

State University of Campinas - UNICAMP

*IMPACTO DO CRESCIMENTO ECONÔMICO  
SOBRE O USO DE ENERGIA RENOVÁVEL*

*Trabalho Final HO-235*

Supervisor: Alexandre Gori Maia

---

*João Paulo F. Fenelon () e Henri Makika (211042)*

---

*Julho 10, 2019*



## 1 INTRODUÇÃO

Este trabalho apresenta e discute os resultados de um modelo de regressão em que se analisa o impacto do crescimento sobre o uso de energia de fontes renováveis. Supõe-se que a medida que a renda das economias aumenta ocorre um processo de substituição das fontes de energia fósseis por mais limpas (renováveis). Utiliza-se na primeira parte um modelo clássico baseado no método de mínimos quadrados ordinários (OLS), de modo a estimar os resultados para o ano de 2010; em seguida, usa-se o modelo de dados em painel para avaliar um período mais longo, de 1964 a 2018. Os dados foram extraídos do world development indicators do Banco Mundial.

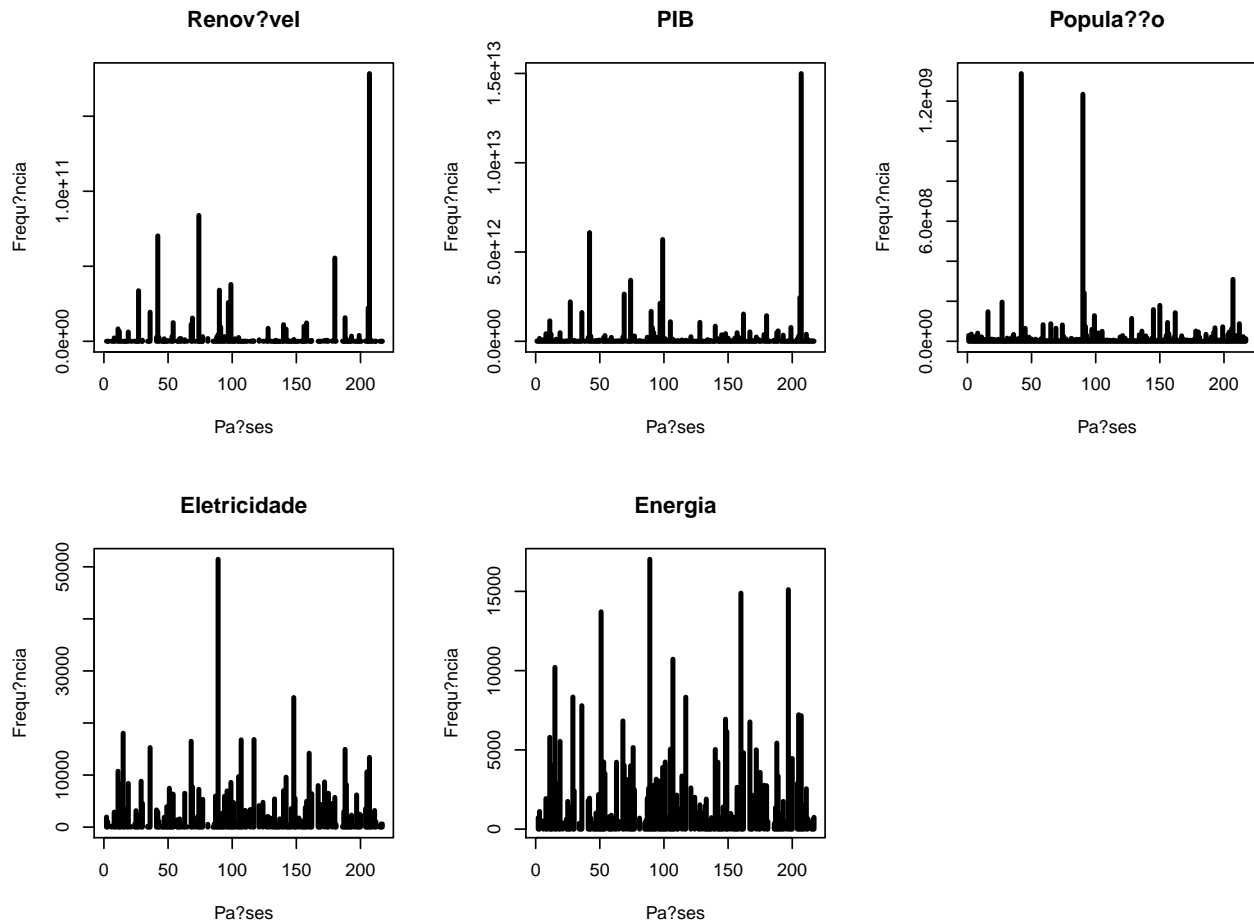
## 2 ESTIMAÇÃO E DISCUSSÃO DOS RESULTADOS

São utilizados os seguintes dados retirados do World Development Indicators do Banco Mundial:

##	renovavel	pib	populacao
##	Min. :0.000e+00	Min. :3.182e+07	Min. :1.000e+04
##	1st Qu.:0.000e+00	1st Qu.:4.895e+09	1st Qu.:6.897e+05
##	Median :1.040e+08	Median :2.027e+10	Median :5.824e+06
##	Mean :5.445e+09	Mean :3.194e+11	Mean :3.180e+07
##	3rd Qu.:1.927e+09	3rd Qu.:1.466e+11	3rd Qu.:2.053e+07
##	Max. :1.784e+11	Max. :1.499e+13	Max. :1.338e+09
##	NA's :77	NA's :12	
##	eletricidade	energia	
##	Min. : 24.52	Min. : 135.4	
##	1st Qu.: 775.95	1st Qu.: 687.3	
##	Median : 2491.63	Median : 1474.7	
##	Mean : 4259.60	Mean : 2641.6	
##	3rd Qu.: 5700.86	3rd Qu.: 3347.6	
##	Max. :51439.91	Max. :17023.2	
##	NA's :76	NA's :76	

Os dados utilizados referem-se ao ano de 2010. Há 217 observações (países). As seguintes variáveis são consideradas ao longo do trabalho:

*renovavel* = produção de energia de fontes renováveis em kWh, excluindo hidroelétricas; *pib* = produto interno bruto em US\$ a preços de 2010; *populacao* = total da população; *eletricidade* = consumo de energia elétrica em kWh per capita; *\*energia* = uso de energia em kg per capita.



Supõe-se que a produção de energia renovável (renovavel) depende positivamente das variáveis produto interno bruto (pib) e negativamente da população (populacao). Argumenta-se que países com maior PIB tendem, ao longo do tempo, substituir as fontes de energia fósseis por fontes mais limpas; países que apresentam maior população tendem a consumir maior quantidade energia fósseis. O que foi dito pode ser estimado por meio de um modelo de regressão linear baseado no método de mínimos quadrados ordinários (OLS) na forma:

$$\widehat{renovavel}_i = \hat{\alpha} + \hat{\beta}_1 pib_i + \hat{\beta}_2 populacao_i + \hat{e}_i$$

Como forma de encontrar o melhor ajuste na forma funcional do modelo, deve-se testar algumas especificações. As especificações do modelo na forma linear, linear-log, log-linear e log-log são descritas a seguir:

Dependent variable:			
	renovavel	log(renovavel + 1)	
	(1)	(2)	(3)
pib	0.0115*** (0.0004)		0.0000*** (0.0000)
populacao	4.7875 (4.0306)		0.0000 (0.0000)

```

## log(pib)                    5,188,506,697.0000***          3.0299***
##                             (1,001,320,220.0000)          (0.5329)
##
## log(populacao)              919,936,393.0000             -0.7444
##                             (1,171,945,667.0000)          (0.6237)
##
## Constant                    -135,113,932.0000  -139,634,102,491.0000***  12.7654*** -50.1408***
##                             (605,477,858.0000)  (19,017,008,819.0000)    (0.8557)   (10.1206)
##
## -----
## Observations                137                137                137                137
## R2                          0.8801                0.3044                0.0860                0.2479
## Adjusted R2                 0.8783                0.2940                0.0724                0.2367
## F Statistic (df = 2; 134)   491.9049***          29.3220***          6.3050***  22.0888***
## =====
## Note:                                                                *p<0.1; **p<0.05; ***p<0.01
##
## studentized Breusch-Pagan test
##
## data: linear
## BP = 11.479, df = 2, p-value = 0.003216
##
## studentized Breusch-Pagan test
##
## data: linear
## BP = 0.12266, df = 1, p-value = 0.7262
##
## studentized Breusch-Pagan test
##
## data: linlog
## BP = 12.845, df = 2, p-value = 0.001625
##
## studentized Breusch-Pagan test
##
## data: linlog
## BP = 0.018464, df = 1, p-value = 0.8919
##
## studentized Breusch-Pagan test
##
## data: loglin
## BP = 2.4055, df = 2, p-value = 0.3004
##
## studentized Breusch-Pagan test
##
## data: loglin
## BP = 0.7205, df = 1, p-value = 0.396
##
## studentized Breusch-Pagan test
##
## data: loglog

```

```

## BP = 7.8889, df = 2, p-value = 0.01936
##
## studentized Breusch-Pagan test
##
## data: loglog
## BP = 0.32147, df = 1, p-value = 0.5707
##
## Jarque-Bera test for normality
##
## data: linear$residuals
## JB = 3498, p-value < 2.2e-16
##
## Shapiro-Wilk normality test
##
## data: linear$residuals
## W = 0.55193, p-value < 2.2e-16
##
## Jarque-Bera test for normality
##
## data: linlog$residuals
## JB = 13019, p-value < 2.2e-16
##
## Shapiro-Wilk normality test
##
## data: linlog$residuals
## W = 0.56979, p-value < 2.2e-16
##
## Jarque-Bera test for normality
##
## data: loglin$residuals
## JB = 20.799, p-value = 0.002
##
## Shapiro-Wilk normality test
##
## data: loglin$residuals
## W = 0.74816, p-value = 4.805e-14
##
## Jarque-Bera test for normality
##
## data: loglog$residuals
## JB = 14.421, p-value = 0.0125
##
## Shapiro-Wilk normality test
##
## data: loglog$residuals
## W = 0.89399, p-value = 1.966e-08
##
## pib populacao
## 1.253587 1.253587

```

```

##      log(pib) log(populacao)
##      1.784668      1.784668

##      pib populacao
##      1.253587  1.253587

##      log(pib) log(populacao)
##      1.784668      1.784668

##
## RESET test
##
## data: linear
## RESET = 4.1673, df1 = 2, df2 = 132, p-value = 0.01758

##
## RESET test
##
## data: linlog
## RESET = 178.25, df1 = 2, df2 = 132, p-value < 2.2e-16

##
## RESET test
##
## data: loglin
## RESET = 7.6217, df1 = 2, df2 = 132, p-value = 0.0007371

##
## RESET test
##
## data: loglog
## RESET = 0.9187, df1 = 2, df2 = 132, p-value = 0.4016

##      df      AIC
## linear  4 6591.5635
## linlog  4 6832.4482
## loglin  4 1008.1595
## loglog  4  981.4439

##      df      BIC
## linear  4 6603.2434
## linlog  4 6844.1282
## loglin  4 1019.8395
## loglog  4  993.1238

```

Escolhe-se o modelo linear, uma vez que foi aquele que apresentou melhor ajuste de acordo com o  $R^2$ . Ademais, o teste de White não verificou suspeitas de heterocedasticidade em nenhum dos modelos; a hipótese de normalidade dos resíduos foi rejeitada em todos os modelos, conforme os testes Jarque-Bera e Shapiro-Wilk; e o teste geral de especificação Ramsey-Reset rejeitou a hipótese de correta especificação para todos os modelos, exceto o modelo log-log.

As variáveis relacionadas à hipótese teórica mostram que no modelo linear apenas o PIB é estatisticamente significativo a 5%. O sinal positivo do coeficiente associado ao PIB está de acordo com a hipótese levantada. O aumento de 100 dólares no PIB gera um aumento de 1,16 kWh no uso de energia renovável, ceteris paribus. A variável população, neste caso, não é estatisticamente significativa na determinação do maior uso da energia renovável. Quanto ao modelo, pode-se concluir pelo teste F que o mesmo é estatisticamente significativo a qualquer nível de significância.

Pode-se testar a hipótese de maior uso de energia renovável no Grupo das sete economias mais avançadas

(G7) em rela  o ao resto do mundo. Para isto, inclui-se uma vari  vel bin  ria em que 1 representa o uso de energia renov  vel dos pa  ses do G7 e 0 para o caso contr  rio.

```
##
## =====
##                               Dependent variable:
##                               -----
##                               renovavel
##                               Linear          Linear-Bin  ria
##                               (1)            (2)
## -----
## pib                0.0115***             0.0115***
##                   (0.0004)             (0.0006)
##
## populacao          4.7875                4.9181
##                   (4.0306)             (4.2438)
##
## binaria_G7                363,258,057.0000
##                   (3,564,800,222.0000)
##
## Constant          -135,113,932.0000       -140,929,245.0000
##                   (605,477,858.0000)     (610,399,688.0000)
##
## -----
## Observations              137              137
## R2                        0.8801            0.8801
## Adjusted R2               0.8783            0.8774
## F Statistic  491.9049*** (df = 2; 134) 325.5182*** (df = 3; 133)
## =====
## Note:                      *p<0.1; **p<0.05; ***p<0.01
```

Os pa  ses do G7 usam, em m  dia, mais de 3,5 milh  es de kwh de energia renov  vel a mais que os demais pa  ses. No entanto, a hip  tese n  o    validada, pois a nula n  o    rejeitada; isto   , a vari  vel bin  ria    estatisticamente insignificante. Logo, mant  m-se com a especifica  o do modelo linear.

Feito isso, verifica-se a possibilidade de endogeneidade na vari  vel pib. O problema de endogeneidade pode ocorrer devido a tr  s fatores: (1) omiss  o de vari  veis relevantes correlacionadas ao pib; (2) erros de medi  o com o pib (e.g. uma proxy mal especificada); (3) simultaneidade entre y (renovavel) e uma ou mais vari  veis explicativas (pib e populacao). Para este trabalho, testa-se o caso (1), em que se considera existir uma vari  vel de controle associada ao PIB e que n  o deve ser omitida do modelo, uma vez que se teria um modelo com vi  s de especifica  o. Sup  e-se dois modelos: o primeiro possui, supostamente, vi  s ao omitir vari  veis; o segundo considera vari  veis (proxies) de controle associadas    demanda por energia, as quais ajudam a explicar o uso de energia renov  vel e est  o associadas ao PIB. Estas vari  veis s  o o uso de energia per capita e o consumo de eletricidade per capita.

```
##
## =====
##                               Dependent variable:
##                               -----
##                               renovavel
##                               Linear          Linear com Controles
##                               (1)            (2)
## -----
## pib                0.0115***             0.0115***
##                   (0.0004)             (0.0004)
##
```

```
## populacao          4.7875          5.0456
##                   (4.0306)          (4.1097)
##
## energia            -287,172.8000
##                   (356,869.4000)
##
## eletricidade       178,950.9000
##                   (175,439.3000)
##
## Constant          -135,113,932.0000      -154,382,459.0000
##                   (605,477,858.0000)      (788,217,963.0000)
##
## -----
## Observations              137              137
## R2                        0.8801              0.8811
## Adjusted R2               0.8783              0.8775
## Residual Std. Error 6,662,336,361.0000 (df = 134) 6,686,198,842.0000 (df = 132)
## F Statistic           491.9049*** (df = 2; 134)    244.4613*** (df = 4; 132)
## =====
## Note:                                *p<0.1; **p<0.05; ***p<0.01
```

Uma vez que a inclus?o das vari?veis de controle n?o melhorou a signific?ncia dos estimadores ou o valor do  $R^2$ , mant?m-se o modelo original.

Uma desvantagem da an?lise com dados em corte transversal s?o os limitados graus de liberdade. Isso significa que ? poss?vel incluir apenas um conjunto restrito de vari?veis no modelo de regress?o; caso contr?rio, reduz-se sua efici?ncia: a vari?ncia tende a se tornar muito elevada e os testes de hip?tese inv?lidos. Por outro lado, omitir vari?veis pode gerar estimativas tendenciosas. A principal forma de enfrentar este trade-off entre efici?ncia e vi?s ? aumentar o n?mero de observa??es (n). Para um mesmo n?mero de ind?viduos (pa?ses) a solu??o passa por incluir mais per?odos. Para o caso analisado, ? poss?vel estender o per?odo da amostra para um intervalo maior que 2010. A base do WDI do Banco Mundial permite analisar os per?odos de 1960 a 2018. Para lidar com esses dados alguns modelos s?o considerados.

Uma alternativa ? usar o m?todo OLS com dados empilhados. Os resultados s?o descritos a seguir comparando-os ao modelo com dados em corte transversal.

```
## # A tibble: 6 x 9
##   ano tempoano pais  codigo renovavel      pib populacao eletricidade
##   <dbl> <chr>   <chr> <chr>      <dbl>      <dbl>      <dbl>      <dbl>
## 1  1960 YR1960 Afgh~ AFG          NA NA          8996973          NA
## 2  1960 YR1960 Alba~ ALB          NA NA          1608800          NA
## 3  1960 YR1960 Alge~ DZA          NA 2.74e10  11057863          NA
## 4  1960 YR1960 Amer~ ASM          NA NA           20123          NA
## 5  1960 YR1960 Ando~ AND          NA NA           13411          NA
## 6  1960 YR1960 Ango~ AGO          NA NA          5454933          NA
## # ... with 1 more variable: energia <dbl>

##
## =====
##                               Dependent variable:
##   -----
##                               renovavel
##   Transversal      Empilhado(1)
##   (1)              (2)
##   -----
## pib                0.0115***      0.0085***
```



```

##          (0.0004)          (0.0001)
##
## populacao          4.7875          4.5036***
##          (4.0306)          (0.9423)
##
## Constant          -135,113,932.0000          -891,335,566.0000***
##          (605,477,858.0000)          (131,251,620.0000)
##
## -----
## Observations          137          5,447
## R2          0.8801          0.5631
## Adjusted R2          0.8783          0.5629
## F Statistic  491.9049*** (df = 2; 134) 3,508.2790*** (df = 2; 5444)
## =====
## Note:          *p<0.1; **p<0.05; ***p<0.01
##
## =====
##                               Dependent variable:
##                               -----
##                               renovavel
##                               Empilhado(2)          Empilhado(3)
##                               (1)          (2)
## -----
## pib          0.0084***          -0.9413***
##          (0.0001)          (0.0138)
##
## populacao          4.4111***          2.7253***
##          (0.9254)          (0.6658)
##
## periodoano          138,072,157.0000***          -21,375,901.0000***
##          (10,354,291.0000)          (7,789,194.0000)
##
## periodoperiodo          -1,430,672,511.0000***          285,493,622.0000
##          (548,972,966.0000)          (428,549,807.0000)
##
## int_pibperiodoano          0.0005***
##          (0.00001)
##
## int_pibperiodoperiodo          -0.0078***
##          (0.0005)
##
## Constant          -274,834,424,559.0000***          42,195,236,554.0000***
##          (20,378,191,495.0000)          (15,333,904,330.0000)
##
## -----
## Observations          5,447          5,447
## R2          0.5790          0.7825
## Adjusted R2          0.5787          0.7823
## Residual Std. Error  8,896,506,732.0000 (df = 5442) 6,395,404,772.0000 (df = 5440)
## F Statistic          1,871.1040*** (df = 4; 5442) 3,262.3140*** (df = 6; 5440)
## =====
## Note:          *p<0.1; **p<0.05; ***p<0.01

```

Nota-se que s?o descritas tr?s especifica??es para o modelo com dados empilhados: (i) intercepto e coeficientes

angulares constantes; (ii) intercepto diferente e coeficientes angulares constantes; e (iii) intercepto e coeficientes angulares diferentes. Assim, permite-se identificar mudan?as na rela??o entre o crescimento econ?mico e a produ??o de energia renov?vel ao longo do tempo. Embora o  $R^2$  n?o aumente nos modelos com dados empilhados, a vari?vel popula??o passa a ser estatisticamente significativa.

Alternativamente, pode-se estar interessado em controlar os efeitos de vari?veis omitidas no modelo e que se sup?e relevantes. Assim, ? poss?vel usar estimadores que permitem controlar os efeitos dessas vari?veis omitidas sobre as demais. Esse controle pode ser feito por meio da cria??o de bin?rias para cada pa?s da amostra menos um. Este ? m?todo ? feito por meio do estimador de efeitos fixos. O problema disso ? que se gastam muitos graus de liberdade e, conseq?entemente, reduz-se a efici?ncia do modelo. Outro m?todo ? tratar os efeitos das vari?veis omitidas como valores aleat?rios; isto ?, por meio do estimador de efeitos aleat?rios incorpora-se o efeito das vari?veis omitidas (as quais s?o tratadas com uso de bin?rias no estimador de efeitos fixos) dentro do erro de previs?o. ? uma alternativa v?lida quando n?o se deseja reduzir a efici?ncia do modelo. No entanto, o uso do estimador de efeitos aleat?rios requer que a correla??o entre as vari?veis explicativas e o erro de previs?o seja igual ou pr?xima a zero.

Os resultados estimados para o modelo linear de corte transversal, modelo com dados empilhados e os modelos com estimadores fixos e aleat?rios s?o descritos a seguir.

```
## Warning: use of 'plm.data' is discouraged, better use 'pdata.frame' instead

##
## F test for individual effects
##
## data: renovavel ~ log(pib) + log(populacao)
## F = -13.194, df1 = 137, df2 = 5307, p-value = 1
## alternative hypothesis: significant effects

##
## =====
##                                     Dependent variable:
##                                     -----
##                                     renovavel
##                                     OLS          panel
##                                     Empilhado    linear
##                                     (1)          Fixos-Oneway
##                                     (2)
## -----
## pib                                0.0085***
##                                   (0.0001)
##
## populacao                          4.5036***
##                                   (0.9423)
##
## log(pib)                                8,921,010,678.0000*
##                                   (485,023,641.0000)
##
## log(populacao)                       -7,916,785,859.0000
##                                   (970,420,026.0000)
##
## cbind(log(pib), log(populacao), as.factor(pais))1
##
## cbind(log(pib), log(populacao), as.factor(pais))2
##
```

```

## Constant                -891,335,566.0000***
##                          (131,251,620.0000)
## -----
## Observations                5,447                5,447
## R2                        0.5631                0.0736
## Adjusted R2                0.5629                0.0493
## F Statistic                3,508.2790*** (df = 2; 5444) 210.7651*** (df = 2; 5444)
## =====
## Note:
##
## F test for individual effects
##
## data: renovavel ~ log(pib) + log(populacao)
## F = -13.194, df1 = 137, df2 = 5307, p-value = 1
## alternative hypothesis: significant effects
##
## =====
##                               Dependent variable:
##                               -----
##                               (renovavel)          cbind(renovavel)
##                               Aleat?rios-Oneway      Aleat?rios-Twoway
##                               (1)                    (2)
## -----
## log(pib)                4,577,517,890.0000***
##                          (290,307,828.0000)
##
## log(populacao)          -1,510,224,786.0000***
##                          (401,222,476.0000)
##
## cbind(log(pib), log(populacao), as.factor(ano))1      2,705,342,490.0000***
##                                                         (330,552,446.0000)
##
## cbind(log(pib), log(populacao), as.factor(ano))2      -1,364,844,186.0000***
##                                                         (402,790,911.0000)
##
## cbind(log(pib), log(populacao), as.factor(ano))3      175,656,976.0000***
##                                                         (14,437,704.0000)
##
## Constant                -86,298,066,517.0000***      -49,005,568,320.0000***
##                          (5,624,562,766.0000)          (6,424,800,674.0000)
## -----
## Observations                5,447                5,447
## R2                        0.0566                0.0817
## Adjusted R2                0.0563                0.0812
## F Statistic                163.2555*** (df = 2; 5444) 161.3444*** (df = 3; 5444)
## =====
## Note:                                                         *p<0.1; **p<0.05; ***p<0.01
##
## Hausman Test
##

```

```
## data: renovavel ~ log(pib) + log(populacao)
## chisq = 136.43, df = 2, p-value < 2.2e-16
## alternative hypothesis: one model is inconsistent
```

O modelo com estimador de efeitos fixos por país (oneway) indica que a energia de fontes renováveis aumenta 0,017 kWh, em média por país, quando o PIB aumenta em uma unidade monetária, ceteris paribus. Mais uma vez, a população não foi estatisticamente significativa. Uma extensão desse modelo é modelo twoway. Neste, controla-se, além dos efeitos entre os países, o efeito entre os períodos (anos). Assim, o aumento em uma unidade monetária do PIB gera um aumento médio, por país e ano, de 0,017 kWh na produção de energia renovável. O teste F em relação aos efeitos por país (e por país e ano) rejeita a nula em que afirma ser o modelo OLS melhor que o modelo com estimador de efeitos fixos.

O modelo com estimador aleatório por país (oneway) Multiplicador de Lagrange de Breusch-Pagan para efeitos aleatórios:

```
##
## Lagrange Multiplier Test - (Breusch-Pagan) for unbalanced panels
##
## data: renovavel ~ pib + populacao
## chisq = 2027.9, df = 1, p-value < 2.2e-16
## alternative hypothesis: significant effects
```

Teste de dependência transversal:

```
## Warning in pcdres(tres = tres, n = n, w = w, form =
## paste(deparse(x$formula)), : Some pairs of individuals (0.021 percent) do
## not have any or just one time period in common and have been omitted from
## calculation
##
## Breusch-Pagan LM test for cross-sectional dependence in panels
##
## data: renovavel ~ log(pib) + log(populacao)
## chisq = 130850, df = 9452, p-value < 2.2e-16
## alternative hypothesis: cross-sectional dependence
## Warning in pcdres(tres = tres, n = n, w = w, form =
## paste(deparse(x$formula)), : Some pairs of individuals (0.021 percent) do
## not have any or just one time period in common and have been omitted from
## calculation
##
## Pesaran CD test for cross-sectional dependence in panels
##
## data: renovavel ~ log(pib) + log(populacao)
## z = 96.422, p-value < 2.2e-16
## alternative hypothesis: cross-sectional dependence
```

Teste de correlação serial:

```
##
## Breusch-Pagan test
##
## data: renovavel ~ pib + populacao + factor(pais)
## BP = 111230, df = 139, p-value < 2.2e-16
```

Teste de Heterocedasticidade:

```
##
## Breusch-Pagan test
```

```
##
## data: renovavel ~ pib + populacao + factor(pais)
## BP = 111230, df = 139, p-value < 2.2e-16

Controlando a heteroscedasticidade:

##
## t test of coefficients:
##
##           Estimate Std. Error t value Pr(>|t|)
## log(pib)      8921010678   485023641 18.3929 < 2.2e-16 ***
## log(populacao) -7916785860   970420026 -8.1581 4.212e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##
## t test of coefficients:
##
##           Estimate Std. Error t value Pr(>|t|)
## log(pib)      8921010678   3081164906  2.8953 0.003803 **
## log(populacao) -7916785860   3731830630 -2.1214 0.033933 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##
## =====
##                               Dependent variable:
##                               -----
##                               renovavel
##                               panel
##                               linear
##                               coefficient
##                               test
##                               Fixos - Oneway
##                               Fixos-Oneway-Consistente
##                               (1)
##                               (2)
## -----
## log(pib)      8,921,010,678.0000***      8,921,010,678.0000***
##              (485,023,641.0000)          (3,081,164,906.0000)
##
## log(populacao) -7,916,785,859.0000***      -7,916,785,859.0000**
##              (970,420,026.0000)          (3,731,830,630.0000)
##
## -----
## Observations      5,447
## R2                 0.0736
## Adjusted R2        0.0493
## F Statistic      210.7651*** (df = 2; 5307)
## =====
## Note:                                     *p<0.1; **p<0.05; ***p<0.01
```