tabelas 4, 5 e 6.

Os resultados relativos à análise de viés (tabela 4) mostram que os vieses dos estimadores consistentes são significativamente reduzidos quando não há pontos de alavancagem na matriz-modelo. Por exemplo, quando $n = 50$ e $\lambda = 21.19$, o viés relativo total do estimador de White é de cerca de 18% contra cerca de 44% no caso onde há pontos de alavancagem (neste caso com $\lambda = 20.86$ na tabela

⁴Os estimadores corrigidos propostos por Cribari-Neto et alii (2000), que são definidos como modificações do estimador de White, também apresentam sensibilidade à existência de pontos de alavancagem.

1). É interessante notar que o estimador HC4 aqui se mostrou menos viesado do que o estimador HC3 sob homocedasticidade ($\lambda = 1$). Em suma, os vieses de todos os estimadores consistentes foram reduzidos substancialmente quando os pontos de alavancagem foram substituídos por pontos que não apresentavam indícios de alavancagem. As raízes quadradas dos erros quadráticos médios totais, por outro lado, aumentaram no novo experimento (tabela 5 comparada à tabela 2). Isto ocorre porque pontos de alavancagem tendem a atuar como atratores da reta de regressão, atraindo-a para próximo de si e induzindo, assim, baixa variabilidade. A tabela 6 apresenta os tamanhos estimados dos diferentes testes quase- t para o novo experimento. Nota-se que as distorções de tamanho dos testes são menores relativamente às verificadas na tabela 3. Em particular, o teste que utiliza o estimador de White para construção da estatística de teste se mostra consideravelmente mais confiável. É importante ainda notar que: os desempenhos dos testes que utilizam os estimadores HC3 e de *bootstrap* ponderado são mais uma vez semelhantes; o estimador HC4 conduziu novamente a testes associados quase- t mais confiáveis do que o estimador HC3, mesmo em situação onde não há pontos de alavancagem; o teste que utiliza o mecanismo de *bootstrap* duplo em geral mostrou-se confiável.

No geral, os resultados de inferência deste segundo experimento numérico, comparativamente ao experimento anterior, revelam que a presença de pontos de alavancagem na matriz X conduz a testes quase- t liberais. Os resultados dos dois experimentos tomados em conjunto sugerem a utilização do estimador HC4, aqui proposto, ou de um mecanismo de teste de *bootstrap* com reamostragem de dois níveis (*bootstrap* duplo) quando da realização de testes quase- t em modelos possivelmente heterocedásticos de regressão.

Tabela 4
Vieses relativos totais, sem alavancagem

n	γ	λ	MQO	White	HC3	HC4	Boot
50	0.00	1.00	0.0067	0.1123	0.1249	0.0874	0.1261
	0.69	4.60	0.3222	0.1378	0.1378	0.1326	0.1389
	1.38	21.19	0.4665	0.1787	0.1353	0.1647	0.1342
	2.07	97.52	0.6984	0.1889	0.1609	0.2215	0.1565
100	0.00	1.00	0.0028	0.0596	0.0622	0.0509	0.0593
	0.69	4.60	0.8299	0.1149	0.0731	0.1269	0.0744
	1.38	21.19	0.2927	0.0764	0.0687	0.0792	0.0664
	2.07	97.52	0.5292	0.0923	0.0782	0.1131	0.0745
150	0.00	1.00	0.0004	0.0328	0.0430	0.0282	0.0391
	0.69	4.60	0.3574	0.0439	0.0442	0.0377	0.0414
	1.38	21.19	0.5219	0.0565	0.0438	0.0457	0.0426
	2.07	97.52	0.6409	0.0686	0.0402	0.0486	0.0372

Tabela 5
REQM totais ($\times\sqrt{5000}$), sem alavancagem

n	γ	λ	MQO	White	HC3	HC4	Boot
50	0.00	1.00	5.8929	8.0601	9.3283	9.2718	9.4734
	0.69	4.60	20.8344	25.7336	30.6714	31.7457	30.9418
	1.38	21.19	108.3848	109.5654	131.8358	139.0886	132.9407
	2.07	97.52	537.1535	523.4387	639.6525	682.9233	641.9539
100	0.00	1.00	2.2533	3.2327	3.4949	3.5168	3.5677
	0.69	4.60	10.0295	10.7744	11.8483	12.2449	12.0145
	1.38	21.19	57.1778	46.3795	51.4658	53.8828	52.0906
	2.07	97.52	296.6819	220.4315	245.0577	257.9918	250.7189
150	0.00	1.00	1.0265	1.4222	1.4932	1.4826	1.5372
	0.69	4.60	5.8635	4.7262	4.9922	5.0114	5.1013
	1.38	21.19	33.7676	20.6284	21.8315	22.0286	22.1987
	2.07	97.52	169.7126	95.4794	101.0987	102.3114	102.1514

Tabela 6
Taxas de rejeição de testes quase- t ao nível nominal de $\alpha = 5\%$, sem alavancagem

n	γ	λ	MQO	white	HC3	HC4	boot	boot teste	boot duplo
50	0.00	1.00	5.14	5.96	4.84	4.88	4.82	4.76	4.02
	0.69	4.60	8.66	7.38	5.36	5.48	5.40	5.26	4.72
	1.38	21.19	15.72	8.46	6.40	6.38	6.36	6.44	5.38
	2.07	97.52	20.96	8.20	6.22	6.08	6.30	6.20	6.62
100	0.00	1.00	5.70	6.50	5.70	5.74	5.70	5.64	4.54
	0.69	4.60	9.54	6.84	6.04	5.98	5.92	5.88	4.58
	1.38	21.19	15.90	7.24	6.24	6.14	6.32	6.40	5.28
	2.07	97.52	22.92	7.28	6.12	5.82	6.22	6.52	7.00
150	0.00	1.00	5.24	5.70	5.14	5.18	5.22	5.18	5.02
	0.69	4.60	9.98	6.26	5.52	5.58	5.72	5.52	4.74
	1.38	21.19	15.90	6.36	5.48	5.50	5.68	5.96	5.32
	2.07	97.52	20.82	6.58	6.02	5.98	5.96	6.22	5.84

6. Discussão e conclusões

A modelagem de dados de corte transversal encontra tipicamente padrões de heterocedasticidade, ou seja, variâncias condicionais que não são constantes ao longo de todas as observações. Uma prática comum é a estimação dos parâmetros da estrutura linear de regressão por mínimos quadrados ordinários, uma vez que este estimador permanece não-viesado e consistente, mesmo quando há heterocedasticidade de forma desconhecida no modelo. Contudo, usa-se algum estimador consistente de sua matriz de covariância para a realização de inferências, uma vez que o estimador tradicional, $\hat{\sigma}^2(X'X)^{-1}$, é viesado e não é consistente quando a suposição de homocedasticidade é violada.

O estimador mais comumente utilizado em aplicações práticas é o proposto por

Halbert White em um influente artigo publicado na *Econometrica* em 1980. Nossas simulações mostram, todavia, que a utilização deste estimador para construção de testes quase- t sobre os parâmetros de regressão conduz a testes liberais. Uma consequência deste fato é que podemos encontrar significância para variáveis que, de fato, não são significantes.

O estimador HC3 (Davidson e MacKinnon, 1993) e (Long e Ervin, 2000) tem sido apontado como o que possui melhor desempenho em pequenas amostras para realização de testes quase- t . Long e Ervin (2000) chegam a afirmar que, quando o tamanho amostral for menor ou igual a 250, o estimador HC3 deve ser preferido, e complementam estimulando autores de softwares estatísticos a incluir este estimador em seus programas. Os nossos resultados mostram que esse estimador possui um comportamento em amostras finitas semelhante ao estimador de *bootstrap* ponderado. Ambos, contudo, conduzem a testes liberais quando há pontos de alavancagem na matriz de regressores. Propomos então uma modificação desse estimador, especificamente desenhada para levar em consideração o efeito desses pontos sobre inferências resultantes, e os nossos resultados numéricos revelam que este novo estimador, denominado HC4, possui desempenho em amostras pequenas superior ao do estimador HC3. De fato, este novo estimador possui desempenho equivalente, quando utilizado para inferência via testes quase- t , ao de testes realizados via esquemas de *bootstrap* duplo, que também se mostraram muito eficazes. O estimador HC4, contudo, é substancialmente mais simples do ponto de vista prático e computacional.

Referências

- Chesher, A. & Jewitt, I. (1987). The bias of a heteroskedasticity consistent covariance matrix estimator. *Econometrica*, 55:1217–22.
- Cribari-Neto, F. (1999). C for econometricians. *Computational Economics*, 14:135–149.
- Cribari-Neto, F., Ferrari, S. L. P., & Cordeiro, G. M. (2000). Improved heteroscedasticity-consistent covariance matrix estimators. *Biometrika*, 87:907–918.
- Cribari-Neto, F. & Gois, M. C. A. (2002). Uma análise de Monte Carlo do desempenho de matrizes de covariância sob heterocedasticidade de forma desconhecida. *Revista Brasileira de Economia*, 56:309–334.

- Cribari-Neto, F. & Zarkos, S. G. (1999). Bootstrap methods for heteroskedastic regression models: Evidence on estimation and testing. *Econometric Reviews*, 18:211–228.
- Cribari-Neto, F. & Zarkos, S. G. (2001). Heteroskedasticity-consistent covariance matrix estimation: White’s estimator and the bootstrap. *Journal of Statistical Computation and Simulation*, 68:391–411.
- Davidson, R. & MacKinnon, J. G. (1993). *Estimation and Inference in Econometrics*. Oxford University Press, New York.
- Davison, A. C. & Hinkley, D. V. (1997). *Bootstrap Methods and their Application*. Cambridge University Press, New York.
- Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Annals of Statistics*, 7:1–26.
- Efron, B. & Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. Chapman & Hall, New York.
- Judge, G. C., Hill, R. C., Griffiths, W. E., Lutkepohl, H., & Lee, T. (1988). *Introduction to the Theory and Practice of Econometrics*. Wiley, New York, 2 edition.
- Long, J. S. & Ervin, L. H. (2000). Using heteroskedasticity-consistent standard errors in the linear regression model. *The American Statistician*, 54:217–224.
- MacKinnon, J. G. & White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics*, 29:305–325.
- Stallman, R. M. (1999). *Using and Porting the GNU Compiler Collection*. The Free Software Foundation, Boston.
- White, H. (1980). A heteroskedasticity-consistent covariance matrix and a direct test for heteroskedasticity. *Econometrica*, 48:817–838.
- Wooldridge, J. M. (2000). *Introductory Econometrics: A Modern Approach*. South-Western College Publishing, Cincinnati.
- Wu, C. F. J. (1986). Jackknife, bootstrap and other resampling methods in regression analysis. *Annals of Statistics*, 14:1261–1295.