

Universidade de São Paulo
Faculdade de Filosofia, Letras e Ciências Humanas
Departamento de Ciência Política
FLS 5028- Métodos Quantitativos e Técnicas de Pesquisa em Ciência Política
FLP 0406 - Métodos e Técnicas de Pesquisa em Ciência Política
1º Semestre/ 2018

Profº Dr. Glauco Peres da Silva

LISTA DE EXERCÍCIO 09

Data de entrega: 11/06/2018 (noturno) e 13/06/2018 (vespertino)

Exercício 01 (2 pontos)

Nas alternativas abaixo, marque “Verdadeiro” (V) ou “Falso” (F) e justifique as suas escolhas em, no máximo, 5 linhas. Quando marcar falso, justifique porque discorda da afirmação, e quando marcar verdadeiro, justifique o raciocínio no qual sua escolha se baseia.

() – Ao elaborar um *survey* a ser aplicado a uma amostra da população da cidade de São Paulo, o único viés com que a pesquisadora deve se preocupar é o viés (tendenciosidade) amostral, que está relacionado com a aleatoriedade e/ou a cobertura desta amostra.

FALSO. Além do viés amostral, corretamente descrito na questão, a pesquisadora também deve se preocupar com o viés na resposta, devido à formulação, à ordem das perguntas ou à maneira com que a entrevistadora as realiza, e com o viés de não-resposta, no qual as entrevistadas se recusam a responder perguntas ou não são encontradas (Agresti e Finlay, 2012, p. 35-38).

() – Em pesquisas de n-grande nas Ciências Sociais utilizamos dados observacionais ou dados experimentais. Contudo, só podemos apontar relações de causalidade em pesquisas que utilizam dados experimentais.

FALSO. Desenhos de pesquisa que utilizam dados observacionais podem levar a conclusões causais. Para tanto, é necessário as pesquisadoras adotem desenhos de pesquisa que simulem a lógica experimental a dados observacionais (Silva, 2017, seção 3.1.1).

() – Duas variáveis contínuas (X e Y), independentes entre si, possuem correlação igual a zero. Contudo, diante somente da informação de que a correlação entre duas variáveis (X e Y) é igual a zero, não podemos afirmar a inexistência de causalidade entre elas.

VERDADEIRO. O teste de correlação não garante que duas variáveis sejam independentes entre si. A medida de correlação “reflete, assim, uma propriedade dos valores assumidos para cada uma das variáveis e não considera a existência de outras variáveis” (Silva, 2017, p. 60). Para afirmarmos a não causalidade entre elas seria necessário considerar possíveis variáveis Z que afetem a relação entre X e Y.

() – King, Keohane e Verba (1994) apontam dois pressupostos para a estimação de efeitos causais. O primeiro destes pressupostos (*unit homogeneity*) afirma que duas unidades são homogêneas se os valores esperados da variável dependente para cada uma delas são iguais quando o valor da variável explicativa é o mesmo. Uma versão menos rigorosa desse pressuposto (o pressuposto de efeitos constantes) assume que oscilações pequenas na variável explicativa não afetam a possibilidade de se obter valores esperados idênticos da variável dependente.

FALSO. A caracterização do pressuposto de homogeneidade das unidades é correta. Porém, segundo King, Keohane e Verba (1994, p. 92-93) o pressuposto de efeitos constantes assume que o efeito causal é o mesmo para unidades com o mesmo valor da variável explicativa, ainda que o valor da variável dependente seja diferente.

Exercício 02 (3,5 pontos)

Pesquisas quantitativas em Ciências Sociais estão primordialmente preocupadas em avaliar o efeito de uma variável explicativa (X) em outra dependente (Y). Dessa forma, a relação entre elas é comumente ilustrada como $X \rightarrow Y$. Contudo, conforme demonstram os textos obrigatórios e complementares deste tópico, a preocupação central destas pesquisas na atualidade é com o estabelecimento de relações causais. Nesse sentido, na quase totalidade dos casos, a relação ilustrada anteriormente não é suficiente para estabelecer relações causais em nossa “realidade multivariada”.

- a) **(0,5 ponto)** Tendo em vista o texto acima, explique porque a relação $X \rightarrow Y$ apresentada não é suficiente para estabelecer relações causais em pesquisas nas Ciências Sociais (**Tamanho máximo: 5 linhas**).

A realidade social é bastante complexa e a relação causal entre duas variáveis neste cenário são dificilmente encontradas, sem que sejam consideradas outras variáveis (Z) que possam afetar esta relação.

- b) **(2 pontos)** De que maneiras outras variáveis (Z) podem afetar a relação entre $X \rightarrow Y$? Dê ao menos um exemplo concreto para cada forma elencada na resposta (**Tamanho máximo: 20 linhas**)

[Dica: reformule a representação $X \rightarrow Y$, de forma a incluir as variáveis Z nestas ilustrações, explicando como Z está afetando a relação original]

Conforme Silva (2017), tais variáveis podem afetar a relação $X \rightarrow Y$ das seguintes formas:

- (1) $X \rightarrow Y \leftarrow Z$ (Silva, 2017, p. 59)

Neste caso, as variáveis representadas por Z afetam a variável dependente (Y) simultaneamente a X. Isto é, tanto X quanto Z estão contribuindo para o valor observado em Y. X e Z podem ter um efeito positivo sobre Y, caso em que Z estará superdimensionando o efeito observado. X e Z podem ter efeitos opostos, o que subdimensionaria (ou até mesmo anularia) o efeito de X em Y.

Silva (2017, p. 59) apresenta o seguinte exemplo: “Em nosso exemplo sobre a avaliação do PBF, o pesquisador pode considerar que não foi apenas o fato de as famílias estarem submetidas ao programa (X) que provocou alterações nas suas condições materiais de vida (Y). Seria possível elencar um outro conjunto de fatores que tenham afetado Y que não tenham a ver com X, como, por exemplo, as condições macroeconômicas do país (Z). O

ambiente macroeconômico do país está evidenciado pelos índices de desemprego, de inflação, pela variação do produto interno bruto etc”.

Outro exemplo é a escolaridade e a chuva afetando o comparecimento eleitoral. São duas variáveis diferentes que podem afetar a variável dependente. Uma eleição em um dia muito chuvoso pode diminuir o comparecimento eleitoral, diminuindo o resultado observado, comparativamente ao valor esperado pela relação $X \rightarrow Y$.

(2)  (Silva, 2017, p. 66)

Neste caso, as variáveis representadas por Z afetam tanto a variável dependente (Y) quanto a variável explicativa (X). Nesse caso, o efeito de X em Y pode, inclusive, advir da existência de Z.

Para o caso do efeito de X em Y ser devido à presença de Z, Silva (2012, p. 67) apresenta o seguinte exemplo: “o cálculo da correlação entre o número de ataques de tubarões a pessoas (Y) em relação ao consumo de sorvetes por habitantes (X). Não há razão alguma para supor que exista uma correlação desse tipo entre as variáveis, mas, em termos práticos, ela deve ser alta. Nesse caso, pode-se supor que as cidades nas quais ocorrem os ataques de tubarões são cidades litorâneas, com grande incidência solar. O elevado calor dessas regiões (Z) explica também o alto consumo de sorvetes.”

Dois exemplos de Kellstedt e Whitten (2015) também são úteis: pesquisas sobre raça (X) e participação política (Y) que podem ser afetadas pelo status socioeconômico (Z) (Kellstedt e Whitten, 2015, p. 86-87) ou o tipo de escola frequentado por alunos que migraram de uma escola pública para uma particular após uma política de *vouchers* (X) e o resultado obtido por alunos em um teste padronizado (Y), ambos afetados pelo envolvimento dos pais na vida dos alunos (Z).

c) **(1 ponto)** O seguinte experimento será conduzido para avaliar o efeito da inclusão de uma questão na média dos resultados da prova de alunos de uma turma de Métodos do curso de Ciências Sociais da USP:

Os 300 alunos de uma turma de Métodos do curso de Ciências Sociais da USP realizarão uma prova sobre seus conhecimentos ao final do semestre. Desejamos avaliar o impacto da inclusão de uma nova questão na média padronizada das notas desta avaliação. O tempo de prova oferecido aos dois grupos é o mesmo. Partindo do pressuposto de que é possível classificar a ordem de dificuldade das questões, todas as questões da prova possuem o mesmo grau de dificuldade. Os alunos da turma são divididos em 2 grupos, com 150 indivíduos em cada. O primeiro grupo realizará uma prova de 5 questões e o outro grupo uma avaliação de 6 questões. A nota final da avaliação para os dois grupos será idêntica: 10 pontos. O efeito da inclusão de uma questão será avaliado pela diferença entre a nota média dos dois grupos.

Como podemos garantir, no desenho deste experimento, que outras variáveis Z não interferirão na estimação do efeito causal investigado? **(Tamanho máximo: 10 linhas)**

Trata-se de um caso no qual a aleatorização na atribuição do tratamento garante que outras variáveis contextuais Z não vão intervir na estimação do efeito causal médio investigado. Temos uma amostra de 300 alunos que deveriam ter se preparado para a prova, seja ao longo do semestre ou com estudos de véspera. Porém, certamente existirão alunos com diferentes níveis de preparação para a avaliação.

Além disso, alguns alunos podem ter trabalhado por 8h no dia, ou passado a noite em claro, indispostos por alguma doença. Ao aleatorizar quais dos alunos receberão o tratamento (a inclusão de uma 6ª questão na prova) e quais não receberão, dentro de um universo de n razoavelmente grande ($n = 300$), controlamos por estes efeitos contextuais relacionados à preparação para prova, condições físicas para realização da prova no dia, etc. Na média, os grupos de tratamento e controle deverão ser iguais nessas características.

Exercício 03 (4,5 pontos)

Interpretar os resultados de estudos quantitativos é uma atividade comum para diversos cientistas sociais, seja na academia ou nos setores público, privado e terceiro. Suponha que você é um gestor de políticas públicas de saúde no nível federal e sua secretaria está muito preocupada com os índices de automedicação no Brasil. Você precisa decidir se sua secretaria fará ou não uma campanha em vídeo contra automedicação para os brasileiros adultos. Infelizmente, a secretaria responsável pelo planejamento do projeto, sem informá-lo anteriormente, já preparou o vídeo da campanha e encomendou quatro estudos de efetividade de campanha para quatro institutos de pesquisa diferentes, sob a alegação de que “quantos mais estudos, melhor”. Essas pesquisas (hipotéticas) já estão prontas. Em **todas** as pesquisas, a pergunta feita aos participantes foi “Você usaria o medicamento X sem anteriormente consultar um médico?”, que iremos chamar de **Pergunta I**. Eles podiam responder “Sim” ou “Não” e não houve indivíduos que deixaram de responder alguma das pesquisas. Responda as questões abaixo. Atente-se ao tamanho máximo das respostas, pois as linhas excedentes a ele **não** serão corrigidas.

- a) **(0,5 ponto)** Qual a população de interesse para sua secretaria? (**Tamanho máximo: 1 linha**).

A população de interesse é a população de brasileiros adultos.

A Tabela 1 mostra os resultados das quatro pesquisas (em um formato razoavelmente similar ao que geralmente usamos em publicações científicas). As duas primeiras linhas são as proporções dos indivíduos que responderam “Sim” à Pergunta I nos grupos de tratamento e controle, respectivamente. A diferença de médias se refere à diferença entre as médias (proporções) dos grupos de tratamento e de controle. O erro-padrão é o erro-padrão da diferença das médias. A estatística t é a divisão da diferença das médias pelo erro-padrão. IC 95% é o intervalo de 95% de confiança para a diferença das médias. O p -valor é referente ao teste da hipótese nula de não diferença de médias entre grupos de controle e tratamento, contra a hipótese alternativa de diferença de médias entre controle e tratamento. Finalmente, a última linha mostra o tamanho da amostra.

[**Dica:** O aluno pode revisar a seção “Robustez para violações da suposição de normalidade” em Agresti e Finlay, p. 146, para verificar as consequências de amostras não aleatórias e pequenas. Os resultados são apresentados no contexto de um intervalo de confiança de uma média, mas se mantêm, de maneira geral, para os testes de hipóteses de diferenças de média, à exceção do tamanho. Uma amostra de 15 observações pode ser ainda muito pequena para um teste de diferenças de médias. Outra fonte sobre testes de diferenças de médias que o aluno pode usar é a seção 7.4.2 de Kellstedt e Whitten (p. 177-182).]

Tabela 1

	Pesquisa A	Pesquisa B	Pesquisa C	Pesquisa D
Proporção (Tratamento)	0,12	0,32	0,34	0,30
Proporção (Controle)	0,15	0,70	0,41	0,36
Diferença de médias	-0,03	-0,38	-0,07	-0,06
Erro-padrão	0,02	0,10	0,02	0,02
Estatística t	-1,24	-3,68	-3,24	-2,56
P-valor	0,214	0,000	0,001	0,011
IC 95%	[-0,08; 0,02]	[-0,58; -0,18]	[-0,11; -0,03]	[-0,11; -0,01]
Tamanho da amostra	800	80	2.000	1.600

Pesquisa A. A consultoria responsável pela pesquisa A enviou um e-mail de convite para todos os quase 60.000 estudantes da Universidade de São Paulo e 5.000 voluntários se interessaram em participar. Então, a consultoria escolheu aleatoriamente 800 voluntários para comporem a amostra. Destes, 400 foram escolhidos aleatoriamente para o grupo de tratamento e assistiram o vídeo da campanha, ao passo que os outros 400 foram para o grupo de controle e assistiram um vídeo placebo, sem relação qualquer com temas de saúde. Após assistirem o vídeo, todos os indivíduos responderam a Pergunta I e essas respostas foram documentadas.

- b) **(1 ponto)** A diferença de médias indica que a pesquisa A estimou o efeito de redução de 3 pontos percentuais (na Tabela 1, -0,03), associado ao vídeo da campanha, na proporção de indivíduos que responderam que se automedicariam com o medicamento X. Sobre a pesquisa A, responda às seguintes perguntas. A diferença de médias é significativa ao nível de 95%? A amostra é grande? A amostra é representativa da população de interesse? O desenho de pesquisas dá motivos para supor que os grupos de tratamento e controle estão balanceados? Você acredita que o efeito estimado de menos 13 pontos percentuais é causal? Justifique as respostas brevemente. **(Tamanho máximo: 10 linhas).**

O efeito estimado não é significativo ao nível de 5%, pois o p-valor é mais de 21% e o intervalo de confiança de 95% contém zero. A amostra é grande, mas não é representativa da população brasileira adulta, pois só contém estudantes de graduação da USP (majoritariamente caracterizados por idade jovem, renda relativamente elevada e origem paulista). Além disso, há auto-seleção, pois só 5 mil dos 90 mil alunos se voluntariaram, ainda que a aleatorização da amostra dentre voluntários ajude a mitigar esse problema. Como houve aleatorização do tratamento, os grupos de tratamento e controle devem estar balanceados. Não há efeito causal significativo estimado (pois não conseguimos rejeitar a hipótese nula). Mesmo que houvesse, não se referiria à população de interesse.

Pesquisa B. A fundação contratada fez um grupo focal com 40 indivíduos, selecionados para a amostra por ordem de inscrição em um formulário postado publicamente na página do Facebook do Ministério da Saúde (MS). Todos os 40 indivíduos responderam a Pergunta I antes de assistirem o vídeo da campanha, respostas estas que a fundação considerou como compondo o grupo de controle. Para construir o grupo de tratamento, a fundação registrou a resposta à Pergunta I após os indivíduos assistirem o vídeo. Note que o tamanho da amostra é 80, pois cada indivíduo foi contado duas vezes.

- c) **(1 ponto)** Sobre a pesquisa B, responda às seguintes perguntas. Qual o efeito estimado? O efeito estimado é significativo ao nível de 95%? A amostra é grande? A amostra é representativa da população de interesse? O desenho de pesquisas dá motivos para supor que os grupos de tratamento e controle estão balanceados? Você acredita que o efeito estimado é causal? Justifique as respostas brevemente. **(Tamanho máximo: 15 linhas)**.

O efeito estimado é a redução de 38 pontos percentuais na proporção de declaração de intenção de automedicação. O efeito é significativo ao nível de 5%, pois o p-valor é próximo de zero e o intervalo de confiança de 95% não contém zero. A amostra é relativamente pequena e não é representativa da população brasileira adulta, pois houve auto-seleção dos participantes (amostra de conveniência). Como o mesmo indivíduo é comparado logo antes e logo depois do tratamento, espera-se haver um balanceamento entre controle e tratamento. No entanto, fazer a mesma pergunta duas vezes para os indivíduos, uma delas logo após o vídeo, tende a induzir a resposta contra a automedicação. Não à toa, o efeito é tão pronunciado. O resultado desse experimento tenderá a não replicar o comportamento dos indivíduos em seu ambiente usual. Nesse sentido, a própria pressão do pesquisador age como um fator pelo qual o desenho de pesquisa não provê controle. O efeito estimado pode até ser causal, mas apenas nas condições do experimento (não se observaria no mundo real).

Observação: O aluno que refizer as contas perceberá que usamos a fórmula do erro-padrão da diferença de proporções como se as observações fossem todas independentes entre si. Como observamos o mesmo indivíduo antes e depois do tratamento, não existe essa independência e o uso dessa fórmula está errado. O correto seria fazer o teste de McNemar (vide Agresti e Finlay, cap. 7).

Pesquisa C. O instituto de pesquisa responsável selecionou uma amostra aleatória estratificada em multiestágios que contemplava toda a população brasileira, com 2000 indivíduos, entrevistados em suas respectivas casas. O grupo de tratamento foi composto por 1.000 indivíduos que quiseram, voluntariamente, ver o vídeo da campanha. Os outros 1.000 indivíduos viram um vídeo placebo sem relação com saúde e compuseram o grupo de controle. Após assistirem os vídeos, todos os indivíduos responderam a Pergunta I.

- d) **(0,5 ponto)** Sobre a pesquisa C, responda às seguintes perguntas. Qual o efeito estimado? O efeito estimado é significativo ao nível de 95%? A amostra é grande? A amostra é representativa da população de interesse? O desenho de pesquisas dá motivos para supor que os grupos de tratamento e controle estão balanceados? Você acredita que o efeito estimado é causal? Justifique as respostas brevemente. **(Tamanho máximo: 10 linhas)**.

O efeito estimado é a redução de 7 pontos percentuais na proporção de declaração de intenção de automedicação. O efeito é significativo ao nível de 5%, pois o p-valor é 0,1% e o intervalo de confiança de 95% não contém zero. A amostra é grande e representativa da população brasileira adulta, devido à aleatorização estratificada multinível. No entanto, os indivíduos escolheram se receberiam o tratamento ou não, de maneira que não há motivos para crer que os grupos de tratamento e controle estejam balanceados. É possível que os indivíduos que escolheram assistir o vídeo sejam sistematicamente diferentes dos indivíduos que não o fizeram. O grupo de tratamento pode, por exemplo, ser mais interessado nos temas de saúde. Por isso, não se pode considerar o efeito como causal.

Pesquisa D. A consultoria contratada também utilizou uma amostra aleatória estratificada em multiestágios que contemplava toda a população brasileira. A amostra tinha 1600 indivíduos, divididos aleatoriamente em dois grupos de igual tamanho, tratamento e controle. O grupo de tratamento recebeu um panfleto que afirmava que a automedicação é muito perigosa (e **não** assistiu o vídeo da campanha). O grupo de controle recebeu um panfleto placebo, sem relação com saúde. Após lerem os panfletos, ambos os grupos responderam a Pergunta I.

- e) **(0,5 ponto)** Sobre a pesquisa D, responda às seguintes perguntas. Qual o efeito estimado? O efeito estimado é significativo ao nível de 95%? A amostra é grande? A amostra é representativa da população de interesse? O desenho de pesquisas dá motivos para supor que os grupos de tratamento e controle estão balanceados? Você acredita que o efeito estimado é causal? Justifique as respostas brevemente. **(Tamanho máximo: 10 linhas).**

O efeito estimado é a redução de 6 pontos percentuais na proporção de declaração de intenção de automedicação. O efeito é significativo ao nível de 5%, pois o p-valor é 1,1% e o intervalo de confiança de 95% não contém zero. A amostra é grande e representativa da população brasileira adulta, devido à aleatorização estratificada multinível. Como o tratamento é aleatorizado, é razoável esperar que os grupos de tratamento e controle sejam balanceados. Pode-se considerar o efeito como causal. Observação: Não obstante termos um efeito causal, não sabemos se o efeito estimado para um panfleto pode ser generalizado para o vídeo.

- f) **(1 ponto)** Você considera que a reputação das quatro consultorias é a mesma e sabe que as pesquisas foram executadas aproximadamente no mesmo período e usaram exatamente a mesma pergunta de pesquisa (Pergunta I), de maneira que não parece problemático comparar seus resultados, desde que se considere as particularidades de cada desenho de pesquisa. Além disso, você sabe que a campanha de vídeo planejada é bastante cara e, caso seja inócua, terá desperdiçado recursos que poderiam ter sido dirigidos a serviços essenciais aos cidadãos, recursos estes que serão retirados do orçamento da sua secretaria. *Com base nos resultados das quatro pesquisas e tendo em vista a população de interesse, você decide promover a campanha ou não? Justifique.* Obs.: Como o objetivo da questão é estimular a reflexão do aluno sobre interpretação de resultados, argumentos de textos sobre comunicação de políticas públicas e mesmos citações de materiais do curso não serão admitidas como corretas nessa

questão. Reflita com base nos conhecimentos do curso e argumente. (**Tamanho máximo: 20 linhas**).

Os efeitos estimados pelas pesquisas A e B não são representativos para a população de interesse. Sobre a pesquisa A, se estamos preocupados com indivíduos adultos do Brasil, não é razoável considerar apenas jovens de uma mesma universidade, particularmente de uma elitizada e no estado mais rico do país. Os resultados da pesquisa A sugerem que os estudantes (ou os jovens, da USP, de São Paulo ou do Brasil, a depender de quanto você achar crível extrapolar a amostra) não são muito suscetíveis à campanha (possivelmente por já se automedicarem relativamente pouco antes mesmo da campanha - note que a proporção do grupo de controle é baixa). Sobre a pesquisa B, os seguidores do MS no Facebook não são representativos de todos os brasileiros adultos. Além disso, o desenho de pesquisa B é bastante frágil, pelos motivos citados na resposta 3-c. Os resultados das pesquisas C e D parecem mais confiáveis, ainda que tenham seus problemas.

As pesquisas C e D são representativas da população de interesse. No entanto, a pesquisa C incorreu em manipulação da atribuição do tratamento pelos participantes e não se sabe bem se o experimento da pesquisa D pode ser generalizado para a campanha, uma vez que seu tratamento é diverso do tratamento produzido pela campanha (panfleto x vídeo). De todas as pesquisas, a D parece mais confiável, pois é a única que estima algum efeito causal de redução da proporção reportada de intenção de automedicação provocado por uma estratégia de comunicação para a população de interesse. Tipicamente, os experimentos não replicam exatamente o comportamento dos agentes em seu ambiente usual e faz-se o possível para aproximá-lo. Assim, é razoável acreditar no resultado da pesquisa D e decidir promover a campanha.

Como a questão é aberta, outras respostas com justificativas bem embasadas serão aceitas.

Exercício 04 (5 pontos): Apenas para a Pós-Graduação

Neste exercício, iremos estudar o desenho de pesquisa e alguns resultados do Working Paper “*Avis, E., Ferraz, C., & Finan, F. (2017). Do Government Audits Reduce Corruption?*”. Como o trabalho é extenso (82 páginas), selecionamos alguns excertos, que foram disponibilizados no Moodle. Leia a seleção disponibilizada para responder às questões propostas. Recomendamos que você não leia o trabalho inteiro para responder os exercícios, pois as questões foram pensadas especificamente para a seleção. Além disso, note que o paper fala em estimacões estruturais e em formas reduzidas, mas abordaremos apenas os resultados em forma reduzida, que já podem ser lidos com os conhecimentos do curso até agora (ainda que você provavelmente não saiba os significados desses termos econométricos). Esperamos que esse exercício ajude os alunos que ainda não tiveram contato com métodos quantitativos a começar a ler resultados desse tipo de trabalho. Aproveite o exercício para se familiarizar com a análise de desenhos de pesquisa e dos resultados e construa respostas breves. Você só precisa justificar os pontos que explicitamente exigirem isso.

- a) **(1 ponto)** Qual a principal pergunta teórica da pesquisa? Expresse sua resposta como uma única relação monocausal entre conceitos (não entre as variáveis), como no exemplo a seguir: “maiores investimentos em infraestrutura causam crescimento econômico”. Os autores expressam essa pergunta de maneira determinística ou probabilística? Quais as variáveis usadas para operacionalizar esses conceitos? Há uma relação clara entre medida e conceito? (**Tamanho máximo: 10 linhas**).

A pergunta da pesquisa é se auditorias promovidas pelo governo reduzem a corrupção dos políticos no nível local e foi expressa de maneira determinística (“In this paper, we investigate the role government audits play in reducing political corruption in local government through the promotion of electoral and judicial accountability.”, p. 1).

A variável dependente é a corrupção em nível local. Para medi-la, os autores utilizaram o número de irregularidades classificadas pela CGU como moderada ou severa. A variável explicativa é uma *dummy* que indica se a CGU já havia auditado o município no passado. Há ainda uma terceira medida para improbidade (*mismanagement*), que indica o número de irregularidades administrativas e procedimentais no município.

Há uma relação clara entre as medidas e os conceitos, ainda que medidas outras (melhores ou piores) pudessem ser usadas hipoteticamente, tais como o valor desviado em corrupção. Os autores fazem uma discussão detalhada de que as diferenças da categorização de irregularidade como improbidade e corrupção é clara, mas que a diferença entre corrupções severa e moderada feita pela CGU não é muito nítida (não à toa, eles agrupam esses casos). Uma crítica possível é que a medida é de corrupção auditada (isto é, aquilo que os auditores conseguiram achar), não de corrupção realizada.

- b) **(2 pontos)** Qual a população de interesse do estudo? Qual foi a amostra selecionada? Essa amostra é grande? A amostra é representativa da população de interesse? Por quê? Quais unidades formam o grupo de controle? Quais unidades foram o grupo de tratamento? **(Tamanho máximo: 15 linhas).**

A população de interesse são os municípios brasileiros.

A amostra é composta pelos municípios que foram auditados pela CGU pelo menos uma vez entre julho de 2006 e março de 2013 (p. 10). A amostra é grande (cerca de 900 a 1000 municípios, a depender da especificação, conforme a tabela 2).

A amostra é aleatória estratificada, pois a CGU audita municípios aleatoriamente, dentro de cada estado (p. 7-8). Sendo assim, a probabilidade de ser auditado é a mesma para os municípios de cada estado, mas difere entre estados. No entanto, apenas as cidades com até 500 mil habitantes participam da loteria (com pequenas mudanças no critério ao longo do tempo, segundo os autores). Assim, a amostra é representativa para municípios pequenos e médios (que são praticamente toda a população dos municípios brasileiros).

Os municípios auditados apenas uma vez são o grupo de controle. Os municípios auditados mais de uma vez são o grupo de tratamento.

- c) **(1 ponto)** O desenho de pesquisa é observacional, experimental ou quase-experimental? Há aleatorização do tratamento? Se houver, os autores dão argumentos contextuais para que o leitor acredite que o tratamento foi aleatorizado? Se sim, cite dois argumentos. Os autores mostram evidências de estatística inferencial que indiquem que os grupos estão balanceados e que, portanto, a aleatorização parece ter sido bem feita (isto é, que não houve manipulação arbitrária da atribuição do tratamento)?

[Dica: Veja a tabela 1. Note que as notas de rodapé das tabelas trazem informações importantes. Os autores citam que fizeram controles por certos efeitos fixos. Por ora, você pode ler os resultados como testes de diferenças de médias (exatamente como fizemos no exercício 3), tendo em mente que houve o cuidado de controlar por outras variáveis. Aprenderemos mais sobre o que significa “controlar” até o fim do curso]. (**Tamanho máximo: 10 linhas**).

O desenho é quase-experimental, pois o tratamento é aleatorizado, embora a atribuição de tratamento não esteja sobre controle dos pesquisadores.

Para convencer o leitor de que houve aleatorização do tratamento, os autores explicam que o sorteio é público (pela Loteria Federal), que as ordens de inspeção com os projetos a serem inspecionados são aleatórias, que os auditores não escolhem quais projetos vão auditar, que os auditores atuam em times (e não tem muita margem para manipulação individual) e que a CGU foi um órgão bastante autônomo no período. Além disso, eles explicam que a probabilidade de ser auditado é a mesma para todos os municípios do Estado. Um argumento importante que os autores não mencionam é que a CGU é independente dos prefeitos e, portanto, estes não têm como escolher se serão auditados ou não.

Uma análise rápida da tabela 1 mostra que nenhuma das diferenças médias é mais que duas vezes maior que os erros-padrões (o que indica que não há diferenças significativas, a um nível de 5%, entre as médias dos grupos de controle e tratamento), em várias características que não o tratamento (população, desigualdade de renda, etc.). Particularmente, não há diferenças significativas no número de ordens de inspeção recebidas. No texto (pg. 13), os autores explicitamente indicam que apenas uma das variáveis tem diferença de média significante a 10%. De maneira geral, os grupos de tratamento e controle estão bem balanceados, portanto. Esse conjunto de evidências dá bastante robustez à ideia de que a condição de aleatorização do tratamento foi atendida.

- d) **(1 ponto)** Há evidências de um efeito do tratamento significativo para a pergunta principal de pesquisa? Quanto é esse efeito estimado? Se sim, os autores dão algum indicativo de que o efeito estimado é substantivo para a política pública? (Um efeito de 0,0001% no crescimento econômico, por exemplo, ainda que estatisticamente significativo, não tem muita importância substantiva). Você acredita que o efeito é causal?

[Dica: Veja a tabela 2. Como é normal na apresentação desse tipo de resultados, a primeira linha tem o efeito estimado para o tratamento e as outras têm os efeitos estimados dos controles, com os quais você não precisa se preocupar, por ora. Novamente, você pode ler a primeira linha como um teste de diferença de médias (sabendo que houve controles). Note que a Tabela 2 responde a pelo menos duas perguntas de pesquisa e que o enunciado deste exercício apenas te indaga algo apenas sobre a pergunta principal – supostamente a mesma que você respondeu no item 4-a.] (**Tamanho máximo: 10 linhas**).

Há evidência de redução de aproximadamente 8% (ou 6% ou 7%, a depender da especificação) no número de casos de corrupção naqueles municípios reauditados pela CGU. Os autores indicam (p.22) que esse efeito corresponde a uma redução de R\$ 355 mil de corrupção por município por ano, segundo a média de transferências federais recebidas anualmente pelos municípios da amostra (R\$ 15 milhões) - ou R\$ 567 mil, quando os autores estimam estruturalmente o efeito, incluindo os efeitos de *spillover* (a estimação estrutural não foi objeto do exercício), o que parece substantivo.

Dado que a amostra é grande e que há bastantes evidências de que a amostra foi aleatorizada e de que o tratamento foi aleatorizado, tudo indica que o efeito estimado é causal.

Obs.: Deve-se lembrar, no entanto, que a amostra é representativa para a população de municípios brasileiros médios e pequenos, não para a população total. Na prática, como os municípios com mais de 500 mil habitantes são pouco mais de 40 hoje (e eram menos à época abordada pela pesquisa) e que provavelmente apenas uma parte deles iria participar de uma eventual amostra aleatória de todos os municípios brasileiros, é pouco provável que os resultados mudassem muito se eles fossem incluídos.