FUNDAÇÃO GETULIO VARGAS

EPGE

Escola de Pós-Graduação em Economia

Ensaios Econômicos

Escola de

Pós Graduação

em Economia

da Fundação

Getulio Vargas

Nº 516 ISSN 0104-8910

Variáveis intrumentais e o MGM: Uso de momentos condicionais

Renato G. Flôres Jr

Novembro de 2003

VARIÁVEIS INSTRUMENTAIS E O MGM; USO DE MOMENTOS CONDICIONAIS

Renato G. Flôres Jr.

(Outubro, 2003)

RESUMO

O presente texto desenvolve, com fins didáticos, as aplicações do Método Generalizado dos Momentos (MGM) ao procedimento de variáveis instrumentais, em modelos lineares e não-lineares. Faz parte de obra (livro) em elaboração.

ABSTRACT

This text presents how the Generalised Method of Moments (GMM) can be used to derive instrumental variables estimators for linear and non-linear models. It makes for a chapter of a work (textbook) in progress.

Capítulo 2

VARIÁVEIS INSTRUMENTAIS E O MGM; USO DE MOMENTOS CONDICIONAIS

1. Variáveis instrumentais e o MGM: modelos lineares.

Consideremos o modelo linear usual, a k explicativas e T observações

$$y = X\,\beta + e \tag{1.1}$$

na situação onde se sabe existir uma outra matriz de projeto Z, Txm, $m \geq k$ (ou dim X $\leq$ dim Z[1]), com os Z's não necessàriamente distintos dos X's, para a qual se tem a condição

$$E\,(\,y - X\,\beta\mid Z\,) = E\,(e\mid Z) = 0 \qquad . \tag{1.2}$$

Traduzindo (2) como as condições de momento:

$$E\,(\,Z'(y - X\,\beta\,)\,) = 0 \qquad , \tag{1.3}$$

o MGM procuraria minimizar uma forma quadrática do tipo

$$(y - X\,\beta\,)'\,Z\,A_T Z'\,(y - X\,\beta\,) \qquad , \tag{1.4}$$

---

[1] O leitor deve notar dois abusos correntes que estão sendo usados aqui: os X's e os Z's referem-se também às variáveis que compõem as colunas das matrizes de observações X e Z, resp.; dimX e dimZ sendo tanto a dimensão dos vetores aleatórios X e Z como a dos subespaços vetoriais gerados pelas

sendo a ponderação ótima dada imediatamente por

$$A_T \propto E\,(Z'ee'Z)^{-1} = [E(E\,(Z'ee'Z \mid Z))]^{-1} = [E(Z'E\,(ee' \mid Z\,)\,Z)]^{-1} \quad . \tag{1.5}$$

É importante examinar cuidadosamente as identidades em (1.5), pois grande parte das questões afetas às variáveis instrumentais está relacionada a como se pode calcular a matriz de covariância que nela figura. Sob homocedasticidade e não correlação entre os erros, tem-se

$$[E(Z'E\,(ee' \mid Z\,)\,Z)]^{-1} = [E(Z'(\sigma^2 I_T)\,Z)]^{-1} \propto [E(Z'Z)]^{-1} \quad .$$

A esperança acima, que nada mais é do que a matriz de momentos cruzados das variáveis Z, é aproximada por $1/T\ (Z'Z)$. É também comum trabalhar-se na distribuição condicionada pelas observações Z. Em ambos os casos, (1.4) se torna:

$$(y - X\beta)'\, Z\ (Z'Z)^{-1}\ Z'\,(y - X\beta) = (y - X\beta)'\, P_z\,(y - X\beta) \tag{1.6}$$

e, consequentemente, o estimador MGM coincide com o estimador de variáveis instrumentais (VI) dado por:

$$\beta_{VI} = (X'\,P_z\,X)^{-1}\, X'\,P_z\,y \tag{1.7}$$

devido originalmente a Sargan (1958). Pela Proposição 3 do capítulo1, (1.7) é o estimador MGM ótimo e, portanto, a sua variância assintótica será:

$$T.\sigma^2\,(X'\,Z\ (Z'Z)^{-1}\ Z'\,X)^{-1} = T.\sigma^2\,(X'\,P_z\,X)^{-1} \tag{1.8}$$

É interessante comparar essa obtenção imediata da varass, pelos resultados do MGM, com a usual, que partiria da expressão (1.7).

Suponhamos, no entanto, que

$$E\,(ee' \mid Z\,) = \Omega \quad , \tag{1.9}$$

sendo Ω, em princípio, uma matriz de covariância qualquer. Neste caso, o ponderador otimal se torna $A_T \propto (Z'\Omega Z)^{-1}$, e a forma a ser minimizada

---

colunas de X e Z, resp. Seria mais ortodoxo escrever posto(X)≤posto(Z) , uma vez que, naturalmente,

$$(y - X\beta)' Z\ (Z'\Omega Z)^{-1}\ Z' (y - X\beta) \qquad (1.10)$$

o que dá, por exemplo, o correto estimador de VI sob heterocedasticidade:

$$\beta_{VIh} = [\ X'Z\ (Z'\Omega Z)^{-1}\ Z'X\ ]^{-1} X'Z\ (Z'\Omega Z)^{-1}\ Z'y \quad , \qquad (1.11)$$

podendo o conjunto de variáveis que gera o mesmo subespaço que $Z\,(Z'\Omega Z)^{-1} Z'$ ser considerado como o de instrumentos otimais. Obviamente, requer-se agora a estimação de Ω.

Se Ω é diagonal, por exemplo[2], caracterizando a existência apenas de heterocedasticidade, duas maneiras são possíveis: ou pelos resíduos de um MQO inicial, ou pelos de um MQ2 ou VI inicial (como em (1.7)). Chamando de $\hat{e}^0$ o vetor Tx1 resultante dessa primeira estimação, isso significa que a diagonal de $(\hat{e}^0\,\hat{e}^{0\prime})$ estaria aproximando a de Ω. Tal aproximação, que é melhor do que nada na ausência de informação adicional sobre a heterocedasticidade, pode entretanto ser demasiado grosseira.

Note-se ademais que, se posto(Z)=posto(X) (ou, como visto, se as duas matrizes tiverem o mesmo número de colunas/variáveis), de tal forma que (X'Z) seja inversível,

$$\beta_{VIh} = (Z'X)^{-1} Z'y = \beta_{VI} \quad ,$$

sendo inócua a "correção da heterocedasticidade". Esse fato constitui um interessante exemplo de que a condição de otimalidade para a variância assintótica do estimador MGM não é garantia de eficiência (ou, em outras palavras, não é suficiente).

Com essas ressalvas, a varass do estimador (1.11) é:

$$T.[\ X'Z\ (Z'\Omega Z)^{-1}\ Z'X\ ]^{-1} \quad , \qquad (1.12)$$

que deve ser comparada com (1.8).

T>m e k.

[2] Matrizes Ω não-diagonais podem muitas vezes ser obtidas diretamente da estrutura do modelo. Vide os exemplos no capítulo sobre dados em painel para um destes casos.

Se Ω não é diagonal, no contexto de processos ao longo do tempo, estaríamos diante de uma estrutura de auto-correlação entre os erros. Cumby et al. (1983), em um artigo hoje clássico, discutiram as variáveis instrumentais sob heterocedasticidade contemporânea e correlação serial das variáveis envolvidas. Neste caso, os dois métodos sugeridos anteriormente para a estimação preliminar da matriz de covariâncias não se aplicam mais, pois, devido à correlação entre as observações, é melhor se obter diretamente uma estimativa de

$$\lim_{T\uparrow\infty} (1/T)\, E(\, Z'ee'Z\,) \qquad ,$$

a matriz de covariância a longo prazo do processo $\{\, e_t Z_t' \,\}$[3], suposto estacionário de segunda ordem, que aproximará o ponderador otimal.

A forma proposta é obter resíduos preliminares $\hat{e}^0$, por VI, transformá-los para $\{\hat{e}_t^0 \;\; Z_t'\}_{t=1,\ldots,T}$ e, então, através de um procedimento não-paramétrico como os discutidos no capítulo 4, obter um estimador F da matriz (1/T) E(Z'ee'Z). Este estimador possui dois usos igualmente importantes.

O primeiro é que, nesse caso, a correta aproximação da varass do estimador β_{VI} , definido em (1.7), é dada pela expressão

$$T^2.(\, X'\, P_z\, X\,)^{-1}\, X'Z\;\; (Z'Z)^{-1}\, F\, (Z'Z)^{-1}\, Z'X\, (X'\, P_z\, X)^{-1} \qquad . \qquad (1.13)$$

O segundo é que o estimador MGM resultará[4], como sabido, da minimização de

$$(y - X\,\beta\,)'\, Z\, F^{-1} Z'\, (y - X\,\beta\,) \qquad , \qquad (1.14)$$

assumindo a forma

$$\beta_{MGM} = (X'Z\, F^{-1}\, Z'X)^{-1}\, X'Z\, F^{-1}\, Z'y \qquad (1.15)$$

e tendo a variância assintótica aproximada, naturalmente, por

$$T^2.(X'Z\, F^{-1}\, Z'X)^{-1} \qquad . \qquad (1.16)$$

---

[3] Z_t é a t-ésima linha da matriz Z.

A comparação entre (1.13) e (1.16) não é evidente à primeira vista[5]. Entretanto, (1.7) advém do uso do ponderador sugerido por (1.5). Desse modo, (1.13) é também uma versão da fórmula (3.13), na Proposição 2 do capítulo 1, e, portanto, (1.16) é assintoticamente mais eficiente.

2. Variáveis instrumentais e o MGM: modelos não-lineares.

A situação descrita por (1.2) e (1.3) pode ser naturalmente generalizada para o caso de um modelo não linear $h(X_t;\, \theta) = h_t(\theta) = e_t$, onde valem as q condições de momento

$$E\,(\,Z'h(\theta)\,) = 0 \tag{2.1}$$

sendo $h(\theta)$ o vetor Tx1 obtido pelo empilhamento dos "erros" $h_t(\theta)$. Caso, ao invés de uma equação, tenha-se um modelo não linear a n equações, chamando de $H(\theta)'$ a matriz nxT com as T observações para cada uma das n equações, (2.1) se transforma, na versão matricial, em

$$E\,(\,H(\theta)'Z\,) = 0 \qquad , \tag{2.2}$$

matriz que engloba as q=n.m condições de momento. Versão essa que pode ser escrita de forma vetorial como

$$E\,[\,\mathrm{vec}\,(H(\theta)'Z)\,] = 0 \qquad , \tag{2.3}$$

que é mais próxima a (2.1). Por essa razão, sem grande perda de generalidade, continuaremos a desenvolver, a seguir, o caso uni-equacional.

O estimador MGM é obtido pela minimização de

---

4 Este estimador é também, algumas vezes, denotado por CHO, devido às iniciais dos três autores do artigo citado neste exemplo.

5 Novamente, é valioso também compará-las com desenvolvimentos clássicos, como Sargan (1959), e o próprio Cumby et al. (1993).

$$h(\theta)'Z\, A_T\, Z'h(\theta) \quad , \tag{2.4}$$

e o ponderador otimal, trabalhando-se na lei condicionada pelos Z, proporcional ao inverso de

$$E(\, Z'h(\theta)h(\theta)'Z \mid Z\,) = Z'\, E(\, h(\theta)h(\theta)' \mid Z\,)\, Z = Z' \textstyle\sum_h Z \quad . \tag{2.5}$$

Note-se que a matriz $\sum_h$, como a sua correspondente Ω em (1.8), é de dimensão TxT. A varass do estimador eficiente será aproximada por

$$T\, (\, [\partial h/\partial\theta_T]'Z\, (Z'\, S_h\, Z)^{-1}\, Z'\, [\partial h/\partial\theta_T]\,)^{-1} \quad , \tag{2.6}$$

onde S_h é um estimador consistente de $\sum_h$. Para tal, faz-se necessária uma estimativa inicial θ_T^0 do parâmetro, que pode ser obtida, por exemplo, minimizando a função objetivo resultante ao se fazer $A_T = I_m$ em (2.4).

Obviamente, o número de instrumentos é, em princípio, arbitrário. No entanto, é possível mostrar que *dado um conjunto linearmente independente Z, Txq, de candidatos a instrumentos*, existe um limite inferior para a variância acima (sempre segundo a ordem de Loewner), entre todos os possíveis subespaços de instrumentos definidos a partir de Z. Esse limite é atingido quando se usa *todos os instrumentos disponíveis*.

Para demonstrar esse fato, suponha-se uma nova matriz de instrumentos ZB, onde B é $m \times m'$, $m' < m$. Se B é a matriz identidade de ordem m', completada inferiormente com zeros, isto equivale a selecionar os m' "primeiros" instrumentos. Chamando de Δ, para não carregar a notação, o $\text{Plim}\ (1/T)\ (Z'\ [\partial h/\partial\theta_T])$, a diferença entre os inversos da antiga e da nova varass será

$$\Delta'\Phi^{-1}\Delta - \Delta'B(B'\Phi\, B)^{-1}B'\Delta \quad , \tag{2.7}$$

onde $\Phi = \text{Plim}\ (1/T)\ Z'h(\theta)\ h(\theta)'Z$. Tomando-se, como na seção 3 do capítulo 1, a ψ, raiz quadrada de Φ^{-1}, a expressão acima equivale a

$$\Delta'\psi\ (\ I_q - \psi^{-1}B(B'\psi^{-1}.\psi^{-1}B)^{-1}B'\psi^{-1}\ )\ \psi\Delta \qquad .$$

Como no interior do parêntese tem-se um operador projeção, a diferença (2.7) é semidefinida positiva. Mas como ela é a diferença entre os inversos das matrizes de covariância, isto implica em que é válida a

Proposição 1. Sob as condições (2.1), e sendo válida a expressão (2.5), se se "reduz" por combinações lineares o conjunto dos instrumentos, a variância do estimador obtido da minimização de (2.4), com $\lim_{T\uparrow\infty} A_T = \Phi^{-1}$, tende a aumentar.

Este resultado pode levar à crença de que "quanto mais instrumentos melhor". Três pontos se contrapõem a isto. Os ganhos de eficiência com o aumento do número de instrumentos são, em geral, decrescentes, a Proposição 1 fornecendo na realidade uma garantia negativa: ao se reduzir os instrumentos a varass *não* diminue. Em segundo lugar, um número excessivo de instrumentos pode ocasionar sérios vieses em amostras finitas. Por fim, o resultado a seguir mostra que, em muitos casos, o ganho de fato será nulo.

Considerando que, em nosso modelo não linear $h(X_t;\ \theta) = e_t$, h seja de classe C^1 em Θ, e sempre restringindo-se ao universo das combinações lineares finitas de todos os instrumentos possíveis, *existe um conjunto ótimo de instrumentos*. Esse conjunto é dado pela matriz Txp:

$$Z = \Sigma_h^{-1}\ [\partial h/\partial\theta_0] \qquad . \tag{2.8}$$

O uso desses instrumentos fornece uma varass para o estimador igual a

$$T.\ (\ [\partial h/\partial\theta_0]\ \Sigma_h^{-1}\ [\partial h/\partial\theta_0]\ )^{-1} \tag{2.9}$$

que é aproximada pela seguinte versão

$$T. \left([\partial h/\partial \theta_T]\, S_h^{-1}\, [\partial h/\partial \theta_T] \right)^{-1} \quad . \qquad (2.10)$$

A prova do fato acima segue passos muito semelhantes à anterior, mostrando-se que a inversa de (2.9) domina a inversa da varass associada a qualquer outro conjunto de instrumentos, bastando apenas usar a raiz quadrada de Σ_h^{-1} . Pode-se então enunciar a

Proposição 2. Sob as condições (2.1), e sendo válida a expressão (2.5), se h é de classe C^1 em Θ, o estimador dado pelos instrumentos na matriz de projeto (2.8) atinge a varass mínima na classe de todos os estimadores por variáveis instrumentais.

A proposição acima é uma das poucas afirmativas em econometria a fornecer um conjunto de instrumentos *ótimos*; ideal muito procurado mas difícil de ser obtido em geral, sendo ainda objeto de pesquisa. Note-se que, curiosamente, os instrumentos ótimos são como que definidos endògenamente, a partir da própria estrutura do modelo. Cabe então naturalmente perguntar se a condição (2.1), fundamental, é satisfeita por (2.8). O leitor é convidado a mostrar esse fato no Exercício 1, ao final do capítulo.

Confirma, também, a Proposição 2, o comentário feito à anterior. Caso se esteja trabalhando com uma matriz Z* de instrumentos, que *contenha* a coleção (2.8) como submatriz, a varass do estimador MGM associado à Z* será *idêntica* a (2.9).

Como definidos em (2.8), a obtenção dos instrumentos ótimos necessita, além de um estimador consistente S_h de Σ_h, que um valor inicial θ_T^0 seja estimado para o parâmetro. Nesse caso, se um conjunto Z qualquer de instrumentos não fôr disponível, uma primeira estimação, por mínimos quadrados não-lineares, por exemplo, terá que ser efetuada.

O Exercício 2 relembra que toda a discussão dessa seção pode ser generalizada para o caso em que $h(X_t; \theta) = e_t$ é um sistema de n equações não-lineares, o processo "de erros" $\{e_t\}$ sendo, conformemente, n-dimensional.

3. Uso de momentos condicionais.

Considere-se, nos últimos casos da seção anterior, modelos em série de tempo onde as condições de ortogonalidade, para cada instante t, sejam dadas não pelos momentos absolutos, como no Capítulo 1 e nas seções precedentes, porém condicionando-se a esperança por uma filtração $\{J_t\}$. Para os não familiarizados com o têrmo, assinalamos que uma filtração é uma sequência de conjuntos de informação[6] $\{J_t\}$ tal que, se $t<t'$ então $J_t \subseteq J_{t'}$. Os momentos condicionais de base, fornecidos pelo contexto do problema passam então a se escrever, para cada t:

$$E\,(\,H\,(Y_t,X_t;\theta)\mid J_t\,) = 0 \qquad , \qquad (3.1)$$

onde $H(Y_t,X_t;\theta)=H_t$, nx1, é o modelo não-linear em questão.

Supondo-se disponíveis m instrumentos, na forma de um processo estocástico vetorial $\{\,Z_t\,\}$, adaptado a $\{J_t\}$, ou seja, composto por séries Z_{jt} , $j=1,...,m$, tais que, para todo t, os Z_{jt} estejam (ou sejam mensuráveis) em J_t ; as equações (3.1) se traduzem nas $q=n.m$ condições de momento, para cada t:

$$E\,(\,Z_t \otimes H_t(\theta)\,) = 0 \qquad . \qquad (3.2)$$

[6] Ou, mais técnicamente, de σ-álgebras.

As igualdades acima implicam que o processo $\{ Z_t \otimes g_t(\theta) \}$ é estacionário de primeira ordem, o que autoriza definir um estimador MGM pela minimização de:

$$(\Sigma_t(Z_t \otimes H_t(\theta)))' A_T (\Sigma_t(Z_t \otimes H_t(\theta))) \quad . \qquad (3.3)$$

Se o processo $\{ Z_t \otimes H_t(\theta) \}$ fôr, ao menos, estacionário de segunda ordem, o ponderador otimal será dado pelo inverso de

$$\Omega^* = \text{Plim } 1/T \; (\Sigma_t Z_t \otimes H_t(\theta_T)) (\Sigma_t Z_t \otimes H_t(\theta_T))' \quad . \qquad (3.4)$$

Denotando por D(Z) a matriz n.m x p[7]

$$D(Z) = E (Z_t \otimes [\partial H_t(\theta)/\partial\theta_0] \mid J_t) \quad , \qquad (3.5)$$

a varass do estimador correspondente será

$$(D(Z)' (\Omega^*)^{-1} D(Z))^{-1} \quad . \qquad (3.6)$$

Mais do que na seção anterior, o número de instrumentos possíveis é, em princípio, muito elevado, ficando a sua escolha ainda mais arbitrária. Cabe então, de modo análogo, perguntar se não existe um conjunto ótimo de candidatos. Para processos estocásticos gerais $\{Z_t\}$ e $\{H_t(\theta)\}$, ainda que estacionários, a resposta pode ficar muito complicada. Hansen (1985) e Hansen et al. (1989) lançam algumas luzes sobre a questão. No segundo artigo, em um contexto bem geral - onde as condições de ortogonalidade podem valer relativamente a diversos períodos anteriores - os autores conseguem encontrar um limite inferior para a variância do estimador definido por (3.3). No entanto, este limite é expresso em função dos coeficientes de uma determinada representação *à la* Wold para o processo de erros; representação esta que, na maioria dos casos práticos, é impossível de se identificar. As exceções conhecidas compreendem a estimação de coeficientes de certos processos ARMA - sob a hipótese de que as inovações sejam diferenças de martingales homocedásticas - que já havia sido estudada por Stoica et al. (1985), e os exemplos constantes no próprio Hansen (1985) e em Heaton and Ogaki (1991).

Sob a hipótese iid para os processos envolvidos, Chamberlain (1987) obtém de forma elegante um limite inferior válido em um contexto semi-paramétrico. Newey (1993), também supondo $\{Z_t\}$ e $\{H_t(\theta)\}$ iid, trabalhando na condicional dada pelos instrumentos, obtém a matriz de funções $A(Z_t)$, rxn , $r \geq p$, que define o estimador eficiente, *função dos instrumentos* Z_t, que melhor "mistura" as condições condicionais dadas pelos $H_t(\theta)$, obtendo as r condições de momento:

$$E\,(\,A(Z_t)\;g_t(\theta)\,) = 0 \qquad . \qquad (3.7)$$

A condição iid permite que se defina as matrizes

$$D°(Z_t) \;=\; E\,(\,[\partial H_t(\theta)/\partial\theta_0] \mid Z_t\,)\;,\;\; nxp \qquad , \qquad (3.8)$$

$$\Omega°(Z_t) \;=\; E\,(\,H_t(\theta_0)\,H_t(\theta_0)' \mid Z_t\,)\;,\; nxn \qquad , \qquad (3.9)$$

a partir das quais a solução ótima é dada por:

$$A°(Z_t) \;=\; C\,D°(Z_t)'\;\Omega°(Z_t)^{-1}\;,\;\; pxn \qquad , \qquad (3.10)$$

onde C é qualquer matriz pxp, inversível.

Utilizando (3.10) em (3.7), tem-se exatamente p condições de momento e, portanto, a varass do estimador associado aos instrumentos (3.10) será o inverso de:

$$E(A°(Z_t)[\partial H_t(\theta)/\partial\theta_0])'\;(E(A°(Z_t)H_t(\theta_0)H_t(\theta_0)'A°(Z_t)'))^{-1}\;E(A°(Z_t)[\partial H_t(\theta)/\partial\theta_0])\;, \quad (3.11)$$

o que, pelas propriedades da esperança condicional, dará:

$$E(D°(Z_t)'\;\Omega°(Z_t)^{-1}\;D°(Z_t)) \qquad . \qquad (3.12)$$

Para demonstrar a otimalidade, considerando-se um estimador definido por um $A(Z_t)$ qualquer, basta notar que a diferença entre (3.12) e o inverso da fórmula de sua varass - exatamente idêntica à expressão (3.11) - pode ser escrita como:

---

[7] O que implica que o processo $\{Z_t \otimes [\partial H_t(\theta)/\partial\theta_0]\}$ é estacionário de primeira ordem (na esperança

$$E(D°(Z_t)'\Omega°(Z_t)^{-1}D°(Z_t)) - E(A(Z_t)D°(Z_t))' \ (E(A(Z_t)\Omega°(Z_t)^{-1}A(Z_t)'))^{-1} \ E(A(Z_t)D°(Z_t)) \ .$$

Tomando-se a raiz quadrada de (3.12), a prova prossegue como feito nas seções anteriores (vide Exercício 3).

Os instrumentos ótimos definidos por (3.8) a (3.10) necessitam, para serem implementados, o conhecimento ao menos de uma estimativa consistente do parâmetro. No entanto, só isto não basta. As esperanças condicionais (3.8) e (3.9) devem também ser computadas. Newey (1993) sugere dois métodos não-paramétricos para tal, o "do vizinho mais próximo"[8], inspirado em Robinson (1987), e um por aproximação de funções. Advogamos, entretanto, o uso do estimador de Nadaraya-Watson[9], que é extremamente indicado para o cálculo de esperanças condicionais em tais contextos.

Exemplo 1. Seja um problema de regressão não-linear, com observações iid porém erros heterocedásticos:

$$y = f(x \ ; \alpha_0) + e \quad , \quad E[e \mid x] = 0 \quad , \quad E[e^2 \mid x] = h(x \ ; \alpha_0 , \beta_0) \quad ;$$

o que, seguindo a notação em (3.7) equivale a duas ortogonalidades de base

$$E[\ y - f(x \ ; \alpha_0) \mid x] = 0 \quad , \quad E[\ (y - f(x \ ; \alpha_0))^2 - h(x \ ; \alpha_0 , \beta_0) \mid x] = 0 \quad ,$$

a serem usadas para compor os momentos. Supondo que as explicativas x comporão os instrumentos ótimos, estes serão dados a partir da matriz D°(x):

---

condicionada à filtração).

8 Em inglês, *nearest-neighbour*.

9 Excelentes exposições sobre este estimador se encontram na monografia de Rosenblatt (1991) ou no capítulo 8 de Bierens (1996).

$$\begin{bmatrix} \dfrac{\partial f(x,\alpha)}{\partial \alpha} & 0 \\ \dfrac{\partial h(x;\alpha,\beta)}{\partial \alpha} & \dfrac{\partial h(x;\alpha,\beta)}{\partial \beta} \end{bmatrix}$$

e por $\Omega°(x) = \mathrm{var}\,(\,(\,e, e^2\,)' \,|\, x\,)$. Neste caso, a varass do estimador otimal será, em geral, inferior à do de mínimos quadrados com correção para a heterocedasticidade. O leitor deve verificar (Exercício 4) que ambas coincidem se

$$E\,[e^3 | \, x] \neq 0 \quad \text{ou} \quad h(x\,;\alpha_0\,,\beta_0) = h(x\,;\beta_0)\ ,$$

ou seja, os coeficientes da equação não figurarem na especificação da heterocedasticidade. Para maiores detalhes vide Newey (1993).

Exercícios

1. Prove que a condição (2.1) é satisfeita pelo conjunto de instrumentos ótimos definido pela matriz de projeto (2.8).

2. Refaça os resultados da seção 2 para um sistema de n equações não lineares H_t.

3. Prove rigorosamente que a inversa de (3.12) é mínima na classe das variâncias assintóticas definidas pela inversa de (3.11).

4. Prove a afirmativa sobre a varass dos mínimos quadrados com correção de heterocedasticidade, feita no Exemplo 1.

Referências

Burguete, J. F. 1980. Asymptotic theory of instrumental variables in nonlinear regression. Ph.D. Dissertation, North Carolina State University, Raleigh, North Carolina.

Chamberlain, G. 1987. Asymptotic efficiency in estimation with conditional moment restrictions. *J. of Econometrics* 34, 305 - 34.

Cumby, R. E., J. Huizinga and M. Obstfeld. 1993. Two-step two-stage least squares estimation in models with rational expectations. *J. of Econometrics* 21, 333-55.

Davidson, R. and J. G. MacKinnon. 1993. *Estimation and Inference in Econometrics*. Oxford University Press, New York.

Gallant, A. R. and D. W. Jorgenson. 1979. Statistical inference for a system of simultaneous, nonlinear implicit equations in the context of instrumental variables estimation. *Journal of Econometrics* 11, 275 - 302.

Hansen, L. P. 1985. A method for calculating bounds on the asymptotic covariance matrices of generalized method of moments estimators. *J. of Econometrics* 30, 203 - 38.

Hansen, L. P., J. C. Heaton and M. Ogaki. 1988. Efficiency bounds implied by multi-period conditional moment restrictions. *JASA* 83, 863 - 71.

Hansen, L. P. and K. J. Singleton. 1982. Generalized instrumental variables estimation of nonlinear rational expectations models. *Econometrica* 50, 1269 - 86.

Heaton, J. C. and M. Ogaki. 1991. Efficiency bound calculations for a simple model with conditional heteroskedasticity. *Econ. Letters* 35, 167 - 71.

Nelson, C. R. and R. Startz. 1990. The distribution of the instrumental variables estimator and its t-statistic when the instrument is a poor one. *J. of Business* 63, 125 - 40.

Newey, W. K. 1993. Efficient estimation of models with conditional moment restrictions, in G. S. Maddala, C. R. Rao and H. D. Vinod, eds., *Handbook of Statistics*, vol. 11, Elsevier Science Publishers B.V., Amsterdam.

Robinson, P. 1987. Asymptotically efficient estimation in the presence of heteroskedasticity of unknown form. *Econometrica* 55, 875 -891.

Rosenblatt, M. 1991. *Stochastic Curve Estimation.* Institute of Mathematical Statistics, Hayward, Cal.

Sargan, D. 1958. The estimation of economic relationships using instrumental variables. *Econometrica,* vol. 26, 393-415.

Sargan, D. 1959. The estimation of relationships with autocorrelated residuals by the use of instrumental variables. *J. R. Stat. Society,* series B, vol. 21, 91-105.

Stoica, P., T. Soderstrom and B. Friedlander. 1985. Optimal instrumental variable estimates of the AR parameters of an ARMA process. *IEEE Trans. of Automatic Control* 30, 1066 - 74.

ENSAIOS ECONÔMICOS DA EPGE

471. CUSTO DE CICLO ECONÔMICO NO BRASIL EM UM MODELO COM RESTRIÇÃO A CRÉDITO - Bárbara Vasconcelos Boavista da Cunha; Pedro Cavalcanti Ferreira – Janeiro de 2003 – 21 págs.

472. THE COSTS OF EDUCATION, LONGEVITY AND THE POVERTY OF NATIONS - Pedro Cavalcanti Ferreira; Samuel de Abreu Pessoa – Janeiro de 2003 – 31 págs.

473. A GENERALIZATION OF JUDD'S METHOD OF OUT-STEADY-STATE COMPARISONS IN PERFECT FORESIGHT MODELS - Paulo Barelli; Samuel de Abreu Pessoa – Fevereiro de 2003 – 7 págs.

474. AS LEIS DA FALÊNCIA: UMA ABORDAGEM ECONÔMICA - Aloísio Pessoa de Araújo – Fevereiro de 2003 – 25 págs.

475. THE LONG-RUN ECONOMIC IMPACT OF AIDS - Pedro Cavalcanti G. Ferreira; Samuel de Abreu Pessoa – Fevereiro de 2003 – 30 págs.

476. A MONETARY MECHANISM FOR SHARING CAPITAL: DIAMOND AND DYBVIG MEET KIYOTAKI AND WRIGHT – Ricardo de O. Cavalcanti – Fevereiro de 2003 – 16 págs.

477. INADA CONDITIONS IMPLY THAT PRODUCTION FUNCTION MUST BE ASYMPTOTICALLY COBB-DOUGLAS - Paulo Barelli; Samuel de Abreu Pessoa – Março de 2003 – 4 págs.

478. TEMPORAL AGGREGATION AND BANDWIDTH SELECTION IN ESTIMATING LONG MEMORY - Leonardo R. Souza - Março de 2003 – 19 págs.

479. A NOTE ON COLE AND STOCKMAN - Paulo Barelli; Samuel de Abreu Pessoa – Abril de 2003 – 8 págs.

480. A HIPÓTESE DAS EXPECTATIVAS NA ESTRUTURA A TERMO DE JUROS NO BRASIL: UMA APLICAÇÃO DE MODELOS DE VALOR PRESENTE - Alexandre Maia Correia Lima; João Victor Issler – Maio de 2003 – 30 págs.

481. ON THE WELFARE COSTS OF BUSINESS CYCLES IN THE 20TH CENTURY - João Victor Issler; Afonso Arinos de Mello Franco; Osmani Teixeira de Carvalho Guillén – Maio de 2003 – 29 págs.

482. RETORNOS ANORMAIS E ESTRATÉGIAS CONTRÁRIAS - Marco Antonio Bonomo; Ivana Dall'Agnol – Junho de 2003 – 27 págs.

483. EVOLUÇÃO DA PRODUTIVIDADE TOTAL DOS FATORES NA ECONOMIA BRASILEIRA: UMA ANÁLISE COMPARATIVA - Victor Gomes; Samuel de Abreu Pessoa;Fernando A . Veloso – Junho de 2003 – 45 págs.

484. MIGRAÇÃO, SELEÇÃO E DIFERENÇAS REGIONAIS DE RENDA NO BRASIL - Enestor da Rosa dos Santos Junior; Naércio Menezes Filho; Pedro Cavalcanti Ferreira – Junho de 2003 – 23 págs.

485. THE RISK PREMIUM ON BRAZILIAN GOVERNMENT DEBT, 1996-2002 - André Soares Loureiro; Fernando de Holanda Barbosa - Junho de 2003 – 16 págs.

486. FORECASTING ELECTRICITY DEMAND USING GENERALIZED LONG MEMORY - Lacir Jorge Soares; Leonardo Rocha Souza – Junho de 2003 – 22 págs.

487. USING IRREGULARLY SPACED RETURNS TO ESTIMATE MULTI-FACTOR MODELS: APPLICATION TO BRAZILIAN EQUITY DATA - Álvaro Veiga; Leonardo Rocha Souza – Junho de 2003 – 26 págs.

488. BOUNDS FOR THE PROBABILITY DISTRIBUTION FUNCTION OF THE LINEAR ACD PROCESS – Marcelo Fernandes – Julho de 2003 – 10 págs.

489. CONVEX COMBINATIONS OF LONG MEMORY ESTIMATES FROM DIFFERENT SAMPLING RATES - Leonardo R. Souza; Jeremy Smith; Reinaldo C. Souza – Julho de 2003 – 20 págs.

490. IDADE, INCAPACIDADE E A INFLAÇÃO DO NÚMERO DE PESSOAS COM DEFICIÊNCIA - Marcelo Neri ; Wagner Soares – Julho de 2003 – 54 págs.

491. FORECASTING ELECTRICITY LOAD DEMAND: ANALYSIS OF THE 2001 RATIONING PERIOD IN BRAZIL - Leonardo Rocha Souza; Lacir Jorge Soares – Julho de 2003 – 27 págs.

492. THE MISSING LINK: USING THE NBER RECESSION INDICATOR TO CONSTRUCT COINCIDENT AND LEADING INDICES OF ECONOMIC ACTIVITY - JoãoVictor Issler; Farshid Vahid – Agosto de 2003 – 26 págs.

493. REAL EXCHANGE RATE MISALIGNMENTS - Maria Cristina T. Terra; Frederico Estrella Carneiro Valladares – Agosto de 2003 – 26 págs.

494. ELASTICITY OF SUBSTITUTION BETWEEN CAPITAL AND LABOR: A PANEL DATA APPROACH - Samuel de Abreu Pessoa ; Silvia Matos Pessoa; Rafael Rob – Agosto de 2003 – 30 págs.

495. A EXPERIÊNCIA DE CRESCIMENTO DAS ECONOMIAS DE MERCADO NOS ÚLTIMOS 40 ANOS – Samuel de Abreu Pessoa – Agosto de 2003 – 22 págs.

496. NORMALITY UNDER UNCERTAINTY – Carlos Eugênio E. da Costa – Setembro de 2003 – 08 págs.

497. RISK SHARING AND THE HOUSEHOLD COLLECTIVE MODEL - Carlos Eugênio E. da Costa – Setembro de 2003 – 15 págs.

498. REDISTRIBUTION WITH UNOBSERVED 'EX-ANTE' CHOICES - Carlos Eugênio E. da Costa – Setembro de 2003 – 30 págs.

499. OPTIMAL TAXATION WITH GRADUAL LEARNING OF TYPES - Carlos Eugênio E. da Costa – Setembro de 2003 – 26 págs.

500. AVALIANDO PESQUISADORES E DEPARTAMENTOS DE ECONOMIA NO BRASIL A PARTIR DE CITAÇÕES INTERNACIONAIS - João Victor Issler; Rachel Couto Ferreira – Setembro de 2003 – 29 págs.

501. A FAMILY OF AUTOREGRESSIVE CONDITIONAL DURATION MODELS - Marcelo Fernandes; Joachim Grammig – Setembro de 2003 – 37 págs.

502. NONPARAMETRIC SPECIFICATION TESTS FOR CONDITIONAL DURATION MODELS - Marcelo Fernandes; Joachim Grammig – Setembro de 2003 – 42 págs.

503. A NOTE ON CHAMBERS'S "LONG MEMORY AND AGGREGATION IN MACROECONOMIC TIME SERIES" – Leonardo Rocha Souza – Setembro de 2003 – 11págs.

504. ON CHOICE OF TECHNIQUE IN THE ROBINSON-SOLOW-SRINIVASAN MODEL - M. Ali Khan – Setembro de 2003 – 34 págs.

505. ENDOGENOUS TIME-DEPENDENT RULES AND THE COSTS OF DISINFLATION WITH IMPERFECT CREDIBILITY - Marco Bonomo; Carlos Viana de Carvalho – Outubro de 2003 – 27 págs.

506. CAPITAIS INTERNACIONAIS: COMPLEMENTARES OU SUBSTITUTOS? - Carlos Hamilton V. Araújo; Renato G. Flôres Jr. – Outubro de 2003 – 24 págs.

507. TESTING PRODUCTION FUNCTIONS USED IN EMPIRICAL GROWTH STUDIES - Pedro Cavalcanti Ferreira; João Victor Issler; Samuel de Abreu Pessoa – Outubro de 2003 – 8 págs.

508. SHOULD EDUCATIONAL POLICIES BE REGRESSIVE ? Daniel Gottlieb; Humberto Moreira – Outubro de 2003 – 25 págs.

509. TRADE AND CO-OPERATION IN THE EU-MERCOSUL FREE TRADE AGREEMENT - Renato G. Flôres Jr. – Outubro de 2003 – 33 págs.

510. OUTPUT CONVERGENCE IN MERCOSUR: MULTIVARIATE TIME SERIES EVIDENCE - Mariam Camarero; Renato G. Flôres Jr; Cecílio Tamarit – Outubro de 2003 – 36 págs.

511. ENDOGENOUS COLLATERAL - Aloísio Araújo; José Fajardo Barbachan; Mario R. Páscoa – Novembro de 2003 – 37 págs.

512. NON-MONOTONE INSURANCE CONTRACTS AND THEIR EMPIRICAL CONSEQUENCES - Aloísio Araujo; Humberto Moreira – Novembro de 2003 – 31 págs.

513. EQUILIBRIA IN SECURITY MARKETS WITH A CONTINUUM OF AGENTS - A. Araujo; V. F. Martins da Rocha; P. K. Monteiro – Novembro de 2003 – 17 pág's.

514. SPECULATIVE ATTACKS ON DEBTS AND OPTIMUM CURRENCY AREA: A WELFARE ANALYSIS - Aloisio Araujo; Márcia Leon – Novembro de 2003 – 50 págs.

515. O MÉTODO GENERALIZADO DOS MOMENTOS(MGM): CONCEITOS BÁSICOS - Renato G. Flôres Jr – Novembro de 2003 – 27 págs.

516. VARIÁVEIS INTRUMENTAIS E O MGM: USO DE MOMENTOS CONDICIONAIS - Renato G. Flôres Jr – Novembro de 2003 – 27 págs.

517. O VALOR DA MOEDA E A TEORIA DOS PREÇOS DOS ATIVOS - Fernando de Holanda Barbosa – Dezembro de 2003 – 17 págs.

518. EMPRESÁRIOS NANICOS, GARANTIAS E ACESSO À CRÉDITO - Marcelo Côrtes Néri; Fabiano da Silva Giovanini - Dezembro de 2003 – 23 págs.

519. DESENHO DE UM SISTEMA DE METAS SOCIAIS - Marcelo Côrtes Néri; Marcelo Xerez - Dezembro de 2003 – 24 págs.

520. A NEW INCIDENCE ANALYSIS OF BRAZILIAN SOCIAL POLICIES USING MULTIPLE DATA SOURCES - Marcelo Côrtes Néri - Dezembro de 2003 – 55 págs.