

Trabalho Final

Econometria I - EAE0324

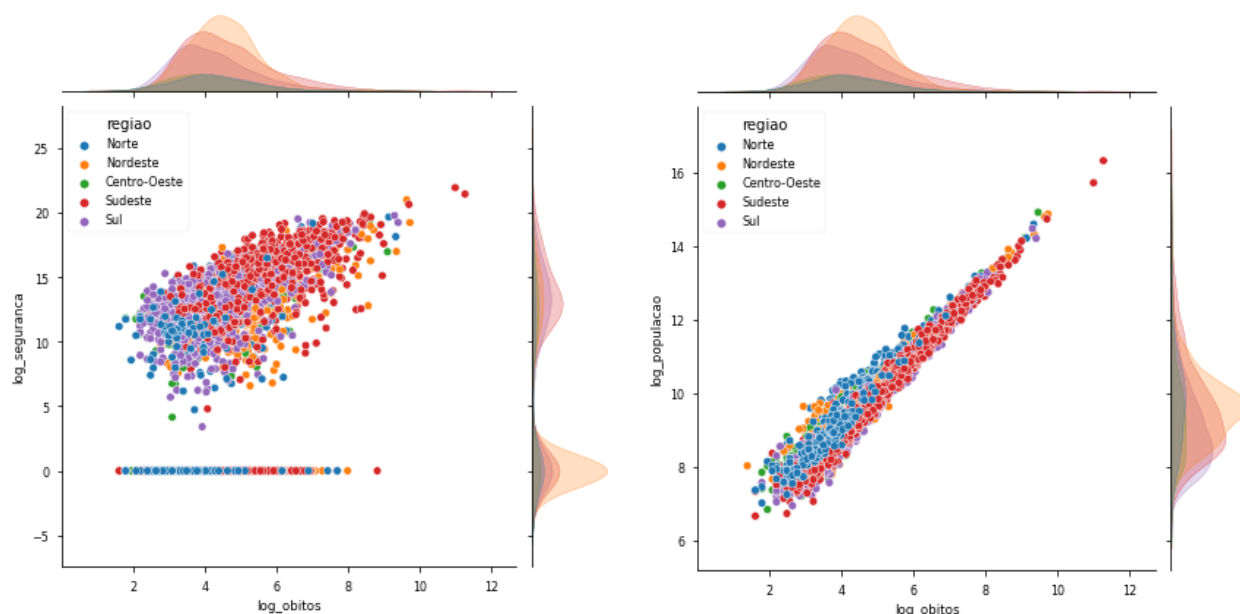
VITOR BALDOINO - 11766857

Julho 2022

Este trabalho visa entender como algumas variáveis econômicas impactam o número de óbitos de cidades brasileiras. Para tal, é utilizada uma *cross-section* de municípios que contém a variável explicada, o número de óbitos, e possíveis variáveis explicativas, como as despesas orçamentárias do município, população total, região geográfica e indicadores educacionais. Esse trabalho foi inspirado pelo artigo *US mortality by economic, demographic, and social characteristics: the National Longitudinal Mortality Study* de Sorlie, Backlund e Keller, publicado no *American Journal of Public Health* em 1995.

Descrição dos Dados

A base utilizada neste trabalho é de elaboração própria, as bases brutas usadas para aquisição dos dados podem ser obtidas através do portal [Base dos Dados](#) (a lista completa das bases utilizadas se encontra nas referências). Adicionalmente, os *scripts* utilizados para o tratamento, agregação e modelagem dos dados podem ser encontrados em [Perfil de Vitor Balduino - GitHub](#).



Como apresentado no gráfico da esquerda, os municípios de todas as regiões apresentam uma associação positiva entre o \log do número de óbitos e o \log da despesa total em segurança, isto é contra-intuitivo, visto que se espera que o gráfico apresente relação negativa. Essa contradição será melhor explorada nos modelos de regressão.

No gráfico da direita, está a associação entre os \log s do número de óbitos e da população. Um fato interessante no gráfico de distribuição superior é que o \log do número de óbitos da região nordeste, comparado às outras regiões, concentra-se em um nível ligeiramente maior e possui menor variância.

Modelos de Regressão

A abordagem para estimar o impacto de variáveis econômicas no número de óbitos de um município consiste em estimar três modelos:

$$\log(\text{mortes}) = \beta_0 + \beta_1 \log(\text{Pop}) + \beta_2 \text{Emprego} + \beta_3 \log(\text{Despesa}) + \beta_4 \text{Evasão-EM} + \beta_5 \text{Nordeste} + u_i$$

$$\log(\text{mortes}) = \beta_0 + \beta_1 \log(\text{Pop}) + \beta_2 \text{Emprego} + \beta_3 \log(\text{Seg}) + \beta_4 \text{Evasão-EM} + \beta_5 \text{Nordeste} + u_i$$

$$\log(\text{mortes}) = \beta_0 + \beta_1 \log(\text{Pop}) + \beta_2 \text{Emprego} + \beta_{i,i=\{3,4,5\}} \log(\text{Despesas}) + \beta_{i,i=\{6,7\}} \text{Educação} + \beta_8 \text{Nordeste} + u_i$$

Nestes modelos, as variáveis “log(Pop)”, “Emprego”, “Evasão-EM” representam o *log* da população, o tempo médio que um trabalhador permanece empregado e a taxa de evasão do ensino médio de cada município, respectivamente. A variável “Nordeste” é uma *dummy* que assume o valor 1 se a cidade observada é na região nordeste do país e 0 caso contrário. Espera-se que as variáveis “log(Pop)” e “Evasão-EM” possuam um sinal positivo, a primeira porque é natural que municípios maiores apresentem um maior número de óbitos e a última para refletir a intuição de que uma maior evasão escolar está relacionada à um baixo nível de instrução e, conseqüentemente, a uma qualidade de vida menor e expectativa de vida mais curta. O sinal esperado da variável “Emprego” é positivo, refletindo a ideia de que uma permanência maior em um emprego, na média, tende a aumentar a qualidade de vida de um trabalhador e, portanto, diminuir as chances de óbito.

No primeiro e segundo modelo, as variáveis “log(Despesa)” e “log(Seg)” são o *log* da despesa orçamentária total do município e o *log* do gasto em segurança público do município, respectivamente. No terceiro modelo, a variável representada como “log(Despesas)” contém o *log* da despesa orçamentária do município com segurança, assistência social e previdência. Esses gastos foram isolados e representados separadamente por se pensar que eles sejam mais importante para explicar o número de óbitos da cidade. Adicionalmente, espera-se que o sinal das variáveis de despesas orçamentárias seja negativo em todos os modelos, intuitivamente, esse sinal captaria a ideia que um gasto público maior tem a capacidade de aumentar o bem-estar geral e diminuir o número de óbitos do município.

No terceiro e último modelo, o componente “Educação” e seus subscritos indicam a variável “Evasão-EM”, presente nos modelos anteriores, e a variável “Aluno/Turma-EM”, que é a média de alunos por turma para aquele município. Essas variáveis foram condensadas no componente “Educação” para poupar espaço. Assim como nos outros modelos que possuem indicadores relacionados à educação, o sinal esperado dessas variáveis é positivo, transmitindo a ideia de que a piora nos indicadores educacionais aumente o número de óbitos.

	Modelo 01	Modelo 02	Modelo 03
const	-4.5198*** (0.0849)	-4.4103*** (0.0370)	-4.5053*** (0.0552)
nordeste	-0.0515*** (0.0086)	-0.0321*** (0.0095)	-0.0049 (0.0102)
log_populacao	0.9482*** (0.0035)	0.9399*** (0.0039)	0.9269*** (0.0056)
taxa_abandono_em	-0.0143*** (0.0011)	-0.0138*** (0.0011)	-0.0126*** (0.0011)
tempo_medio_emprego	0.0010*** (0.0002)	0.0011*** (0.0002)	0.0010*** (0.0002)
log_seguranca		0.0039*** (0.0007)	0.0037*** (0.0007)

log_assistencia		0.0144***	
		(0.0048)	
log_previdencia		0.0041***	
		(0.0005)	
log_despesa	0.0032		
	(0.0040)		
atu_em		-0.0023***	
		(0.0009)	
=====			
R-squared	0.9405	0.9409	0.9419
R-squared Adj.	0.9405	0.9409	0.9418
Prob(F-statistic):	0.0000	0.0000	0.0000
Prob(JB):	0.0000	0.0000	0.0000
Covariance type:	HC1	HC1	HC1
=====			
* p<.1, ** p<.05, ***p<.01			

Começando pelo termo de erro, o valor-p do teste de Jarque-Bera indica sucesso ao rejeitar a hipótese nula de que os erros dos modelos seguem uma distribuição normal. Apesar disso, seguiremos com a interpretação geral dos modelos e com o teste de hipóteses.

Os coeficientes reportados para as variáveis “log(Pop)” e “Emprego” confirmam as nossas suposições feitas *a priori*, indicando que um aumento de 1% na população de uma cidade está associado a um aumento de cerca de 0,9% no número de mortes e que o aumento no tempo médio que um trabalhador fica empregado está associado a uma diminuição de, aproximadamente, 0,1% no número de óbitos. Ambos os resultados são estatisticamente significantes ao nível de 5% em todos os modelos.

Apesar dos resultados consoantes com a intuição para as variáveis acima, as variáveis “Evasão-EM”, “Aluno/Turma-EM” apresentam um sinal discordante das suposições feitas *a priori*. Para a variável de “Evasão-EM”, os resultados obtidos reportam que um aumento na taxa de abandono do ensino médio *diminui* a o número de óbitos de um município em cerca de 1,3%, sendo esse resultado estatisticamente significativo ao nível de 5%. Analogamente, o coeficiente da variável “Aluno/Turma-EM” no terceiro modelo indica que um aumento na média de alunos por turma em uma cidade diminui o número de óbitos em 0,23% e esse resultado também é estatisticamente significativo ao nível de 5%. Inicialmente, essa contradição nos coeficientes das variáveis educacionais pode indicar a necessidade de se utilizar um controle para o perfil de renda do município. Isto é, em municípios muito pobres, um par de braço a mais no trabalho - e portanto fora da escola - pode significar um aumento da qualidade de vida e, conseqüentemente, um aumento na expectativa de vida.

Além dos indicadores educacionais acima, as variáveis orçamentárias “log(Despesa)” e “log(Despesas)” também contradizem as suposições feitas *a priori*. No primeiro modelo, a variável “log(Despesa)” apresentou um sinal positivo, indicando que um aumento de 1% no gasto público total está associado, na média, a um aumento de 0,0032% no número de óbitos, entretanto, esse resultado não é estatisticamente diferente de zero. Os resultados dos dois últimos modelos indicam que um aumento no gasto público em segurança está associado, *ceteris paribus*, a um aumento de aproximadamente 0,0038% no número de óbitos. Semelhantemente, no terceiro modelo, valores de 0,0144% e 0,0041% são reportados para a assistência social e para a previdência. Esses resultados são estatisticamente significantes ao nível de 5%.

A discordância entre as suposições iniciais e os resultados nas variáveis de gasto público apresentam algumas limitações dos modelos estimados. Uma das justificas para os sinais encontrados é a ausência de um controle por perfil etário do município. Intuitivamente, municípios que se encontram na parte superior da pirâmide etária apresentam um gasto com previdência maior sem que isso produza um efeito sob o número de óbitos.

Adicionalmente, a relação entre o número de óbitos e o município estar situado no nordeste, apontada na etapa de descrição dos dados, é significativa para os dois primeiros modelos, mas não da forma que a distri-

buição dos dados permite inferir. Como indicado na tabela, municípios localizados no nordeste do país possuem um número de óbitos entre 3,2% e 5,15% menor. Para o terceiro modelo, o coeficiente da *dummy* “Nordeste” não é estatisticamente diferente de zero.

Por fim, dois comentários: (i) o valor-p das estatísticas F de todos os modelos permite rejeitar a hipótese nula de que as variáveis explicativas, de forma conjunta, são iguais a zero; e (ii) a matriz de variância-covariância utilizada é robusta (HC1).

Referências

Sorlie, P. D., Backlund, E., & Keller, J. B. (1995). US mortality by economic, demographic, and social characteristics: the National Longitudinal Mortality Study. *American Journal of Public Health*, 85(7), 949-956.

Base dos Dados. Sistema de Informações sobre Mortalidade (SIM). Disponível em: https://basedosdados.org/dataset/br-ms-sim?bdm_table=microdados. Acesso em: 14 de jul. de 2022.

Base dos Dados. Cadastro Geral de Empregados e Desempregados (CAGED). Disponível em: https://basedosdados.org/dataset/br-me-caged?bdm_table=microdados_antigos. Acesso em: 14 de jul. de 2022.

Base dos Dados. Projeto Acesso a Oportunidades. Disponível em: https://basedosdados.org/dataset/br-ipea-acesso-oportunidades?bdm_table=estatisticas_2019. Acesso em: 14 de jul. de 2022.

Base dos Dados. Sistema de Informações Contábeis e Fiscais do Setor Público Brasileiro (Siconfi). Disponível em: https://basedosdados.org/dataset/br-me-siconfi?bdm_table=municipio_despesas_orcamentarias. Acesso em: 14 de jul. de 2022.

Base dos Dados. Indicadores Educacionais. Disponível em: https://basedosdados.org/dataset/br-inep-indicadores-educacionais?bdm_table=escola. Acesso em: 14 de jul. de 2022.

Base dos Dados. População Brasileira. Disponível em: https://basedosdados.org/dataset/br-ibge-populacao?bdm_table=municipio. Acesso em: 14 de jul. de 2022.

Scripts:

- [Aquisição e tratamento dos dados.](#)
- [Modelagem e conteúdo da regressão.](#)