

Identificação dos Fatores Relacionados à Exportação de Açúcar no Brasil

Alexandre, Carina e Eduardo

7 de dezembro de 2016

1 - Resumo:

O Objetivo desse breve trabalho é analisar os fatores econômicos relacionados com as exportações de açúcar no Brasil. O trabalho analisa dados econômicos, referentes aos anos de 1974 a 2010, obtidos no site do Instituto de Pesquisa Econômica Aplicada (IPEA), para analisar quais fatores influenciam ou não às exportações brasileiras de açúcar. Para isso foi estimado um modelo de regressão do tipo “Mínimos Quadrados Generalizados” (MQG) e foi constatado que as exportações de açúcar possuem uma relação positiva com a área colhida de cana de açúcar e uma relação negativa com a taxa básica de juros da Economia, a taxa Selic.

2 - Dados Utilizados no Trabalho e Estatísticas Descritivas:

No site do IPEA (www.ipeadata.gov.br) foram obtidos os seguintes dados, referentes aos anos de 1974 a 2010:

- Volume de açúcar exportado (em toneladas)
- Valor das exportações de açúcar (em milhares de US\$)
- Área colhida de cana de açúcar (em Hectares)
- Produção de cana de açúcar (em toneladas)
- Taxa Selic no ano (em % a.a.)

A partir dos valores das exportações de açúcar, foi calculado o preço médio da tonelada de açúcar no ano no mercado internacional.

Sendo assim, a variável dependente dos modelos de regressão testados neste trabalho será o volume de açúcar exportado (em toneladas) e as variáveis independentes a serem analisadas serão: 1) área colhida de cana de açúcar (em Ha), 2) A produção de cana de açúcar no ano (em ton), 3) a taxa Selic (em % a.a.) e 4) o preço médio da tonelada de açúcar no mercado internacional.

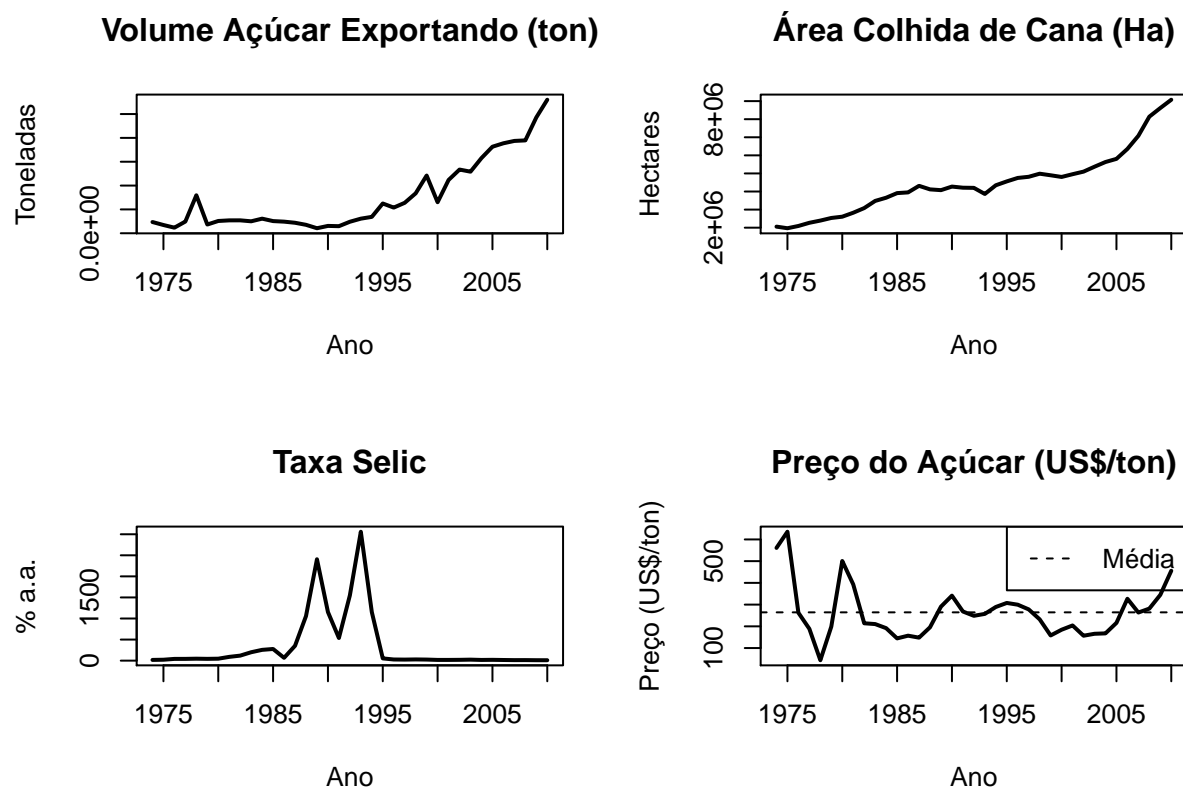
A seguir são apresentadas as principais estatísticas descritivas de cada variável analisada aqui:

Variável	Média \bar{X}	Desvio Padrão σ	Coefficiente de Variação $\left(\frac{\sigma}{\bar{X}} \times 100\right)$
VolExpAcucar_ton	7.55×10^6	7.39×10^6	97.84%
AreaColhidaCana	4.46×10^6	1.74×10^6	39.05%
ProdCana_ton	3.02×10^8	1.57×10^8	51.96%
TaxaSELIC	347.73	697.23	200.51%
PrecoAcucar	264.38	120.8	45.69%

A tabela acima mostra que a variável com menor dispersão (isto é, apresentou o menor coeficiente de variação)

foi **AreaColhidaCana** e a com maior dispersão foi **TaxaSelic**. O motivo para a alta dispersão observada para a variável **TaxaSelic** é o período de hiperinflação que culminou principalmente nos anos de 1981 a 1994.

A Seguir é apresentado um painel com 4 gráficos, apresentando a evolução das exportações de açúcar (em toneladas), a área colhida de cana de açúcar (em Ha), a taxa Selic e o preço médio da tonelada de açúcar no mercado internacional.



Nos painel de gráficos acima são apresentadas as séries temporais com a evolução das variáveis analisadas ao longo do período de 1974 a 2010. Nesse painel deve-se destacar os seguintes fatos:

- Tendência crescente do volume de açúcar exportado ao longo do período analisado.
- Os altos níveis da taxa Selic entre a segunda metade dos anos 1980 e a primeira metade dos anos 1990.
- As grande volatilidade (variabilidade) do preço do açúcar, no mercado internacional, ao longo do período.

A seguir, são apresentadas as relações lineares individuais de cada variável independente com a variável dependente **VolExpAcucar_ton**.

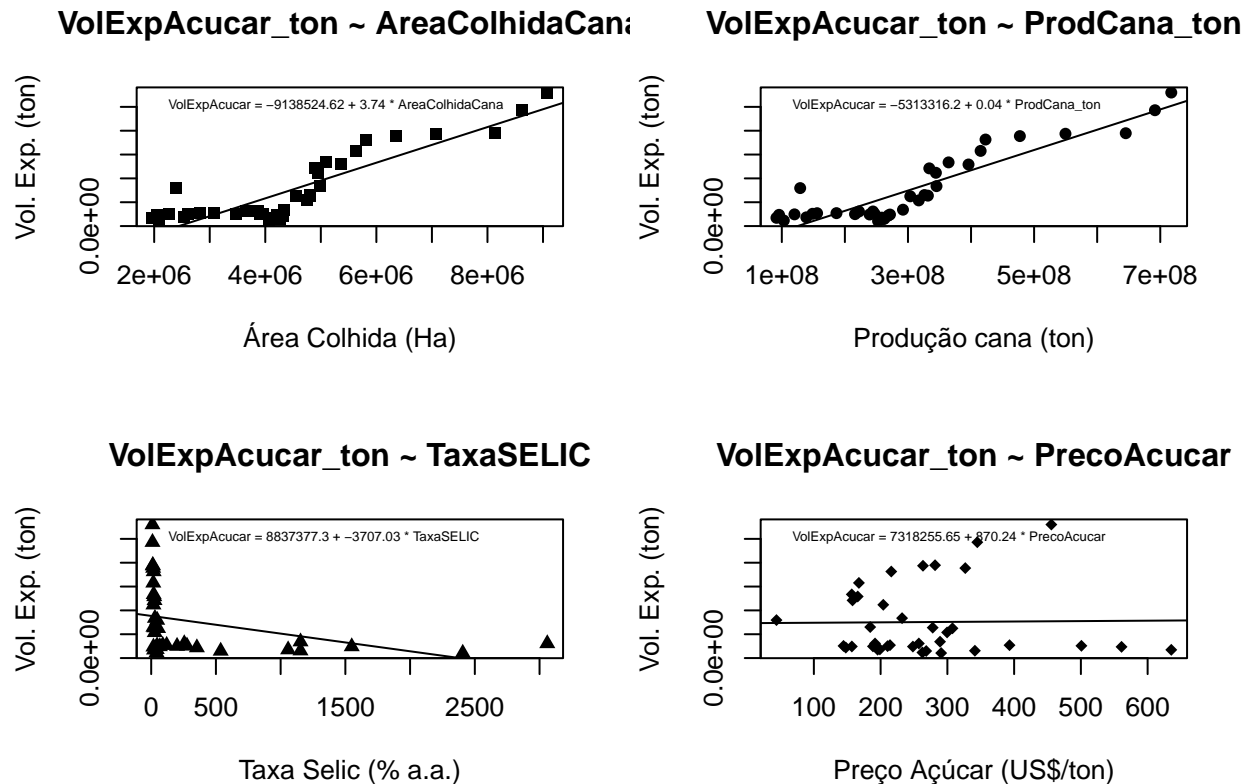
3 - Relações Lineares Entre as Variáveis:

A seguir, é apresentada uma tabela com as correlações entre cada uma das variáveis independentes com a variável dependente **VolExpAcucar_ton**:

Variável	$\rho_{VolExpAcucar_{ton}, Varivel}$
AreaColhidaCana	0.88
ProdCana_ton	0.9
TaxaSELIC	-0.35
PrecoAcucar	0.01

Na tabela acima, pode-se constatar correlações fortíssimas entre o volume exportado de açúcar (VolExpAcucar_ton) e as variáveis referentes à produção de cana: AreaColhidaCana e ProdCana_ton.

A seguir é apresentado um painel com gráficos de dispersão, com os valores de cada uma das variáveis independentes no eixo-X e os valores da variável dependente no eixo Y.



No painel acima, pode-se constatar que o volume de açúcar exportado no ano possui uma relação positiva com a área colhida de cana de açúcar no ano e com a produção de cana no ano. O motivo é bastante óbvio, uma vez que para poder exportar açúcar, é necessário produzir a sua matéria prima, a cana de açúcar.

Em relação à taxa Selic, esta possui uma relação negativa com o volume de açúcar exportado, uma vez que incrementos na taxa básica de juros da economia aumenta o custo de oportunidade do produtor, fazendo este preferir investir o seu dinheiro em títulos públicos do que na sua produção. Isto é, uma alta nas taxas de juros da economia fazem com que:

- Os produtores de cana de açúcar prefiram investir o seu dinheiro em ativos financeiros, uma vez que é possível obter maiores lucros com investimentos financeiros do que com a venda da cana plantada.

- Os produtores de açúcar produzam menos açúcar, tanto por causa da diminuição na oferta da matéria prima, quanto por causa da possibilidade de obtenção de maiores retornos em investimentos financeiros.

No caso do preço médio do açúcar no mercado internacional, o segundo gráfico da parte inferior do painel acima sugere que o preço do açúcar no mercado internacional possui de pouca a nenhuma relação linear com o volume de açúcar exportado no ano. Uma possível explicação para essa baixa relação é que o mercado de açúcar, como um todo, pode ser definido em termos econômicos como um mercado de “**concorrência perfeita**” e, conseqüentemente, os produtores de açúcar são tomadores de preço. Se esse for realmente o caso, pode-se afirmar que os preços do açúcar no mercado internacional não irão influenciar o volume de açúcar exportado no país.

Uma possível segunda explicação, é que os produtores de açúcar também podem vender antecipadamente a sua produção por meio de instrumentos financeiros de *hedge*. Nesse caso, o produtor poderia vender antecipadamente a sua produção e contrair uma obrigação de venda por um determinado preço, o comprador dessas opções receberia uma opção de compra do açúcar pelo mesmo preço e, no caso do preço do açúcar ser maior do que o preço dessas opções, ele terá o direito de comprar o açúcar pelo preço das opções.

4 - Seleção das Variáveis Independentes do modelo:

Um pressuposto importante do modelo de regressão linear múltipla é a ausência de multicolinearidade, isto é, não pode haver correlação entre as variáveis independentes do modelo de regressão. A seguir é apresentada uma matriz de correlação entre as variáveis independentes apresentadas acima:

##	AreaColhidaCana	ProdCana_ton	TaxaSELIC	PrecoAcucar
## AreaColhidaCana	1.000	0.998	-0.119	-0.033
## ProdCana_ton	0.998	1.000	-0.148	-0.018
## TaxaSELIC	-0.119	-0.148	1.000	-0.013
## PrecoAcucar	-0.033	-0.018	-0.013	1.000

A matriz de correlações acima aponta que seria imprudente a inclusão simultânea das variáveis **AreaColhidaCana** e **ProdCana_ton** no modelo de regressão uma vez que a correlação entre as duas variáveis é quase que perfeita, 1.

Para identificar qual das duas será incluída no modelo, deve-se realizar o teste VIF (Variance InFlator) para identificar qual das duas “infla” mais a variância dos resíduos do modelo e, então, retirar essa variável do modelo.

##	AreaColhidaCana	ProdCana_ton	TaxaSELIC	PrecoAcucar
##	254.9	256.7	1.2	1.1

Com os resultados apresentados acima, optou-se por **retirar do modelo a variável “ProdCana_ton”** por apresentar o maior VIF.

5 - Primeiro Modelo de Regressão:

Após a retirada de uma variável independente do modelo que estava fortemente correlacionada com uma outra variável independente, será executado o primeiro modelo de regressão:

$$VolExpAcucar_{ton} = \beta_0 + \beta_1 AreaColhidaCana_t + \beta_2 TaxaSELIC_t + \beta_3 PrecoAcucar_t + \epsilon_t$$

A seguir, os resultados para o modelo de regressão estimado pela técnica de Mínimos Quadrados Ordinários (MQO):

```
##
## Call:
```

```
## lm(formula = VolExpAcucar_ton ~ AreaColhidaCana + TaxaSELIC +
##     PrecoAcucar, data = dados)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4660434 -2498039  249696  1937149  7655897
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -8.33e+06   1.88e+06   -4.43  9.7e-05 ***
## AreaColhidaCana  3.62e+00   2.97e-01   12.20  8.9e-14 ***
## TaxaSELIC      -2.62e+03   7.41e+02   -3.54  0.0012 **
## PrecoAcucar     2.42e+03   4.25e+03    0.57  0.5736
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3080000 on 33 degrees of freedom
## Multiple R-squared:  0.841, Adjusted R-squared:  0.826
## F-statistic: 58.1 on 3 and 33 DF, p-value: 2.95e-13
```

O output acima sugere que as variáveis “AreaColhidaCana” e “TaxaSELIC” contribuem para explicar o volume de açúcar exportado pelo país, enquanto que a variável “PrecoAcucar”, que representa o preço médio do açúcar exportado para o mercado internacional, não contribui para explicar o volume de açúcar exportado. Dado os resultados acima, a próxima seção irá reestimar o modelo de regressão, pelo método de MQO, utilizando apenas as variáveis independentes que foram significativas nesse primeiro modelo de regressão.

6 - Segundo Modelo de regressão

Estimando novamente o modelo de regressão por MQO, **utilizando como variáveis independentes apenas as variáveis “AreaColhidaCana” e “TaxaSELIC”**, foram gerados os seguintes resultados:

```
##
## Call:
## lm(formula = VolExpAcucar_ton ~ AreaColhidaCana + TaxaSELIC,
##     data = dados)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4796071 -2425862  391802  2135340  7108943
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -7.66e+06   1.45e+06   -5.27  7.6e-06 ***
## AreaColhidaCana  3.61e+00   2.93e-01   12.31  4.4e-14 ***
## TaxaSELIC      -2.63e+03   7.34e+02   -3.59  0.001 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3e+06 on 34 degrees of freedom
## Multiple R-squared:  0.839, Adjusted R-squared:  0.83
## F-statistic: 88.7 on 2 and 34 DF, p-value: 3.21e-14
```

Os resultados acima corroboram com os resultados previstos na seção 3 (3 - Relações Lineares Entre as variáveis) acima e, portanto, o modelo de regressão final é dado por:

$$VolExpAcucar_{t,on} = \beta_0 + \beta_1 AreaColhidaCana_t + \beta_2 TaxaSELIC_t + \epsilon_t$$

7 - Verificação dos Pressupostos Acerca dos resíduos:

Com o modelo de regressão final em mãos, essa seção tem o intuito de verificar se os três pressupostos acerca dos resíduos são atendidos. Os três pressupostos a serem testados são:

- 1- Os resíduos do Modelo possuem distribuição normal
- 2- Homocedasticidade dos resíduos, isto é, a variância dos resíduos é homogênea.
- 3- Os resíduos são independentes, isto é, não há correlação serial entre os resíduos da regressão.

7.1 - Normalidade dos resíduos:

A seguir é apresentado o resultado do teste de normalidade de Shapiro-Wilk para os resíduos do modelo de regressão. A hipótese nula e alternativa a serem testadas a seguir são:

$$H_0 : \text{Resduos so normalmente distribudos}$$

$$H_1 : \text{Resduos no so normalmente distribudos}$$

```
##
## Shapiro-Wilk normality test
##
## data:  modReg2$residuals
## W = 1, p-value = 0.4
```

O resultado acima indica que, a um nível de significância de 5%, não se pode rejeitar a hipótese nula de que os erros são normalmente distribuídos e, portanto, a regressão atende ao pressuposto de normalidade dos resíduos.

7.2 - Homogeneidade de Variâncias dos Resíduos (Homocedasticidade dos Resíduos)

A seguir é apresentado os resultados do teste de homogeneidade das variâncias dos resíduos da regressão. Nesse teste, os resíduos são divididos em dois grupos e, então, é testada a seguinte hipótese:

$$H_0 : \sigma_1 = \sigma_2$$

$$H_1 : \sigma_1 \neq \sigma_2$$

```
##
## F test to compare two variances
##
## data:  varParte1 and varParte2
## F = 0.6, num df = 20, denom df = 20, p-value = 0.4
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.25 1.70
## sample estimates:
## ratio of variances
##                0.64
```

O resultado do teste F, acima, mostra que também não se deve rejeitar, a um nível de 5% de confiança, a hipótese nula de que a variância dos resíduos dos dois grupos de resíduos sejam estatisticamente iguais e, portanto, o pressuposto de homogeneidade dos resíduos também é atendido.

7.3 - Teste de Independência dos Resíduos:

A seguir é apresentado os resultados do teste de Durbin-Watson acerca da independência dos resíduos. A hipótese a ser testada aqui será:

$$H_0 : \rho_{i,j} = 0$$

$$H_1 : \rho_{i,j} \neq 0$$

```
##
## Durbin-Watson test
##
## data: modReg2
## DW = 1, p-value = 9e-05
## alternative hypothesis: true autocorrelation is greater than 0
```

O resultado acima indica que deve-se rejeitar a hipótese nula de independência dos resíduos da regressão e, portanto, o modelo de regressão não atende ao pressuposto de independência dos resíduos e deve-se fazer alguma correção para que o modelo seja válido.

8 - Estimando o Modelo pelo Método dos Mínimos Quadrados Generalizados:

Uma vez que o pressuposto de independência dos resíduos não foi atendido, deve-se **utilizar um modelo de regressão com estimativa de erros autoregressivos de primeira ordem (AR1)**. Isto é, deve-se estimar a regressão pelo método de Mínimos Quadrados Generalizados (MQG). A seguir são apresentados os resultados da estimação do modelo de regressão pelo método de MQG, e as próximas duas subseções irão verificar os pressupostos de normalidade dos resíduos e homocedasticidade dos resíduos.

```
## Generalized least squares fit by REML
## Model: VolExpAcucar_ton ~ AreaColhidaCana + TaxaSELIC
## Data: dados
## AIC BIC logLik
## 1174 1182 -582
##
## Correlation Structure: AR(1)
## Formula: ~1 | Ano
## Parameter estimate(s):
## Phi
## 0
##
## Coefficients:
## Value Std.Error t-value p-value
## (Intercept) -7664010 1453239 -5.3 0.000
## AreaColhidaCana 4 0 12.3 0.000
## TaxaSELIC -2631 734 -3.6 0.001
##
## Correlation:
## (Intr) ArClhC
```

```
## AreaColhidaCana -0.92
## TaxaSELIC      -0.28  0.12
##
## Standardized residuals:
##   Min    Q1   Med    Q3   Max
## -1.57 -0.80  0.13  0.70  2.33
##
## Residual standard error: 3e+06
## Degrees of freedom: 37 total; 34 residual
```

Os resultados acima confirmam os resultados encontrados na estimação por MQO e indicam que existe uma relação positiva entre a área colhida de cana de açúcar e o volume exportado de açúcar e uma relação negativa entre o volume de açúcar exportado e a taxa Selic.

8.1 - Pressuposto de Normalidade dos resíduos

```
##
## Shapiro-Wilk normality test
##
## data:  modRegressao$residuals
## W = 1, p-value = 0.4
```

O resultado acima indica que o pressuposto de que os resíduos são normalmente distribuídos continua sendo atendido.

8.2 - Homogeneidade das Variâncias dos resíduos:

```
##
## F test to compare two variances
##
## data:  varParte1 and varParte2
## F = 0.6, num df = 20, denom df = 20, p-value = 0.4
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.25 1.70
## sample estimates:
## ratio of variances
##                0.64
```

Os resultados acima indicam que o pressuposto de homocedasticidade das variâncias dos resíduos continua sendo atendido.

9 - Conclusão:

Através desse trabalho, buscou-se identificar os principais fatores capazes de influenciar o volume (em toneladas) de açúcar exportado pelo Brasil. Nas análises realizadas foi identificado que os principais fatores que contribuem com o volume exportado de açúcar foram a produção de cana de açúcar e a taxa Selic.

O fator produção de cana de açúcar possui uma relação positiva com o volume de açúcar exportado e se refere a oferta de açúcar, isto é, quanto maior for a quantidade de cana plantada, maior será a oferta de açúcar e, conseqüentemente, maior será a quantidade de açúcar exportado.

O fator taxa Selic representa o tradeoff entre investir em ativos financeiros e na produção. Desse modo, a relação entre taxa Selic e volume de açúcar exportado é negativa, visto que taxas de juros altas tornam os investimentos financeiros mais atraentes que, por sua vez, aumentam os custos dos empréstimos na economia. Esses, conseqüentemente, fazem com que os produtores decidam plantar menos cana e, por fim, diminui a oferta de açúcar na economia.