

MATEUS SANTOS PIRES

**É PRECISO PERDER PARA DAR VALOR? UMA
ANÁLISE EMPÍRICA SOBRE O VALOR ESTATÍSTICO
DA VIDA DURANTE A PANDEMIA DA COVID-19**

SÃO PAULO

2021

MATEUS SANTOS PIRES

**É PRECISO PERDER PARA DAR VALOR? UMA ANÁLISE
EMPÍRICA SOBRE O VALOR ESTATÍSTICO DA VIDA DURANTE
A PANDEMIA DA COVID-19**

Monografia apresentada ao Curso de Ciências Econômicas da Universidade de São Paulo, como requisito parcial para a obtenção do título de bacharel em Economia.

Orientadora: Fabiana Fontes Rocha

SÃO PAULO

2021

Resumo

Esta monografia tem como intuito estimar o valor estatístico da vida (VSL) para os Estados brasileiros durante a pandemia da COVID-19 e descobrir se existe relação causal entre a VSL e a preferência eleitoral dos indivíduos, representada pelo resultado do segundo turno das eleições presidenciais de 2018. Para cálculo do valor estatístico da vida, utiliza-se de um método no qual se considera o dispêndio monetário em detrimento da variação de mortos por COVID-19. Para a análise causal, o método adotado foi o de grafos acíclicos dirigidos (DAG). Geradas as regressões, embora exista significância conjunta das variáveis que explicam o valor estatístico da vida, não foi possível afirmar, por meio do método adotado, que haja efeito causal da preferência eleitoral sobre a VSL.

Palavras-chaves: Valor Estatístico da Vida, VSL, Grafo Acíclico Dirigido, DAG, COVID-19, Preferência Eleitoral

Código JEL: C32, H12, I18, J17

Abstract

This paper has the intent to estimate the value of a statistical life (VSL) for the brazilian states during the COVID-19 pandemic and find out if there is a causal relation between the VSL and the electoral preference, represented by the result of the second round of the 2018 presidential elections. The calculus for the value of a statistical life uses a method that considers the monetary outlay at the expense of the variation of deaths by COVID-19. For the causal analysis, the adopted method was the directed acyclic graph (DAG). With the regressions, although there is joint significance for the variables that explain the value of a statistical life, it wasn't possible to assert, with the method used by this paper, that there is a causal effect by the electoral preference over the VSL.

Keywords: Value of a Statistical Life, VSL, Direct Acyclic Graph, DAG, COVID-19, Electoral Preference

JEL Codes: C32, H12, I18, J17

Agradecimentos

A trajetória acadêmica no Brasil é uma luta longa e ardúa, repleta de desafios e conquistas. Com esta monografia, encerro este capítulo da minha vida acadêmica e, como primeiro membro da família a se formar na Universidade de São Paulo, me sinto no dever de agradecer àqueles que me deram suporte durante estes últimos 5 anos.

Inicialmente, agradeço à minha mãe Sandra, meus irmãos Débora e Gustavo e minha avó Izabel, por proporcionarem o impulso necessário e todos os insumos que precisei. Também agradeço aos meus demais familiares e amigos, em especial meu tio Roberto, por me convencer a desistir da graduação em matemática e seguir a carreira de economista.

Gostaria de agradecer as duas entidades que me mostraram como o ambiente universitário pode ensinar através de experiências únicas. Ao atletismo FEA USP, por me ensinar a ter persistência e proporcionar amizades incríveis e à Associação Atlética Acadêmica Visconde de Cairu, agradeço pelos melhores momentos da minha vida universitária e, sobretudo, por me fazer um feano fanático. Em especial, agradeço à gestão 2019 - "Somos um só", por serem os grandes amigos que são.

Agradeço à FEA USP, instituição que durante os últimos 5 anos aprendi a amar e à minha orientadora Fabia Rocha, por me proporcionarem um curso de excelência.

Gostaria ainda de agradecer a Deus, por ter colocado todas estas pessoas no meu caminho e me ajudado durante todo esse período.

Por fim, agradeço às duas pessoas que me incentivaram a seguir a carreira de economista, meu falecido pai, Luiz e minha professora de história do ensino médio, Raquel.

Obrigado, FEA USP!

Mateus Santos Pires

Sumário

1	Introdução	8
2	O Valor Estatístico da Vida	11
3	Avaliando o Efeito Causal	14
4	Mensurando o Valor Estatístico da Vida	16
5	Descrição dos Dados e Modelo Causal	18
5.1	VSL	18
5.2	Modelo DAG	18
6	Resultados	21
6.1	VSL	21
6.2	Modelo DAG	22
7	Conclusão	29
	REFERÊNCIAS	31
1	Anexo I: Do-file do software R	33

Lista de tabelas

Tabela 1 – Estimativas do valor estatístico da vida (em milhões de reais)	21
Tabela 2 – Estatísticas descritivas: Base 1	22
Tabela 3 – Estatísticas descritivas: Base 2	22
Tabela 4 – Estatísticas descritivas: Base 3	23
Tabela 5 – Estatísticas descritivas: Base 4	23
Tabela 6 – Regressões lineares	25
Tabela 7 – Regressões lineares (continuação)	26

Lista de ilustrações

Figura 1 – Relação Causal: Preferência Eleitoral e VSL	15
Figura 2 – Modelo DAG	19

1 Introdução

A carreira de um economista pode tomar diversos rumos, tanto na academia como no restante do mercado de trabalho e, em muitas oportunidades, será solicitado ao economista que encontre o valor de um bem ou serviço. Embora existam inúmeros modelos para precificação, um bem em particular apresenta certa dificuldade para ser mensurado, neste caso, a vida. O valor estatístico da vida (VSL¹) é um método que busca estimar a valoração de uma única vida, por um grupo de indivíduos. Intuitivamente, a VSL tem como objetivo entender qual a dimensão do dispêndio financeiro em troca de uma redução na probabilidade de ocorrer alguma fatalidade, ou seja, busca-se compreender quanto um grupo de indivíduos está coletivamente disposto a pagar para conseguir salvar uma vida.

Formuladores de políticas públicas frequentemente precisam entender a relação entre custo e benefício para que possam tomar decisões quanto aos gastos públicos. Um exemplo desta relação é explicitado no estudo feito por [Miller \(2000\)](#), que apresenta a possibilidade de se realizar a manutenção de uma estrada, tendo como consequência uma potencial redução no risco de acidentes fatais. Este é o tipo de *tradeoff* que políticos estão inclinados a compreender, a fim de decidir se será realizada a manutenção na estrada ou não. A VSL, que mede o valor estatístico da vida, tem como papel fundamental, na aplicação de políticas públicas, a capacidade de mostrar ao político se o custo da manutenção na via, por exemplo, é menor do que o custo das vidas perdidas em futuros acidentes, devido ao mau estado da estrada.

Muitos fatores influenciam na valoração da vida, uma vez que esta depende das preferências dos indivíduos, sendo uma delas a inclinação política. A escolha de um chefe de Estado, por exemplo, pode indicar o perfil de preferência da população, de modo que tais preferências levem a performances diferentes no que diz respeito a políticas públicas, como é possível observar em [Abras, Fava e Kuwahara \(2021\)](#), que apresentam no estudo o melhor desempenho das chefes de Estado mulheres durante a pandemia da COVID-19, se comparadas aos chefes de Estado homens. Também no contexto da pandemia, pode-se observar que municípios que votaram no atual Governo Federal têm um desempenho pior no controle epidemiológico da pandemia, apresentando maior número de infectados e mortos ([AJZENMAN; CAVALCANTI; MATA, 2020](#)).

Segundo [Black \(1948\)](#), o agente público eleito representa as preferências do eleitor mediano. Assim, é intrigante a possível correlação existente entre as preferências que levaram à eleição do político e as preferências que determinam o valor da vida. Por este fator, esse projeto se debruça sobre a determinação da VSL e o efeito causal ocasionado pela relação com fatores ligados à preferência dos indivíduos, neste caso, a preferência política.

¹VSL: *Value of a Statistical Life*

Um dos meios de se estimar a VSL é por meio de pesquisas com a população, a fim de encontrar uma medida de risco-retorno média. Dada a dificuldade de se realizar uma pesquisa ampla o suficiente para atingir os objetivos deste estudo, será atribuído aos governos estaduais a característica de melhor agregador de preferências, de modo que possa representar o eleitor mediano (BLACK, 1948). Deste modo, pesquisas como a [PNAD-COVID19 \(2020\)](#), disponibilizam insumos suficientes para uma estimativa da VSL, que será detalhada no capítulo 4.

Durante a pandemia da COVID-19, o Presidente da República, Jair Bolsonaro, apresentou, constantemente, posturas que demonstram a atribuição de um baixo valor à vida. A proposta de imunidade coletiva, discursos que minimizavam a gravidade do coronavírus e a constante negação à vacinação em massa são atitudes que, comprovadamente, levam a um maior número de mortes e risco de descontrole da pandemia ([FONTANET; CAUCHEMEZ, 2020](#)). As declarações do Presidente da República durante o período pandêmico, assumindo posturas negacionistas, demonstram seu potencial em subestimar a valoração da vida. Como posto anteriormente, a VSL é determinada a partir do esforço coletivo em salvar uma vida, deste modo, é de se ponderar um possível efeito causal existente entre a estimativa da VSL e a preferência eleitoral, representada pelo resultado das eleições de 2018, ou seja, se o indivíduo observado, neste caso um Estado brasileiro, votou majoritariamente no candidato eleito ou no seu adversário. Considerando, portanto, a influência do Governo Federal sobre a atitude da população com relação às medidas sanitárias durante a pandemia da COVID-19 ([AJZENMAN; CAVALCANTI; MATA, 2020](#)), esta monografia coloca como hipótese basilar, a de que existe correlação significante entre a mensuração do valor estatístico da vida e a preferência eleitoral em cada Estado brasileiro.

Definem-se então dois objetivos para esta monografia. O primeiro deles diz respeito à estimativa da VSL para os Estados brasileiros durante o período pandêmico entre maio e novembro de 2020, com base nos dados disponibilizados pelo IBGE na pesquisa [PNAD-COVID19 \(2020\)](#). O segundo objetivo é determinar a relação causal entre a preferência eleitoral dos Estados brasileiros e a estimativa da VSL realizada por esta mesma monografia.

No capítulo 2 são explorados pontos relevantes sobre o valor estatístico da vida, como suas implicações e problemas de estimativa. Também há uma revisão da literatura de modo a direcionar esta monografia ao método de estimativa escolhido. O capítulo 3 explora o método DAG, sigla em inglês para grafos acíclicos dirigidos, como alternativa para a compreensão da inferência causal entre as variáveis. Durante o capítulo ainda é feita uma breve revisão da literatura. No capítulo 4 é apresentada a metodologia de cálculo da VSL, considerando os dados disponibilizados pela [PNAD-COVID19 \(2020\)](#). O capítulo 5 descreve as bases de dados utilizadas nesta monografia, de modo a se compreender a construção das variáveis e do modelo DAG². O capítulo 6, por sua vez, traz os resultados das estimativas da VSL, modeladas no capítulo 4, e da

²DAG: *Directed Acyclic Graph*

inferência causal desenvolvida no capítulo 5. Por fim, o capítulo 7 traz a conclusão do estudo, à luz das discussões expostas nos capítulos anteriores, assim como os resultados obtidos.

2 O Valor Estatístico da Vida

A estimativa do valor estatístico da vida é de suma importância para formuladores de políticas públicas em setores como o da saúde e de infraestrutura. A VSL é utilizada como parâmetro para que as decisões dos formuladores de políticas sejam bem embasadas. Assim, um político pode escolher entre empregar recursos para reformar uma via pública e diminuir o risco de um acidente fatal ou manter a via como está.

Existe uma ampla literatura econômica acerca da VSL, que tem auxiliado governos a tomar decisões e desenvolver projetos levando em conta o caráter analítico de custo-benefício que a VSL intrinsecamente possui. [Miller \(2000\)](#) expõe os principais e mais utilizados métodos para estimar a VSL, sendo eles: (1) pesquisas com os indivíduos, perguntando explicitamente qual valor estes dão à vida, (2) estudos sobre o comportamento de consumidores e (3) estudos de cunho custo-benefício, como por meio da análise do *tradeoff* entre risco e salário.

O método 1, embora útil, encontra dois problemas que podem gerar viés à VSL. Primeiro, o custo para se realizar uma pesquisa de modo a abranger um grupo que exclua o viés de seleção é muito alto e pode inviabilizar o cálculo para um subgrupo. O segundo problema seria o viés da resposta dos indivíduos, que podem não dar a resposta que de fato expressa a sua preferência, haja visto que valorar a própria vida não é uma atividade do cotidiano.

Já o método 2 possui uma implicação de maior grau no que diz respeito a captar a preferência das pessoas. [Garbacz \(1989\)](#) exemplifica esta abordagem ao estimar a valoração dos indivíduos à vida por meio da preferência por comprar um detector de fumaça. O estudo de [Garbacz \(1989\)](#) observa o *tradeoff* dos consumidores entre gastar com um detector de fumaça e a redução que este dispositivo gera na probabilidade de ocorrer uma fatalidade. Outro exemplo desta abordagem é exemplificado pelo estudo de [Ashenfelter e Greenstone \(2004\)](#), no qual se utilizam de uma alteração na velocidade de uma via para descobrir qual a relação entre o valor ganho pelo menor tempo de viagem e o número de acidentes que ocorreram. Os autores se utilizam do aumento da velocidade nas rodovias norte-americanas em 1987 e ponderam qual o ganho no tempo de viagem salvo com relação às vidas perdidas nas rodovias.

O método 3, por sua vez, é o mais utilizado nos estudos que buscam estimar a VSL, em particular, a estimativa por meio da análise de risco-retorno. Neste tipo de estudo, o autor busca entender, por exemplo, a relação entre o ganho marginal no salário e o aumento na chance de sofrer um acidente no trabalho. Deste modo, o autor é capaz de captar o *tradeoff* entre ganhar um salário maior e estar exposto a um risco maior, dado que um risco maior é acompanhado de uma maior remuneração. Estudos como [Pereira, Almeida e Oliveira \(2020\)](#) e [Black e Kniesner \(2003\)](#) exploram esse tipo de análise por meio de regressões lineares, neste caso chamadas de regressões dos salários hedônicos, de modo a controlar fatores que possam gerar viés na

estimação e possibilitando, inclusive, a estimação da VSL para subgrupos específicos.

Existem outras estratégias para a estimação da VSL que são, em alguma medida, uma releitura dos métodos expostos anteriormente. Contudo, todos estes estão sujeitos a problemas intrínsecos ao conceito da VSL. Um destes conceitos a ser posto em discussão é o de observações da VSL com valores negativos. [Doucouliagos, Stanley e Giles \(2012\)](#) exploram essa discussão a medida em que dizem que valores negativos da VSL são contraintuitivos e apontam para uma visão econômica sem sentido. Esses valores podem surgir pela presença de viés de seleção e que, por vezes, passam sem serem reportados ([DOUCOULIAGOS; STANLEY; GILES, 2012](#)). Estudos como [Black e Kniesner \(2003\)](#) exemplificam casos em que há observações negativas da VSL. Sendo assim, é comum que pesquisadores se concentrem nos valores positivos da VSL, dado um valor também positivo no coeficiente de probabilidade de morte da regressão de salários hedônicos.

Quando se trata da VSL é importante salientar, como faz [Colmer \(2020\)](#), que a sua estimação não mensura o valor de uma vida, mas sim quanto um indivíduo está disposto a pagar por uma redução marginal na probabilidade de morrer. O autor ainda traz à discussão o conceito de que a VSL só pode ser considerada em variações pequenas na probabilidade de perder a vida. [Colmer \(2020\)](#) aponta para a diferença entre a disposição a pagar pela redução na probabilidade de ocorrer uma fatalidade e a disposição a aceitar compensações na redução do risco de morte. Para o autor, utilizar a disposição a pagar é válido para os casos em que não se possui a propriedade do bem em questão, sendo verdadeiro o oposto para quando é analisado a disposição a aceitar. [Colmer \(2020\)](#) ainda ressalta que quando se analisam grandes reduções na probabilidade de uma fatalidade, a melhor métrica a se utilizar seria a disposição a pagar.

Outro ponto importante na discussão sobre a VSL, que é abordado na maioria dos estudos, é a heterogeneidade que os dados podem apresentar. Dada a estimação pela regressão de salários hedônicos, existem alguns fatores que podem ser responsáveis pela heterogeneidade entre as estimativas da VSL, como por exemplo a renda ([DOUCOULIAGOS; STANLEY; GILES, 2012](#)). Outro fator que pode gerar heterogeneidade é apontado por [Colmer \(2020\)](#), sendo a disponibilidade a pagar condicionada à idade. Estes tipos de heterogeneidade podem ser aferidos em estimativas da VSL em grupos específicos, por isso em alguns estudos como [Kniesner et al. \(2012\)](#), que consideram efeitos e observações médias da VSL, essas complicações não compõem parcela importante da discussão do autor.

Compreendidos os métodos de estimação e suas consequências, vale ressaltar a relevância dos estudos empíricos que interpretam as estimativas da VSL, sobretudo no que diz respeito à dimensão que podem tomar. Estudos como [Miller \(1990\)](#) se utilizam de outros artigos como [Viscusi \(1981\)](#) e [Leigh \(1987\)](#), que estimaram a VSL das mais diversas maneiras, para criar uma base de comparação entre as estimativas da VSL. O autor até mesmo elenca as prováveis fontes

de viés advindas dos demais estudos, como a idade da população e a renda.

A fim do desenvolvimento desta monografia, é importante destacar as bases de comparação para o valor estatístico da vida no Brasil. O estudo de Pereira, Almeida e Oliveira (2020) faz uso de dados do Registro Anual de Informações Sociais (RAIS) e o Anuário Estatístico de Acidentes de Trabalho (AEAT) para realizar a regressão de salários hedônicos. Os autores consideram os subgrupos de trabalhadores e estimaram uma VSL para homens de R\$ 2,4 milhões e para mulheres R\$ 1,1 milhão.

Outro estudo que apresenta estatísticas para a VSL no contexto brasileiro é Cardoso e Dahis (2020). Como exposto anteriormente, a estimativa do valor estatístico da vida considera, sobretudo, variações marginais no risco de ocorrer alguma fatalidade, sendo as extrapolações do valor estatístico da vida realizadas de modo linear. Dado estes fatores, Cardoso e Dahis (2020) buscam um método para estimar a VSL por meio do benefício gerado com o distanciamento social, de modo que sejam então consideradas grandes variações no valor estatístico da vida. Os autores estimam para o Brasil uma VSL, levando em conta a média dos salários, de US\$ 2,7 milhões, que correspondem a R\$ 15,2 milhões em outubro de 2021. A diferença de valores para a VSL apresentada pelos estudos está condicionada a um fator extremamente relevante, que é a diferença na percepção de risco entre os dois estudos. Este fato ocorre, pois, Cardoso e Dahis (2020), diferentemente de Pereira, Almeida e Oliveira (2020), consideram uma alta variação no risco de mortalidade.

3 Avaliando o Efeito Causal

DAG, sigla em inglês para grafo acíclico dirigido, é um método para determinar inferências causais. Embora os modelos de DAG tenham sido inicialmente introduzidos ainda na década de 1920 por [Wright \(1928\)](#), o método não ganhou robustez a ponto de desenvolver uma discussão pujante. Contudo, recentemente, [Pearl et al. \(2000\)](#) resgatou o debate dos DAGs e seu processo de inferência causal por meio de análise gráfica. Deste modo, outros estudos foram capazes de introduzir a implicação dos DAGs no debate econômico, mais especificamente na econometria, como [Cunningham \(2020\)](#), [Matsueda \(2012\)](#), [Heckman e Pinto \(2015\)](#).

O modelo DAG consiste em um conjunto de nós e setas, de modo que cada nó representa uma variável e cada seta que interliga os nós, indica a relação de causalidade entre as duas variáveis. Um dos objetivos do DAG é facilitar a compreensão de como os indivíduos concebem relações causais ([IMBENS, 2020](#)), assim é possível desenvolver grafos que auxiliam na estratégia de estimativa. Desta forma, é possível encontrar variáveis que possam causar viés, permitindo assim, que o modelo regressivo seja controlado para tais variáveis.

Embora explicitado no nome do modelo, vale enfatizar que os DAGs possuem duas restrições: (1) não é possível haver grafos cíclicos e (2) as setas inferindo causalidade não podem ser bidirecionais, mas sim dirigidas. Se os casos anteriores não forem cumpridos, nenhum modelo DAG será capaz de captar a causalidade entre as variáveis, devendo assim se recorrer a outros métodos que lidem com simultaneidade. Outro ponto de distinção dos modelos DAG, fazendo com que se diferencie dos demais métodos de inferência causal, é o fato do DAG ser uma representação teórica do conhecimento sobre o fenômeno sendo estudado ([CUNNINGHAM, 2020](#)). Assim, é possível compreender a ideia de que um DAG é, de fato, edificado em hipóteses colocadas pelo indivíduo ou grupo que está desenvolvendo o modelo.

Deve-se aindaressaltar a importância de dois conceitos relevantes para o desenvolvimento dos DAGs, o de variáveis colisoras e confusoras. Em um estudo que pretende determinar o efeito causal de uma determinada variável A em B, é possível classificar uma terceira variável X como colisora se esta for causada por ambas as variáveis A e B. Neste caso de variáveis colisoras, não é possível determinar o efeito causal que A tem em B, pois a presença da variável X implica uma dependência espúria entre as variáveis. [Angrist e Pischke \(2010\)](#) ressaltam este caso de correlação espúria e classificam a variável como um controle ruim. Usando o exemplo anterior, cujo objetivo seria de conceber o efeito causal de A em B, a terceira variável X seria classificada como confusora, se estivesse causando ao mesmo tempo A e B. Diferentemente das variáveis colisoras, as confusoras não implicam em correlações espúrias e se for condicionada ao modelo, então a estimativa ocorre sem o viés da variável confusora. Este tipo de variável nada mais é do que como um controle simples, utilizado em modelos econométricos mais básicos, onde buscam-se variáveis que possam gerar viés de omissão, por exemplo.

Como posto no capítulo 1, este estudo se utilizará do método DAG como artifício para analisar a relação causal entre a VSL e a preferência de candidato nas eleições de 2018, a qual será definida por meio do resultado do segundo turno. A hipótese inicial é a de que a VSL, durante a pandemia, sofreu um efeito causal da preferência dos eleitores e que esta não estava ligada a outros fatores que pudessem causar simultaneidade. A ampla divisão da população instigada pelo Governo Federal, baseando-se no combate ao desenvolvimento de vacinas e irrestrita divulgação de procedimentos farmacológicos ineficientes, são fatores que indicam a seta dirigida, ou seja, em apenas um sentido como ilustrado na figura a seguir:



Figura 1 – Relação Causal: Preferência Eleitoral e VSL

O capítulo 5 explora as demais variáveis e suas estatísticas descritivas, a fim de desenvolver o grafo completo para a análise DAG.

4 Mensurando o Valor Estatístico da Vida

Como exposto no capítulo 1, o primeiro objetivo desta monografia é estimar o valor estatístico da vida para os Estados brasileiros entre maio e novembro de 2020, durante a pandemia da COVID-19. Neste caso, para desenvolver a estimativa da VSL, utilizaram-se dados da [PNAD-COVID19 \(2020\)](#), realizada entre os meses de maio e novembro e disponibilizados pelo IBGE.

A abordagem para a estimativa da VSL seguirá a lógica da preferência dos indivíduos, como em [Garbacz \(1989\)](#) e [Ashenfelter e Greenstone \(2004\)](#). A partir da hipótese de [Black \(1948\)](#) de que o governante eleito é o que melhor representa as preferências do eleitor mediano, será considerada nessa monografia que os governadores dos Estados são representativos da vontade do eleitor mediano e, por consequência, da maioria.

Durante o período pandêmico ficou a cargo dos governos estaduais e municipais, segundo a decisão do Supremo Tribunal Federal³, de estabelecer *lockdowns* e restringir horários de circulação e trabalho. Dadas as medidas de restrição de circulação, e considerando os demais choques de oferta por conta da pandemia, ficou evidente o impacto no produto nacional. Em meio à pandemia os governadores tiveram que tomar escolhas no que diz respeito ao *tradeoff* entre geração de produto e controle da disseminação do coronavírus. A estimativa da VSL nesta monografia buscará um valor médio a cada observação, de modo que questões como a heterogeneidade na estimativa não seja um problema trivial, como exposto por [Colmer \(2020\)](#).

Com o intuito de filtrar a base da [PNAD-COVID19 \(2020\)](#), de modo a obter a melhor estimativa possível, foram tomadas as seguintes hipóteses: (1) indivíduos que não declararam rendimentos foram retirados da base, (2) indivíduos que declararam que usualmente tinham rendimentos positivos e efetivamente não tiveram rendimentos na semana da pesquisa foram excluídos da base. Esses dois filtros têm o intuito de manter apenas as pessoas ativas no mercado de trabalho, de modo que as pessoas que tiveram rendimentos indo a zero foram consideradas como excluídas do mercado e não contabilizam o cálculo da VSL. Os cálculos realizados na utilização da [PNAD-COVID19 \(2020\)](#) utilizaram-se dos pesos de contagem, a fim de efetivamente considerar o conjunto da população que aquela observação representa.

A fim de estimar o valor médio da VSL, será considerado que os indivíduos que passaram a ter rendimentos menores do que o usual, foram atingidos pelas restrições impostas pelos governadores de seus respectivos Estados, os quais representam a preferência da maioria. Assim, pela ótica da renda, o produto deixado de ser gerado representa o dispêndio da população em detrimento da redução da disseminação do coronavírus, ou seja, a defasagem entre os valores

³BRASIL. Supremo Tribunal Federal. Ação Direta de Inconstitucionalidade nº6341. **ADI 6341**. Partido Democrático Trabalhista (PDT). Relator: Ministro Marco Aurélio. 23 mar. 2020. Disponível em: <http://portal.stf.jus.br/processos/detalhe.asp?incidente=5880765> . Acesso em: 20 nov. 2021.

usualmente e efetivamente recebidos apontam para o quanto esse grupo de pessoas estaria disposta a abrir mão em troca de salvar mais vidas.

Tanto no cálculo dos salários hedônicos, como nos demais métodos para estimar a VSL, considera-se a redução na probabilidade de morrer. Neste caso, tal percepção é expressa pela redução do número de mortos por COVID-19, dado que esta redução explicita a preferência da população. Ademais, esta monografia não se utiliza da variação percentual de uma fatalidade. Este tipo de análise é inviável pelo fato de que a pandemia apresentou picos no número de mortos por COVID-19, implicando assim em uma alta variação no risco de morte. Um dos fatores que colaboraram para a alta variação no risco de ocorrer uma fatalidade é a própria natureza do vírus de, por exemplo, impor maior índice de contágio em massas populacionais mais densas ([REX; BORGES; KÄFER, 2020](#)).

Deste modo, a estimativa da VSL, a cada mês, é obtida pela relação entre a diferença entre os rendimentos usual e efetivo, multiplicado por 4, dado que a pergunta do IBGE faz referência ao rendimento semanal, e a variação do número de mortos por COVID-19 entre o mês observado e o anterior extraídos do [DATASUS \(2020\)](#):

$$VSL_{i,t} = \frac{(RendimentoEfectivo_{i,t} - RendimentoUsual_{i,t}) * 4}{Mortos_{i,t} - Mortos_{i,t-1}} \quad (1)$$

A equação 1, portanto, apresenta o modelo de estimativa da VSL que considera a diferença nos rendimentos e a variação no número de mortos entre o mês anterior e o mês atual. Os subscritos i e t representam, respectivamente, o Estado brasileiro e o mês observado. Com isso, o modelo captura a preferência dos indivíduos no que diz respeito ao dispêndio monetário em detrimento da variação no número de mortos por COVID-19.

5 Descrição dos Dados e Modelo Causal

5.1 VSL

Como exposto no capítulo anterior, os dados para cálculo da VSL advêm da pesquisa [PNAD-COVID19 \(2020\)](#), realizada pelo IBGE entre os meses de maio e novembro de 2020, e da base de dados do [DATASUS \(2020\)](#) que compila o número de mortos por COVID-19. As manipulações na base de dados da [PNAD-COVID19 \(2020\)](#) foram explicitadas no capítulo 4 e levam à granularidade estadual nas observações. Os dados do [DATASUS \(2020\)](#) também foram agregados ao nível estadual.

5.2 Modelo DAG

A fim de construir o grafo completo e, consequentemente, definir a relação causal entre a VSL e a preferência eleitoral, faz-se necessário acrescentar mais variáveis ao modelo. A primeira dessas é a variação, em valores absolutos, do número de leitos disponíveis em cada observação. Esta variação evidencia a preferência dos indivíduos com relação à disponibilidade de aumentar o gasto público com saúde frente ao risco eminente da COVID-19. Pelo modelo DAG, esta variável de variação no número de leitos tem efeito causal apenas na VSL. Os dados de número de leitos em cada Estado por mês foram obtidos pelo [DATASUS \(2020\)](#), sendo que a variação entre os meses, diz respeito à diferença simples das observações.

Uma segunda variável é a idade média em cada observação, considerando que cada observação diz respeito a um Estado brasileiro e a um mês entre maio e novembro de 2020. A variável idade busca outro perfil da preferência dos indivíduos, que é intrínseca a questões geracionais. No modelo DAG a variável idade tem efeito causal sobre a variação nos leitos, na preferência eleitoral dos indivíduos e sobre a VSL. A terceira variável relevante para o modelo é o rendimento dos indivíduos no período. Esta variável tem como intuito captar outro perfil de preferência dos indivíduos, porém com o caráter econômico, ou seja, a variável tenta incorporar ao modelo a influência da riqueza gerada pelos indivíduos. Seguindo o modelo DAG, a variável rendimento tem efeito causal sobre a VSL e a preferência eleitoral. Ambas as variáveis, idade e rendimento, foram obtidas por meio da [PNAD-COVID19 \(2020\)](#), divulgada pelo IBGE entre os meses de maio e novembro de 2020. Idade e rendimento correspondem, respectivamente, às variáveis A002 e C011A12.

A quarta variável confere uma perspectiva demográfica à preferência dos indivíduos. Para alcançar tal objetivo utilizou-se de dados da [PNAD-COVID19 \(2020\)](#) para estimar a população em cada observação e dados de área dos Estados, obtidos pela base do [IBGE \(2020\)](#), para calcular a densidade demográfica. Apesar de não captar a distribuição espacial dos indivíduos dentro de cada território, esta variável será utilizada a fim de conceber ao modelo uma noção demográfica

da preferência das pessoas. No que diz respeito ao modelo DAG, a variável densidade demográfica tem efeito causal sobre a VSL, a variação do número de leitos e a idade de cada observação.

A última variável, a qual também é a variável de interesse no segundo objetivo da monografia, é a preferência eleitoral. Essa variável é representada por uma variável *dummy*, que assume o valor 1 caso o Estado em questão tenha votado majoritariamente no Presidente Jair Bolsonaro no segundo turno das eleições presidenciais de 2018 e assume o valor 0 caso contrário. Estes valores foram derivados a partir de dados do [ELEIÇÕES2018 \(2018\)](#).

Vale ressaltar que, diferentemente do cálculo da VSL, os demais valores obtidos por meio da [PNAD-COVID19 \(2020\)](#) foram realizados para a totalidade de cada Estado e não considerou os filtros utilizados no cálculo do valor estatístico da vida. Com isso, o modelo DAG para o problema apresentado nesta monografia é o representado pela figura a seguir:

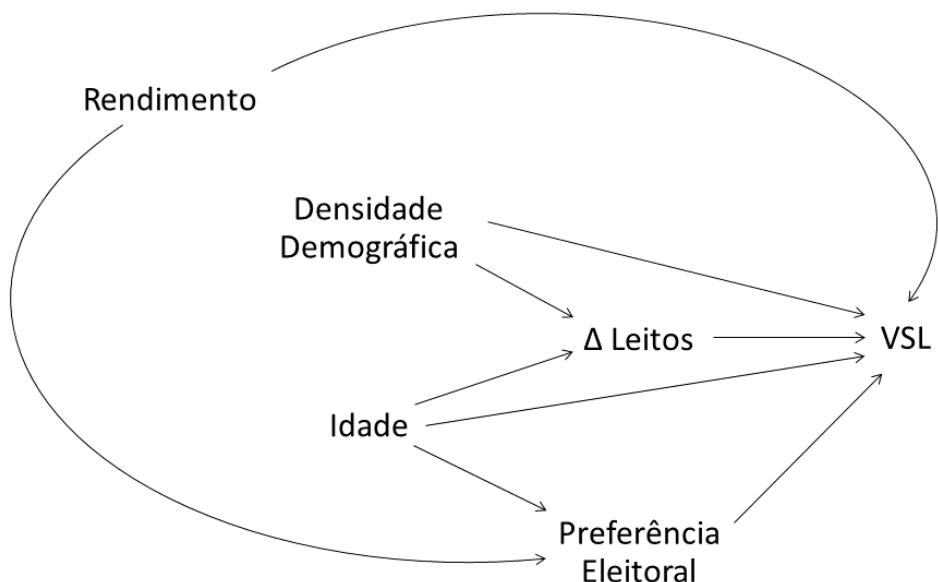


Figura 2 – Modelo DAG

No que diz respeito à base de dados, optou-se por desenvolver dois filtros. O primeiro filtro elimina os *outliers* da VSL, ou seja, observações da VSL que estejam a mais do que dois desvios-padrão da média. Já o segundo filtro exclui da base de dados os valores negativos encontrados para a VSL. Esta monografia considerará, em todas as suas estimativas, quatro bases possíveis, sendo:

- Base 1: Considera todos os valores disponíveis;
- Base 2: Filtra da base 1 as observações que apresentam valores *outliers* da VSL;
- Base 3: Filtra da base 1 as observações que apresentam valores negativos da VSL;

- Base 4: Filtra da base 1, inicialmente, as observações que apresentam valores negativos da VSL e, em seguida, as observações *outliers* da VSL.

6 Resultados

6.1 VSL

De acordo com o modelo desenvolvido no capítulo 4, os valores estimados para a VSL são os seguintes:

Estado	Maio	Junho	Julho	Agosto	Setembro	Outubro	Novembro
Rondônia	-3.20	-2.29	52.03	4.29	4.61	2.39	-60.74
Acre	-1.06	-1.22	1.73	1.03	2.64	4.94	12.06
Amazonas	-0.69	1.05	2.27	10.26	-4.62	4.31	12.92
Roraima	-0.87	-0.99	-85.17	0.59	3.75	2.78	32.86
Pará	-0.68	2.77	1.45	4.90	79.19	5.84	632.06
Amapá	-0.97	-21.20	2.50	1.71	1.30	8.72	-4.31
Tocantins	-6.23	-7.36	-6.00	-2.63	6.64	1.59	2.10
Maranhão	-1.30	-3.23	8.10	1.32	5.06	42.51	6.60
Piauí	-3.45	-1.23	-3.20	2.03	1.61	48.84	4.44
Ceará	-0.68	-3.31	1.08	1.65	8.38	3.81	10.79
Rio Grande do Norte	-4.58	-1.74	-13.69	1.71	2.83	-10.32	5.75
Paraíba	-3.39	-2.53	-3.94	3.70	2.26	5.41	5.49
Pernambuco	-1.60	13.17	8.44	3.05	4.65	4.82	-33.70
Alagoas	-1.13	-2.05	4.27	1.61	1.94	13.59	3.43
Sergipe	-4.41	-1.57	-2.24	1.34	1.74	44.58	5.12
Bahia	-8.09	-6.16	-8.04	-9.80	4.89	4.71	9.80
Minas Gerais	-79.81	-16.54	-6.80	-9.53	11.68	13.53	7.00
Espírito Santo	-4.79	-3.63	11.57	5.27	5.77	14.07	-8.91
Rio de Janeiro	-2.52	-40.46	5.93	9.25	42.98	20.04	48.10
São Paulo	-8.42	-14.25	-23.22	17.26	14.37	7.53	13.00
Paraná	-455.47	-16.68	-6.90	-54.35	28.34	8.85	-17.10
Santa Catarina	-68.35	-32.13	-4.88	-5.34	3.38	7.83	-4.54
Rio Grande do Sul	-58.38	-29.46	-7.01	-20.18	26.11	12.22	-3614.88
Mato Grosso do Sul	-297.85	-16.24	-2.89	-3.39	8.21	2.83	1.52
Mato Grosso	-25.09	-1.76	-1.34	3.59	2.60	2.96	5.75
Goiás	-50.21	-12.69	-3.30	-7.91	-32.62	3.86	3.98
Distrito Federal	-11.69	-4.45	-2.98	-7.00	3.49	1.81	2.48

Tabela 1 – Estimativas do valor estatístico da vida (em milhões de reais)

Como é possível observar na tabela 1, a estimativa da VSL produziu valores negativos e *outliers*, como era esperado, segundo a discussão do capítulo 2. Como é possível observar, cada observação da VSL corresponde a um Estado brasileiro entre os meses de maio e novembro de 2020. A mediana das observações tem valor igual a R\$ 1,5 milhões, valor este que se encontra entre as estimativas para homens e mulheres de Pereira, Almeida e Oliveira (2020). No que diz respeito à média, o valor obtido é negativo, de modo que não tenha significância econômica no contexto desta monografia. A estimativa, contudo, é distante da encontrada por Cardoso e Dahis (2020), fato que pode ser ocasionado pelas observações com valores negativos e pelos *outliers*. Ao analisar os dados da base de dados que exclui ambos os valores negativos e *outliers*

da VSL, têm-se que o valor mediano das estimativas da VSL é de R\$ 4,4 milhões e a média de R\$ 5,6 milhões, valores distantes dos encontrados por [Pereira, Almeida e Oliveira \(2020\)](#). Contudo, estes valores são convenientes com a expectativa de valores maiores, como é discutido por [Cardoso e Dahis \(2020\)](#), dada a alta variabilidade no risco da ocorrência de uma fatalidade.

6.2 Modelo DAG

Os dados obtidos para o desenvolvimento desta monografia consideraram diversos aspectos da preferência dos indivíduos que são parâmetros para as estimativas do valor estatístico da vida. Contudo, seguindo o modelo DAG e os fatores descritos no capítulo 4, é possível identificar que as variáveis rendimento e idade são variáveis confusoras, de modo que necessitam ser controladas no modelo DAG. A variação do número de leitos, por ter impacto na VSL e ser impactada pela idade (variável confusora), também entra no modelo DAG. A densidade demográfica, por sua vez, não é incorporada no modelo DAG, haja visto que, apesar de ter impacto na VSL, a variável não tem impacto direto ou indireto na preferência eleitoral.

No que diz respeito aos dados que serão utilizados no modelo DAG, tem-se as seguintes estatísticas descritivas:

	Rendimento	VSL	Idade	Eleição	Leitos	Densidade
Min	2.744e+08	-3.615e+09	28.34	0.0000	-769.00	2.461
1st Qu.	1.218e+09	-4.580e+06	31.75	0.0000	-48.00	6.963
Median	2.821e+09	1.516e+06	33.54	1.0000	6.00	36.309
Mean	5.634e+09	-2.038e+07	33.05	0.5926	27.24	76.535
3rd Qu.	5.859e+09	5.122e+06	34.80	1.0000	81.00	88.070
Max	4.900e+10	6.321e+08	37.56	1.0000	1070.00	532.381

Tabela 2 – Estatísticas descritivas: Base 1

Sobre as estatísticas descritivas da base 1, é possível observar que a mediana das estimativas da VSL é de R\$ 1,5 milhões, valor coerente com as estimativas de [Pereira, Almeida e Oliveira \(2020\)](#). As demais variáveis apresentam os valores esperados para suas estatísticas descritivas.

	Rendimento	VSL	Idade	Eleição	Leitos	Densidade
Min	2.744e+08	-17098681	28.38	0.0000	-769.0	2.461
1st Qu.	1.148e+09	-3284790	31.70	0.0000	-44.0	6.941
Median	2.592e+09	1630256	33.51	1.0000	9.5	36.283
Mean	5.160e+09	842005	32.83	0.5506	28.3	74.651
3rd Qu.	4.629e+09	4868804	34.12	1.0000	82.5	88.131
Max	4.900e+10	17256977	37.54	1.0000	1070.0	532.381

Tabela 3 – Estatísticas descritivas: Base 2

No caso da base 2, a exclusão das observações que possuíam estimativas *outliers* da VSL colaboraram para que as estatísticas descritivas da variável, como esperado, sofressem alterações nos seus valores extremos, porém sem influenciar muito na média. O destaque fica para a média da VSL, que passa a ser positiva, diferentemente da média na base 1. Como esperado, o número de observações da base 2 é diferente da base 1. Enquanto a base 1 possui 189 observações, a base 2 possui 158.

	Rendimento	VSL	Idade	Eleição	Leitos	Densidade
Min	3.258e+08	592013	28.41	0.0000	-769.00	2.482
1st Qu.	1.211e+09	2549022	30.67	0.0000	-76.50	6.942
Median	2.712e+09	4885956	33.21	1.0000	-11.00	26.437
Mean	5.633e+09	15426783	32.71	0.5437	-21.37	76.985
3rd Qu.	4.659e+09	10031639	34.16	1.0000	50.50	92.921
Max	4.900e+10	632057589	37.54	1.0000	449.00	532.381

Tabela 4 – Estatísticas descritivas: Base 3

A base 3, por sua vez, apresenta valores mais próximos do que se esperava quando são observadas as estatísticas descritivas da VSL. Embora a distância continue alta entre os extremos, os valores centrais são intuitivamente mais concretos. Pode-se considerar que, como as variações de risco são maiores durante a pandemia, esperam-se valores da VSL maiores, o que de fato ocorre nestas observações. O valor médio da VSL, de R\$ 15,4 milhões, está em linha com o que foi estimado por [Cardoso e Dahis \(2020\)](#). Dados que a base 3 filtra da base 1 as observações da VSL com valores negativos, o número de observações não é o mesmo. Neste caso, a base 3 possui 103 observações.

	Rendimento	VSL	Idade	Eleição	Leitos	Densidade
Min	3.258e+08	592013	28.41	0.0000	-769.00	2.482
1st Qu.	1.215e+09	2362645	30.68	0.0000	-80.00	6.019
Median	2.664e+09	4375452	33.36	1.0000	-9.50	31.381
Mean	5.566e+09	5558116	32.65	0.5435	-27.02	74.692
3rd Qu.	4.522e+09	7894974	34.13	1.0000	49.25	90.657
Max	4.900e+10	20039515	37.54	1.0000	449.00	532.381

Tabela 5 – Estatísticas descritivas: Base 4

Por fim, a base 4, assim como a base 3, apresenta valores da VSL mais coerentes com o que se espera de um alto risco de fatalidade, de acordo com a discussão do capítulo 2. Neste caso a média e mediana são mais próximos do que o observado nas bases anteriores, R\$ 5,6 milhões e R\$ 4,4 milhões, respectivamente. Como nas bases 2 e 3, a base 4 não possui o mesmo número de observações que a base 1. No caso da base 4, são excluídas as observações que apresentam valores da VSL negativos e em seguidas são excluídas as observações que possuem valores *outliers* da VSL, deixando a base com 92 observações.

O modelo DAG considerará como relevante para a estimação da VSL: o rendimento, variação no número de leitos, idade e a preferência eleitoral, que é a variável de interesse. Também serão desenvolvidos modelos que consideram todas as variáveis descritas no capítulo 4. Todas as estimativas serão feitas por meio de regressões lineares adotando o método dos mínimos quadrados ordinários (MQO) e considerando as 4 bases de dados disponíveis. O resultado das regressões são apresentados nas tabelas a seguir:

Tabela 6 – Regressões lineares

	Dependent variable: VSL			
	(1)	(2)	(3)	(4)
Idade	-23,936,372.000** (9,583,656.000)	-20,545,358.000** (9,208,080.000)	-620,648.500** (252,235.500)	-561,066.900** (240,267.000)
Eleição	-58,804,583.000 (41,749,927.000)	-54,028,324.000 (41,640,095.000)	-1,425,410.000 (1,043,929.000)	-1,315,544.000 (1,033,228.000)
Leitos	33,940.250 (103,099.300)	60,336.640 (101,081.800)	-12,867.280*** (2,503.705)	-12,579.120*** (2,473.604)
Densidade	229,529.600 (183,132.000)		3,583.222 (4,560.240)	
Rendimento	0.002 (0.003)	0.003 (0.003)	0.0001 *** (0.0001)	0.0001 *** (0.0001)
Constant	774,851,320.000** (308,932,386.000)	674,230,081.000** (298,781,242.000)	21,394,300.000*** (8,147,758.000)	19,616,267.000*** (7,817,437.000)
Observations	189	189	158	158
R ²	0.045	0.037	0.192	0.189
Adjusted R ²	0.019	0.016	0.166	0.168
Residual Std. Error	267,832,771.000 (df = 183)	268,247,955.000 (df = 184)	6,145,256.000 (df = 152)	6,137,567.000 (df = 153)
F Statistic	1.743 (df = 5; 183)	1.781 (df = 4; 184)	7.231 *** (df = 5; 152)	8.906 *** (df = 4; 153)

* p<0.1; ** p<0.05, *** p<0.01

Note:

Tabela 7 – Regressões lineares (continuação)

	Dependent variable: VSL			
	(5)	(6)	(7)	(8)
Idade	-3,069,809.000 (3,204,010.000)	-2,904,768.000 (2,920,483.000)	360,519.300* (192,561.000)	336,406.900* (176,186.100)
Eleição	-16,308,221.000 (13,332,934.000)	-16,064,075.000 (13,130,446.000)	840,262.100 (795,790.500)	802,878.500 (783,039.200)
Leitos	-15,572.060 (41,720.840)	-14,308.980 (40,342.150)	-3,062.691 (2,477.547)	-3,216.360 (2,417.678)
Densidade	8,170.141 (63,588.660)		-1,199.449 (3,760.320)	
Rendimento	0.001 (0.001)	0.001 (0.001)	0.0001 *** (0.00005)	0.0001 *** (0.00005)
Constant	119,809,104.000 (102,416,584.000)	114,831,444.000 (94,329,514.000)	-7,478,162.000 (6,186,613.000)	-6,752,275.000 (5,723,089.000)
Observations	103	103	92	92
R ²	0.024	0.024	0.324	0.323
Adjusted R ²	-0.026	-0.016	0.285	0.292
Residual Std. Error	63,540,632.000 (df = 97)	63,220,992.000 (df = 98)	3,544,673.000 (df = 86)	3,526,326.000 (df = 87)
F Statistic	0.480 (df = 5; 97)	0.601 (df = 4; 98)	8.238*** (df = 5; 86)	10.379*** (df = 4; 87)

*p<0.1; **p<0.05; ***p<0.01

Note:

Os modelos 1, 3, 5 e 7 das tabelas 6 e 7 foram feitos como uma regressão linear múltipla, considerando todas as variáveis que foram exploradas pela base de dados desta monografia e que inferem causalidade na VSL. Por sua vez, os modelos 2, 4, 6 e 8, também das tabelas 6 e 7, consideraram o modelo DAG, excluindo a variável densidade demográfica que foi definida como não significante para a compreensão da variável de interesse, a preferência eleitoral.

Iniciando a análise pelos modelos 1 e 2, que consideram a base 1, tem-se que ambos os modelos apresentam características semelhantes. Tanto o modelo 1 quanto o 2 têm a idade e a constante como significantes para o modelo, ambas têm coeficiente atrelado à preferência eleitoral como efeito negativo sobre a VSL e possuem estatísticas F abaixo do nível de significância a 5%.

Já para os modelos 3 e 4, que utilizam a base 2 para fazer as regressões, também possuem características análogas. Em ambos os modelos as variáveis idade, variação do número de leitos, rendimento e a constante são significativos e possuem coeficientes com a mesma direção. Também vale ressaltar que os dois modelos têm estatística F acima do nível de significância, sendo que o modelo 4, proveniente do DAG, apresenta estatística F e R^2 ajustado maior do que o do modelo 3, que considera todas as variáveis.

Os modelos 5 e 6, derivados da base 3, diferentemente dos modelos anteriores não apresentam nenhuma variável estatisticamente significante ou possuem estatísticas F relevantes. Estes dois modelos corroboram as exposições do capítulo 2 desta monografia, que apontam para a possibilidade de viés que, neste caso, levou os modelos a serem basicamente desprezíveis no que diz respeito à relevância econométrica.

Por fim, os modelos 7 e 8 tomam a base 4 e, embora esta base seja consideravelmente menor do que as demais, ainda apresenta as variáveis idade e rendimento como estatisticamente significantes para o modelo e com estatísticas F acima do nível de 5% de significância. Entretanto, o coeficiente atrelado à variável preferência econômica tem sinal positivo, o que contradiz a lógica inferida pelas demais estimativas, a hipótese basilar desta monografia.

Os coeficientes estimados pelos modelos citados anteriormente representam a sensibilidade da VSL com relação às variáveis utilizadas. Ao analisar a variável de interesse, preferência eleitoral, é possível compreender qual o impacto médio dessa variável sobre a VSL. Assim, no caso do modelo 3, infere-se que os Estados que elegeram o Presidente Jair Bolsonaro possuem uma VSL, em média, R\$ 1,4 milhões menor do que os Estados que não elegeram o presidente. Dado que a variável preferência eleitoral é uma *dummy*, a explicação da magnitude dos coeficientes se estende para os demais modelos derivados. As outras variáveis são contínuas e se encontram em nível. Assim, a compreensão destas é a de que a alteração em uma unidade na variável, impacta a VSL, em média, na dimensão do seu respectivo coeficiente.

Com a derivação dos 8 modelos anteriores fica claro que a omissão dos valores negativos encontrados para a VSL tende a gerar viés de seleção e que, para fins de compreensão dos determinantes da VSL, essas observações são relevantes. Ademais, a base 1, que considera os *outliers* também deve ser tratada com cautela, haja visto que não produziu nenhuma estatística relevante e, provavelmente, esse fato ocorra, pois os *outliers* captam períodos em que a pandemia da COVID-19 estava fora de controle e a resposta dos governadores não foi suficientemente grande para corroborar com a perda de vidas. Este dito descontrole da disseminação do vírus pode ser consequência de diversos fatores, como por exemplo o relaxamento de normas sanitárias por parte das autoridades ou até mesmo um descumprimento de tais normas por parte da sociedade civil.

Por fim, restam as considerações dos modelos que se utilizam da base 2. Estas regressões (3 e 4) apresentam não só variáveis relevantes e têm estatísticas F acima do nível de 5% de significância, mas também possuem a variável preferência eleitoral com o sinal esperado a partir das hipóteses. Há de se pontuar ainda que o modelo 4, que considera o modelo DAG, apresenta estatísticas de significância mais relevantes do que as do modelo 3. Logo, é possível concluir que, de fato, ao analisar o melhor recorte entre as bases de dados e ter a melhor visão do impacto causal entre as variáveis, o modelo mais assertivo no que diz respeito à significância estatística é o que se utiliza do método DAG.

A falta de significância na variável de preferência eleitoral nas regressões leva a acreditar que esta não tem efeito causal sobre a estimativa da VSL. Esta monografia reconhece que a falta de inferência causal pode ser dada por alguma omissão no modelo, especificação não interpretada ou porque, de fato, não há relação causal entre as variáveis. Entretanto, dados estudos prévios como [Ajzenman, Cavalcanti e Mata \(2020\)](#), que apresentam a correlação positiva entre indicadores como maior número de casos e mortes por COVID-19 em Estados brasileiros que elegeram o Presidente Jair Bolsonaro, sustenta-se o questionamento de que é possível haver correlação positiva e significativa entre a VSL e a preferência eleitoral.

7 Conclusão

Embora a concepção da vida como um bem seja complexa, talvez pelo fato de ser intangível, faz com que a valoração também seja de difícil mensuração. O cálculo do valor estatístico da vida é, por vezes, custoso e cercado de questões que o colocam à prova, como a estimativa de valores negativos e da heterogeneidade presente nas amostras. Contudo, certos métodos auxiliam na compreensão da VSL e, como exposto no capítulo 2, existe uma vasta literatura que dá suporte ao desenvolvimento da valoração desse bem normal, porém com características peculiares que é a vida.

A partir de evidências anedóticas e acadêmicas esta monografia buscou entender se existe relação causal entre a preferência eleitoral dos indivíduos e a estimativa da VSL a nível estadual durante o período pandêmico de maio a novembro de 2020. A hipótese basilar desta monografia é a de que a correlação de fato existia e possuía significância estatística.

O método de cálculo da VSL utilizou a diferença de produto gerada dividida pela variação no número de mortos por COVID-19 como *proxy* da preferência dos indivíduos à menor chance de morrer em detrimento da geração de produto. Como as restrições de circulação e serviços essenciais partia, basicamente, dos governos estaduais, e considerando as hipóteses de [Black \(1948\)](#), considerou-se que essas preferências de fato representavam a vontade da maioria da população.

O modelo DAG foi um artifício utilizado para refinar a análise das relações causais que determinam a VSL e, portanto, compreender quais variáveis de fato precisariam ser controladas a fim de definir o efeito causal da preferência política na VSL. A partir da definição do modelo DAG, considerando as demais variáveis relevantes para o modelo, partiu-se para as estimativas por meio do método dos mínimos quadrados ordinários (MQO), tanto para o modelo DAG como para o modelo que avalia a composição de todas as variáveis disponíveis na base de dados.

Cumpridos o primeiro objetivo desta monografia, foi possível realizar a comparação entre os valores das estimativas da VSL, obtidas pelo método adotado por esta monografia, e os valores dos estudos de [Pereira, Almeida e Oliveira \(2020\)](#) e [Cardoso e Dahis \(2020\)](#), observou-se que há coerência nos valores. Quando são filtrados os valores negativos, as estimativas da VSL se aproximaram, em média, dos valores obtidos por [Cardoso e Dahis \(2020\)](#) e quando filtrados os valores *outliers*, a mediana das estimativas foi mais próxima dos valores derivados por [Pereira, Almeida e Oliveira \(2020\)](#).

Ao executar o segundo objetivo, foi possível realizar as estimativas com as 4 bases de dados derivadas nesta monografia, e conclui-se que os valores negativos da VSL de fato são importantes para a compreensão da estimativa média da VSL, uma vez que as bases de dados que

filtram esses valores não tinham a performance semelhante às demais ou não eram conjuntamente significativas. Deste modo, foi possível inferir que a melhor base de dados a ser utilizada foi aquela que filtrou os *outliers* e utilizou-se do modelo DAG.

Ao entender o melhor a base e o modelo a ser utilizado, percebeu-se que de fato o coeficiente atrelado à preferência eleitoral é negativo, ou seja, Estados que votaram, em sua maioria, em Jair Bolsonaro possuíram um menor valor estatístico da vida durante o recorte da pandemia entre maio e novembro de 2020. Entretanto, a variável de interesse, preferência eleitoral, não é estatisticamente significante, como colocava a hipótese inicial desta monografia.

Conclui-se que, apesar das evidências anedóticas e das correlações encontradas entre o número de mortos e apoio ao Governo Federal, a inclinação política de um Estado nas eleições presidenciais de 2018 não possuem um efeito causal sobre o valor estatístico da vida. Este resultado não exime o fato de que haja possibilidade de outro indicador de compatibilidade com o Governo Federal ou até mesmo uma granularidade diferente explice e comprove a hipótese desta monografia.

Referências

- ABRAS, A.; FAVA, A. C. P. e.; KUWAHARA, M. Y. Women heads of state and covid-19 policy responses. *Feminist Economics*, Taylor & Francis, v. 27, n. 1-2, p. 380–400, 2021. Citado na página 8.
- AJZENMAN, N.; CAVALCANTI, T.; MATA, D. D. More than words: Leaders' speech and risky behavior during a pandemic. Available at SSRN 3582908, 2020. Citado 3 vezes nas páginas 8, 9 e 28.
- ANGRIST, J. D.; PISCHKE, J.-S. The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of economic perspectives*, v. 24, n. 2, p. 3–30, 2010. Citado na página 14.
- ASHENFELTER, O.; GREENSTONE, M. Using mandated speed limits to measure the value of a statistical life. *Journal of political Economy*, The University of Chicago Press, v. 112, n. S1, p. S226–S267, 2004. Citado 2 vezes nas páginas 11 e 16.
- BLACK, D. On the rationale of group decision-making. *Journal of political economy*, The University of Chicago Press, v. 56, n. 1, p. 23–34, 1948. Citado 4 vezes nas páginas 8, 9, 16 e 29.
- BLACK, D. A.; KNIESNER, T. J. On the measurement of job risk in hedonic wage models. *Journal of risk and uncertainty*, Springer, v. 27, n. 3, p. 205–220, 2003. Citado 2 vezes nas páginas 11 e 12.
- CARDOSO, D. S.; DAHIS, R. Value of a statistical life under large mortality risk change: Theory and an application to covid-19. Available at SSRN 3599529, 2020. Citado 5 vezes nas páginas 13, 21, 22, 23 e 29.
- COLMER, J. What is the meaning of (statistical) life? benefit–cost analysis in the time of covid-19. *Oxford Review of Economic Policy*, Oxford University Press UK, v. 36, n. Supplement_1, p. S56–S63, 2020. Citado 2 vezes nas páginas 12 e 16.
- CUNNINGHAM, S. Causal inference. *The Mixtape*, v. 1, 2020. Citado na página 14.
- DATASUS. Disponível em: <https://datasus.saude.gov.br/informacoes-de-saude-tabnet/>. Acesso em 20 nov. 2021, SUS, 2020. Citado 2 vezes nas páginas 17 e 18.
- DOUCOULIAGOS, C.; STANLEY, T. D.; GILES, M. Are estimates of the value of a statistical life exaggerated? *Journal of health economics*, Elsevier, v. 31, n. 1, p. 197–206, 2012. Citado na página 12.
- ELEIÇÕES2018. Disponível em: <https://www.tse.jus.br/eleicoes/estatisticas/repositorio-de-dados-eleitorais-1>. Acesso em 20 nov. 2021, TSE, 2018. Citado na página 19.
- FONTANET, A.; CAUCHEMEZ, S. Covid-19 herd immunity: where are we? *Nature Reviews Immunology*, Nature Publishing Group, v. 20, n. 10, p. 583–584, 2020. Citado na página 9.
- GARBACZ, C. Smoke detector effectiveness and the value of saving a life. *Economics Letters*, Elsevier, v. 31, n. 3, p. 281–286, 1989. Citado 2 vezes nas páginas 11 e 16.

- HECKMAN, J.; PINTO, R. Causal analysis after haavelmo. *Econometric Theory*, Cambridge University Press, v. 31, n. 1, p. 115–151, 2015. Citado na página 14.
- IBGE. Disponível em: <https://www.ibge.gov.br/geociencias/organizacao-do-territorio/estrutura-territorial/15761-areas-dos-municípios>. Acesso em 20 nov. 2021, IBGE, 2020. Citado na página 18.
- IMBENS, G. W. Potential outcome and directed acyclic graph approaches to causality: Relevance for empirical practice in economics. *Journal of Economic Literature*, v. 58, n. 4, p. 1129–79, 2020. Citado na página 14.
- KNIESNER, T. J. et al. The value of a statistical life: Evidence from panel data. *Review of Economics and Statistics*, The MIT Press, v. 94, n. 1, p. 74–87, 2012. Citado na página 12.
- LEIGH, J. P. Gender, firm size, industry, and estimates of the value-of-life. *Journal of Health Economics*, Elsevier, v. 6, n. 3, p. 255–273, 1987. Citado na página 12.
- MATSUEDA, R. *Handbook of structural equation modeling*. [S.l.]: Guilford Press, 2012. Citado na página 14.
- MILLER, T. R. The plausible range for the value of life—red herrings among the mackerel. *Journal of forensic Economics*, JSTOR, v. 3, n. 3, p. 17–39, 1990. Citado na página 12.
- MILLER, T. R. Variations between countries in values of statistical life. *Journal of transport economics and policy*, JSTOR, p. 169–188, 2000. Citado 2 vezes nas páginas 8 e 11.
- PEARL, J. et al. Models, reasoning and inference. *Cambridge, UK: Cambridge University Press*, v. 19, 2000. Citado na página 14.
- PEREIRA, R. M.; ALMEIDA, A. N. d.; OLIVEIRA, C. A. d. O valor estatístico de uma vida: estimativas para o brasil. *Estudos Econômicos (São Paulo)*, SciELO Brasil, v. 50, p. 227–259, 2020. Citado 5 vezes nas páginas 11, 13, 21, 22 e 29.
- PNAD-COVID19. Disponível em: <https://www.ibge.gov.br/estatisticas/sociais/trabalho/27946-divulgacao-semanal-pnadcovid>. Acesso em 20 nov. 2021, IBGE, 2020. Citado 4 vezes nas páginas 9, 16, 18 e 19.
- REX, F. E.; BORGES, C. A. d. S.; KÄFER, P. S. Spatial analysis of the covid-19 distribution pattern in são paulo state, brazil. *Ciência & Saúde Coletiva*, SciELO Public Health, v. 25, p. 3377–3384, 2020. Citado na página 17.
- VISCUSI, W. K. Occupational safety and health regulation: Its impact and policy alternatives. *Research in public policy analysis and management*, JAI Press Greenwich, CT, v. 2, p. 281–299, 1981. Citado na página 12.
- WRIGHT, P. G. *Tariff on animal and vegetable oils*. [S.l.]: Macmillan Company, New York, 1928. Citado na página 14.

1 Anexo I: Do-file do software R

```
1 #Bibliotecas -----
2 if(!require(devtools)) install.packages("devtools")
3 devtools::install_github('datazoompuc/PNAD_Covid/R/datazoom_pnad_covid')
4 install.packages("plyr")
5 install.packages("data.table")
6 install.packages("dplyr")
7 install.packages("tidyR")
8 install.packages("plm")
9 install.packages("fastDummies")
10
11 library(datazoomPNADcovid)
12 library(plyr)
13 library(dplyr)
14 library(tidyR)
15 library(data.table)
16 library(plm)
17 library(fastDummies)
18 library(readxl)
19 library(stargazer)
20
21 #Bases Prim rias -----
22 x <- pnad_covid_microdata('/Users/mateu/Desktop/Monografia/Bases PNAD COVID
   , lang = 'pt_br', c(5, 2020),c(6, 2020),c(7, 2020),c(8, 2020),c(9,
   2020),c(10, 2020),c(11, 2020))
23 base_pnad <- dplyr::bind_rows(x)
24 chave_estados <- data.frame(UF =
   ('11', '12', '13', '14', '15', '16', '17', '21', '22', '23', '24', '25',
   '26', '27', '28', '29',
   '31', '32', '33', '35', '41', '42', '43', '50', '51', '52', '53'), ,
25
26                               Estado = c('RO', 'AC', 'AM', 'RR', 'PA', 'AP', 'TO', ,
27 'MA', 'PI', 'CE', 'RN', 'PB', 'PE', 'AL', 'SE', 'BA', 'MG', 'ES', 'RJ', 'SP', 'PR', 'SC
   ', 'RS', 'MS', 'MT', 'GO', 'DF'))
28 Area <- read_excel("C:/Users/mateu/Desktop/Monografia/Bases/Area.xlsx")
29 Medicos_por_estado <- read_excel("C:/Users/mateu/Desktop/Monografia/Bases/
   Medicos por estado.xlsx")
30 Leitos <- read_excel("C:/Users/mateu/Desktop/Monografia/Bases/Leitos.xlsx")
31 Dados_Eleicao_2018 <- read_excel("C:/Users/mateu/Desktop/Monografia/Bases/
   Dados - Eleicao 2018.xlsx")
32
33 #C lculo VSL-----
34 base_principal <- base_pnad %>%
35   mutate(horas_efetiva = C009,
36         horas_habitual = C008,
37         rendimento_habitual = C01012,
38         rendimento_efetivo = C011A12) %>%
```

```

38 filter(!is.na(horas_efetiva) & horas_efetiva != 0 & !is.na(horas_habitual
  ) & horas_habitual != 0) %>%
39 group_by(Ano, Mes = V1013,UF) %>%
40 summarise(horas_efetiva = sum(horas_efetiva, na.rm = TRUE)*weighted.mean(
  V1032),
41           horas_habitual = sum(horas_habitual, na.rm = TRUE)*weighted.
  mean(V1032),
42           rendimento_habitual = sum(rendimento_habitual, na.rm = TRUE)*
  weighted.mean(V1032),
43           Rendimento_efetivo = sum(rendimento_efetivo, na.rm = TRUE)*
  weighted.mean(V1032))
44
45 base_principal <- join(base_principal, chave_estados)
46
47 base_mortes <- read.csv("https://raw.githubusercontent.com/wcota/covid19br/
  master/cases-brazil-states.csv")
48 base_mortes <- base_mortes[base_mortes$state != 'TOTAL',]
49 base_mortes$date <- as.Date(base_mortes$date, format = "%Y-%m-%d")
50 base_mortes <- base_mortes %>%
51   separate(date, sep="-", into = c("Ano", "Mes", "Dia"))
52 base_mortes <- setDT(base_mortes)
53 base_mortes <- base_mortes[, list(Obitos = sum(newDeaths)), by = list(
  Estado = state, Ano, Mes)]
54 base_mortes <- join(base_mortes, chave_estados)
55 base_mortes$Mes <- as.numeric(base_mortes$Mes)
56 base_mortes <- base_mortes[order(UF, Ano, Mes), ]
57 base_mortes$delta_obitos <- base_mortes$Obitos - shift(base_mortes$Obitos)
58
59 base_principal <- join(base_principal, base_mortes)
60 base_principal$delta_horas <- base_principal$horas_efetiva -
  base_principal$horas_habitual
61 base_principal$delta_rendimento <- base_principal$Rendimento_efetivo -
  base_principal$rendimento_habitual
62 base_principal$VSL <- (base_principal$delta_rendimento*4)/
  base_principal$delta_obitos
63
64
65 #Tratamento outros dados-----
66 base_dados <- base_pnad %>%
67   mutate(pop = 1,
68         idade = A002) %>%
69   group_by(Ano, Mes = V1013,UF) %>%
70   summarise(Populacao = sum(pop)*weighted.mean(V1032),
71             Idade = weighted.mean(idade,V1032))
72 as.character(base_dados$UF)
73 dados <- join(base_principal, base_dados)
74 dados <- join(dados, Dados_Eleicao_2018)

```

```

75 Leitos <- Leitos[order(Leitos$UF, Leitos$Mes), ]
76 Leitos$Var_Leitos <- Leitos$Leitos - shift(Leitos$Leitos)
77 Medicos_por_estado <- Medicos_por_estado[order(Medicos_por_estado$UF,
    Medicos_por_estado$Mes), ]
78 Medicos_por_estado$Var_Medicos <- Medicos_por_estado$Medicos - shift(
    Medicos_por_estado$Medicos)
79
80 #Base 1-----
81 dados <- join(dados, Medicos_por_estado)
82 dados <- join(dados, Leitos)
83 dados <- dados[-c(4:6)]
84 dados <- dados[-c(6:9)]
85 dados <- join(dados, Area)
86 dados <- dummy_cols(dados, select_columns = 'Mes')
87 dados <- dados[-c(17)]
88 dados$Densidade <- dados$Populacao/dados$Area
89 dados$Rendimento_Medio <- dados$Rendimento_efetivo/dados$Populacao
90 base1 <- dados
91
92 #Base 2-----
93 outliers <- boxplot(dados$VSL, plot = FALSE)$out
94 dados1 <- dados[!(dados$VSL %in% outliers), ]
95 base2 <- dados1
96
97 #Base 3-----
98 dados2 <- dados %>%
99   filter(VSL > 0)
100 base3 <- dados2
101
102 #Base 4-----
103 dados3 <- dados %>%
104   filter(VSL > 0)
105 outliers1 <- boxplot(dados3$VSL, plot = FALSE)$out
106 dados3 <- dados3[!(dados3$VSL %in% outliers1), ]
107 base4 <- dados3
108
109 #Regress es -----
110 #regress es base 1
111 reg1 <- lm(VSL ~ Idade + 'Dummy eleitoral' + Var_Leitos + Densidade +
    Rendimento_efetivo, data = base1)
112 reg2 <- lm(VSL ~ Idade + 'Dummy eleitoral' + Var_Leitos +
    Rendimento_efetivo , data = base1)
113
114 #regress es base 2
115 reg3 <- lm(VSL ~ Idade + 'Dummy eleitoral' + Var_Leitos + Densidade +
    Rendimento_efetivo, data = base2)
116 reg4 <- lm(VSL ~ Idade + 'Dummy eleitoral' + Var_Leitos +

```

```

Rendimento_efetivo , data = base2)
117
118
119 #regress es base 3
120 reg5 <- lm(VSL ~ Idade + 'Dummy eleitoral' + Var_Leitos + Densidade +
    Rendimento_efetivo, data = base3)
121 reg6 <- lm(VSL ~ Idade + 'Dummy eleitoral' + Var_Leitos +
    Rendimento_efetivo , data = base3)
122
123 #regress es base 4
124 reg7 <- lm(VSL ~ Idade + 'Dummy eleitoral' + Var_Leitos + Densidade +
    Rendimento_efetivo, data = base4)
125 reg8 <- lm(VSL ~ Idade + 'Dummy eleitoral' + Var_Leitos +
    Rendimento_efetivo , data = base4)
126
127
128 #tabelas com as regress es
129 stargazer(reg1,reg2,reg3,reg4, float.env = "sidewaystable")
130 stargazer(reg5,reg6,reg7,reg8, float.env = "sidewaystable")

```