



Universidade de Brasília  
Departamento de Estatística

Proposta de nova taxa de inovação de adoção de políticas públicas por meio  
do modelo de regressão de Cox

Carolina Gomes Casado de Carvalho

Relatório final apresentado para o Departamento de Estatística da Universidade de Brasília como parte dos requisitos necessários para obtenção do grau de Bacharel em Estatística.

Brasília  
2022

**Carolina Gomes Casado de Carvalho**

**Proposta de nova taxa de inovação de adoção de políticas públicas por meio  
do modelo de regressão de Cox**

Orientadora: Juliana Betini Fachini Gomes

Relatório final apresentado para o Departamento de Estatística da Universidade de Brasília como parte dos requisitos necessários para obtenção do grau de Bacharel em Estatística.

**Brasília  
2022**

# 1 Resumo

Compreender o processo de difusão de políticas, em busca de identificar motivos das adoções de políticas públicas por grupos de estados e da não adoção por outros, bem como os fatores que influenciam a velocidade das adesões às políticas são questões que norteiam esse estudo.

Para isso os scores de inovação de Walker e a Adoption Rate, propostos por Walker e por Boehmke e Skinner, são revisitados. Estas medidas estabelecem um grau de inovação no sentido de que um Estado ou Município inovador adotaria, pelo menos em média, novas políticas mais cedo do que estados menos inovadores.

Além de revisar essas teorias seminais é proposta uma nova medida baseada na metodologia estatística de Análise de sobrevivência, afim de preencher as lacunas existentes na literatura.

Para validação, os dados manipulados é um subconjunto dos dados utilizados no artigo (BOEHMKE; SKINNER, 2012). Por fim, os resultados das duas medidas já existentes na literatura são confrontados com os resultados obtidos por meio da nova taxa proposta.

**Palavras-chave:** Estatística, difusão, inovação política, escore de inovação, análise de sobrevivência, função de risco, censura, modelo de Cox.



## Sumário

<b>1</b>	<b>Resumo</b>	<b>3</b>
<b>2</b>	<b>Introdução</b>	<b>7</b>
<b>3</b>	<b>Objetivos</b>	<b>11</b>
3.1	Objetivo Geral	11
3.2	Objetivos Específicos	11
<b>4</b>	<b>Revisão de Literatura</b>	<b>13</b>
4.1	Análise de Sobrevivência	13
4.1.1	Definições gerais	13
4.1.2	Funções para o tempo de falha	15
4.1.3	Estimador de Kaplan-Meier	16
4.1.4	Modelo de Regressão de Cox	17
4.1.5	Estimação	18
4.1.6	Avaliação do Ajuste do Modelo de Cox	22
4.2	Medida de inovações relacionadas a adoção de políticas públicas	24
<b>5</b>	<b>Metodologia</b>	<b>28</b>
5.1	Material	28
5.2	Método	30
<b>6</b>	<b>Resultados e Análises</b>	<b>32</b>
6.1	Análise Descritiva	32
6.1.1	<i>Charter Schools</i>	32
6.1.2	<i>High School Exit Exams</i>	40
6.1.3	<i>School Choice</i>	48
6.2	Modelo de Riscos Proporcionais de Cox	55
6.3	Diagnóstico do Modelo	60
6.4	Comparação da Medida de Inovação com o Score de Walker e a Taxa de adoção	63
<b>7</b>	<b>Conclusão</b>	<b>66</b>
	<b>Referências</b>	<b>68</b>



## 2 Introdução

A difusão é apontada por Rogers (1995) como um processo social que ocorre entre as pessoas em resposta ao aprendizado sobre uma inovação. Esta inovação é comunicada por meio de determinados canais ao longo do tempo entre membros de um sistema social. Assim, a variável dependente comumente avaliada em pesquisas de difusão é o tempo até a adoção de determinada política (COÊLHO; CAVALCANTE; TURGEON, 2016).

O conjunto de prós e contras de inovações, as características dos adotantes e o contexto social e político são variáveis essenciais para explicar a difusão ou a falta dela. Neste contexto, as políticas têm sido estudadas há muito tempo como inovação no processo de difusão.

O desenvolvimento do campo de estudos da difusão de políticas públicas teve início em 1969 por Jack Walker, a partir do interesse em entender de que forma ocorria a difusão de políticas entre os estados americanos. Desde então, para melhor compreensão deste processo estudos dentro deste campo buscam identificar os motivos da adoção de políticas públicas por um grupo de países, estados ou municípios e da não adoção por outros, os fatores que influenciam na velocidade da difusão de políticas em diversas áreas, os obstáculos que impedem a adoção da inovação e as facilidades, entre outras questões.

Este debate ao longo do tempo ocasionou o desenvolvimento de diversas abordagens na literatura da Ciência Política. Assim sendo, as principais abordagens teóricas desenvolvidas são: determinantes internos, determinantes externos, ação coordenada, ação descoordenada, agentes, fatores estruturais e mecanismos.

A abordagem dos determinantes internos e externos abrangem visões opostas sobre os fatores que influenciam a difusão de políticas. Autores como Downs (1976), Regens (1980) e Mintrom (1998) defendem que as características locais influenciam diretamente a decisão de adoção ou não de uma nova política. Esta abordagem dos determinantes internos pressupõe que os tomadores de decisão frente a estratégia política local em um cenário de difusão de políticas considerariam os obstáculos da reforma, a necessidade da inovação e o grau de competição política. Por outro lado, autores como Richards (1975) e Light (1978), apoiadores da abordagem dos determinantes externos, argumentam que variáveis como a influência de outros governos, a proximidade geográfica e o papel das redes formais e informais motiva diretamente os processos de difusão. Berry e Berry (1990) unificaram as duas teorias com o argumento de que é impensável supor que as decisões de mudança política são determinadas exclusivamente ou por influências externas ou pelos tributos locais (COÊLHO; CAVALCANTE; TURGEON, 2016).

Além de explicações centradas nos determinantes internos e externos, a literatura desenvolveu enfoques para comparar o papel dos atores-chaves e os mecanismos direcio-

nais que conduzem à adoção de políticas. Uma visão tradicional indica que o processo decisório é baseado nos fatores internos e a influência dos vizinhos é nula, não havendo qualquer tipo de ação coordenada. Uma posição rival estabelece que os avanços das ideias ocorreria em razão da ação coordenada de uma organização ou um grupo de unidades governamentais interessadas em desencadear uma nova política, nesta posição há uma cooperação ou coerção. Elkins e Simmons (2005) combinou elementos das alternativas anteriores e redefiniu a difusão como uma ação descoordenada e interconectada, descoordenada pois os governos são classificados como instituições independentes, uma vez que tomam suas próprias decisões, e interconectada pois levam em conta a decisão prévia de outros governos.

Outras abordagens de destaque nos debates teóricos acerca da difusão levam em consideração o fator estrutural e o papel dos atores na adoção de políticas. Explicações centradas no fator estrutural privilegiam as condições estruturais, bem como características como o resultados histórico de políticas e o aparato institucional de governo em que as políticas serão implementadas, a proximidade geográfica com regiões pioneiras. Defensores dessa corrente apontam que a difusão ocorre sem agentes e que depende exclusivamente de uma estrutura ideal para ocorrer. Já as explicações voltadas no fator agentes, destacam o papel dos agentes como atores que conduzem a difusão de políticas. Na visão de Marsh e Sharmman (2010) as duas dimensões estão associadas, dado que estruturas inibem ou facilitam o papel dos agentes (COÊLHO; CAVALCANTE; TURGEON, 2016).

Diante das várias abordagens propostas ao longo do tempo observa-se que existe certo consenso entre os pesquisadores da área com relação aos principais mecanismos políticos e institucionais que impulsionam a adoção de novos modelos de políticas (COÊLHO; CAVALCANTE; TURGEON, 2016). Os mecanismos são: competição política, aprendizado, emulação e coerção. O mecanismo de competição política considera que as decisões tomadas visam a sobrevivência política no tempo, com o objetivo de atingir retornos econômicos e reeleição. O mecanismo de aprendizado se refere a uma decisão no sentido de adotar programas ou práticas políticas que potencialmente irão gerar melhores resultados do que outras alternativas (Meseguer, 2005, 2006). O mecanismo de emulação se caracteriza como um processo de imitação de outros modelos, em termos simbólicos ou normativos. O mecanismo de coerção é definido pelo uso de pressões políticas de uma instituição sobre outras no sentido da aplicação de sanções em caso de não adoção de uma política. Alguns autores não reconhecem coerção como um mecanismo de difusão.

No caso brasileiro, a literatura sobre o processo de difusão de políticas públicas parece escassa. Algumas das contribuições de destaque são de Sugiyama (2008), Wampler (2008), Spada (2010), Coêlho (2012), Oliveira (2013), Coêlho, Cavalcanti e Turgeon (2016), Faria (2003) e a publicação dos livros *"Diffusion of Good Government. Social*



*Sector Reforms in Brazil*”, de Sugiyama(2013) e o livro ”Difusão de Políticas Públicas a partir do Brasil”, organizado por Faria, Coêlho e Silva (2016).

Por ser uma federação altamente descentralizada, onde as três esferas de governo (União, Estados e Municípios) possuem autonomia para implementar políticas, o caso brasileiro representa um ambiente potencial para pesquisas na área. O caso brasileiro apresenta simultaneamente vários mecanismos direcionais, a difusão de políticas podem ocorrer horizontalmente, entre municípios e entre estados, verticalmente de cima para baixo (*top down*), de estados para municípios ou do governo federal para os governos subnacionais, e ainda, verticalmente de baixo para cima (*bottom-up*) que tem direção contrária ao *top down* (COÊLHO; CAVALCANTE; TURGEON, 2016).

Nas últimas décadas, os estudos de difusão de políticas têm recebido grande atenção, especialmente nos EUA e na Europa (COÊLHO; CAVALCANTE; TURGEON, 2016). Nesta conjuntura, é possível constatar que apesar de existir um grande número de publicações e pesquisas na esfera internacional, há um enorme espaço de conhecimento a preencher. Desta forma, esse campo se apresenta como um área potencial de pesquisa.

O presente trabalho está vinculado ao projeto “Estudos e Pesquisas para o aprimoramento das Capacidades Institucionais de Órgãos da Administração Pública para a elaboração, implementação e avaliação de políticas públicas e seus mecanismos de governança”, em desenvolvimento pelo Laboratório de Comportamento Político, Instituições e Políticas Públicas (LAPCIPP) do Instituto de Ciência Política (IPOL) da Universidade de Brasília (UnB) em parceria com a Fundação de Apoio à Pesquisa (FUNAPE). O projeto cooperativo visa contribuir com a ENAP no desenvolvimento de estudos e no aprimoramento de soluções ofertadas que possibilitem a produção de conhecimento estratégico na gestão de políticas públicas. Conseqüentemente, o projeto busca contribuir para a transparência da administração, a publicidade das ações e a otimização do uso e racionalização dos custos. O projeto ao qual o presente trabalho se vincula é focado no desenvolvimento de pesquisa sobre o aprimoramento das capacidades institucionais da burocracia pública federal com foco na avaliação de políticas públicas, mecanismos de governança e mapeamento da difusão do portfólio ministerial aos governos subnacionais.

Para o desenvolvimento da análise inicialmente são apresentados os objetivos a serem alcançados ao final do estudo. Em seguida é exposto o arcabouço teórico das técnicas estatísticas de análise de sobrevivência e de ciência política utilizadas. A seção seguinte descreve o material e a metodologia empregada. Posteriormente são feitas análises descritivas das adesões às políticas em estudo, bem como o ajuste modelo de sobrevivência a cada uma das políticas e o cálculo da taxa de inovação proposta. Por fim, tem-se os resultados do estudo, como também a comparação dos resultados com os scores tradicionais.



## **3 Objetivos**

### **3.1 Objetivo Geral**

Propor uma nova taxa de inovação relacionada à adoção de políticas públicas que considere a influência de variáveis explicativas.

### **3.2 Objetivos Específicos**

- Estudar a difusão de 3 políticas públicas dos Estados Unidos da área de educação.
- Identificar variáveis explicativas que influenciem a adoção das políticas públicas educacionais dos EUA.
- Calcular a taxa de inovação proposta para as políticas em estudo.
- Comparar os resultados da taxa de inovação proposta com as taxas de Walker e com a taxa de adoção proposta por Boehmke e Skinner (2012).



## 4 Revisão de Literatura

### 4.1 Análise de Sobrevivência

#### 4.1.1 Definições gerais

Para melhor compreensão do processo de difusão de uma política, buscando identificar os motivos da adoção de políticas públicas por um grupo de países, estados ou municípios e da não adoção por outros, os fatores que influenciam na velocidade da difusão de políticas em diversas áreas, os obstáculos que impedem a adoção da inovação e as facilidades, entre outras questões, técnicas de análise de sobrevivência podem ser utilizadas. A disseminação desse método nesta área decorre do estudo de Berry e Berry (1990), que aplicaram a técnica de análise de sobrevivência em análise da difusão de loterias nos Estados Unidos (GRAHAM; SHIPAN; VOLDEN, 2008).

Técnicas de análise de sobrevivência surgiram a partir de pesquisas da área médica, cujo interesse era o estudo da mortalidade, levando em consideração o tempo transcorrido entre o diagnóstico de uma determinada doença e a morte do paciente. Dessa forma, a análise de sobrevivência visa monitorar a exposição, ou seja, o tempo, até a ocorrência de um dado evento de interesse, ou até o término do estudo, chamado de tempo de falha. Para examinar o tempo de falha o tempo de início do estudo deve ser precisamente definido e os indivíduos devem ser comparáveis no início. Com relação a escala de medida do tempo, a escala comumente utilizada é o tempo real ou "de relógio", apesar de existirem outras alternativas como número de ciclos ou quilometragem. Ainda sobre o tempo de falha, os eventos de interesse que determinam a falha, na maioria dos estudos são eventos indesejados, porém podem haver casos em que o evento é desejado, como no caso da difusão de uma política. Este evento de interesse deve ser claramente definido no início do estudo e pode ocorrer devido a uma causa ou mais. Importante destacar que, no geral, indivíduos são retirados do estudo após a ocorrência de falha.

Apesar de ter origem na área médica, essa técnica estatística é de ampla utilização em diversas áreas do conhecimento. Em engenharia, são comuns estudos em que produtos são testados afim de se verificar o tempo de vida, área denominada de confiabilidade. Na área educacional, a técnica pode ser utilizada, por exemplo, para avaliar o tempo de permanência de alunos até a evasão escolar, como nos estudos Echeveste (1997) e Lisboa (2021). Nas ciências sociais, o objetivo pode ser avaliar o tempo até a reincidência criminal (SCHMIDT; WITTE, 1989), ou até mesmo em estudos de difusão de políticas, em que é investigado o tempo até que uma política pública seja adotada. Em análise de sobrevivência a variável resposta é, geralmente, o tempo até a ocorrência de um evento

de interesse (COLOSIMO; GIOLO, 2006).

Na conjuntura do estudo da difusão de políticas públicas o evento de interesse, ou seja, a ocorrência da falha, segundo a teoria de análise de sobrevivência, é a adoção de determinada política por um país, estado ou município. A falha de um estado neste cenário representa maior oferta de educação ou saúde, por exemplo, à população do estado. Portanto, quanto maior a quantidade de falhas, mais qualidade de vida está sendo ofertada aos indivíduos. É fundamental que os tomadores de decisão trabalhem para que a educação e a saúde estejam à disposição de todos os cidadãos, entre outras questões.

A principal característica de dados de sobrevivência é a presença de censura, que é dada pela observação parcial da resposta, que pode ocorrer por diversos motivos. Ou seja, quando a experimentação do evento de interesse não é observada, sabe-se apenas que o tempo de falha é superior aquele observado. Indivíduos censurados devem ser incluídos na análise estatística, pois mesmo com informações incompletas, estas fornecem certa informação sobre o tempo de vida desses indivíduos e a omissão dessa informação pode acarretar em vieses nas conclusões. Sem a presença de censura, técnicas estatísticas clássicas poderiam ser utilizadas, no entanto, se houver censuras, faz-se necessário o uso de métodos de análise de sobrevivência que possibilitam incorporar nas análises as informações contidas nos dados censurados.

Dessa forma, dados de sobrevivência são caracterizados pelos tempos de falhas e pelas censuras, estes componentes constituem a variável resposta. O tempo de falha é definido pelos seguintes elementos: tempo inicial do estudo, escala de medida e pelo evento de interesse, falha.

Desse modo, a variável resposta é dada pelo par  $(t, \delta)$ , sendo  $t$ , o tempo registrado e  $\delta$ , variável que indica se a observação é de censura ou de falha. A variável indicadora é dada por:

$$\delta_i = \begin{cases} 1 & \text{se } t_i \text{ é tempo de falha} \\ 0 & \text{se } t_i \text{ é tempo de censura.} \end{cases} \quad (4.1.1)$$

A censura pode ainda ser classificada em à direita, a esquerda e intervalar. Ocorre censura a direita quando o evento de interesse está a direita do tempo registrado, ocorre censura a esquerda quando o tempo registrado é maior que o tempo de falha, isto é, o evento de interesse já aconteceu quando o indivíduo foi observado. Vale ressaltar que pode ocorrer simultaneamente censura a direita e a esquerda. Já a censura intervalar ocorre quando não é registrado o tempo exato de experimentação da falha, mas se tem conhecimento do intervalo de tempo que ocorreu a falha. Neste trabalho, o foco estará na censura à direita, pois é a situação que ocorre com maior frequência em estudos, tanto da área médica como na área das ciências sociais.

A censura à direita se subdivide em três tipos, a censura à direita do tipo I, censura à direita do tipo II e a censura à direita aleatória. A censura à direita do tipo I é aquela em que o estudo terá término após um período pré estabelecido de tempo, assim indivíduos que não falharem até esse tempo serão censurados. A censura à direita do tipo II é aquela em que o estudo terá termino após o evento de interesse ter sido experimentado por um número pré estabelecido de indivíduos, os demais serão censurados. Já a censura à direita aleatória ocorre quando um indivíduo é retirado do estudo por algum motivo aleatório, estes são censurados, este último mecanismo é o que mais ocorre na prática. As mesmas técnicas estatísticas são utilizadas na análise de dados oriundos dos três mecanismos de censura à direita.

#### 4.1.2 Funções para o tempo de falha

Em estudos de sobrevivência define-se a variável aleatória  $T$  não negativa, usualmente contínua, como o tempo de falha. A distribuição dessa variável pode ser caracterizada pela função de densidade de probabilidade, pela função de sobrevivência e pela função de risco.

A função de densidade de probabilidade é definida como a probabilidade de um indivíduo falhar em um determinado intervalo de tempo, dado por  $\Delta t$ , com  $\Delta t$  tendendo a 0, e pode ser expressa por:

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t)}{\Delta t}, \quad (4.1.2)$$

com  $f(t) \geq 0 \forall t$  e  $\int_0^{\infty} f(t)dt = 1$  (COLOSIMO; GIOLO, 2006).

A função de sobrevivência, uma das principais utilizadas para descrever estudos de sobrevivência, é definida como a probabilidade de um indivíduo não falhar até um certo tempo  $t$ , ou seja, sobreviver ao tempo  $t$ . Pode ser escrita como:

$$S(t) = P(T \geq t), \quad (4.1.3)$$

com  $t \geq 0$ . É uma função monótona decrescente, com  $\lim_{t \rightarrow 0} S(t) = 1$  e  $\lim_{t \rightarrow \infty} S(t) = 0$  (COX; OAKES, 2018).

Em decorrência, a função de distribuição acumulada, dada por  $F(t) = 1 - S(t)$ , é definida como a probabilidade de uma observação não sobreviver ao tempo  $t$ .

A função de risco  $h(t)$  indica a forma em que as falhas ocorrem no decorrer do

tempo, por esse motivo também é chamada de função taxa de falha, e pode ser definida, segundo Lawless(2011), como:

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}, \quad (4.1.4)$$

representando a probabilidade de um indivíduo falhar no intervalo  $[t, t + \Delta t)$ , dado que a falha não ocorreu até o tempo  $t$ .

A função de risco acumulada, como o próprio nome sugere, representa a taxa de falha acumulada, não possui uma interpretação direta, mas pode ser útil na avaliação da função de taxa de risco, principalmente na estimação não paramétrica, em que apresenta um estimador com propriedades ótimas, ao passo que a função taxa de falha é difícil de ser estimada. Esta função pode ser definida por:

$$H(t) = \int_0^t h(u) du. \quad (4.1.5)$$

Considerando as funções definidas anteriormente, existem algumas relações matemáticas importantes entre elas, (COLOSIMO; GIOLO, 2006):

$$h(t) = \frac{f(t)}{S(t)} = -\frac{d}{dt}(\log S(t)), \quad (4.1.6)$$

$$H(t) = \int_0^t h(u) du = -\log S(t), \quad (4.1.7)$$

e

$$S(t) = \exp\{-H(t)\} = \exp\left\{-\int_0^t h(u) du\right\}. \quad (4.1.8)$$

Essas relações mostram que o conhecimento de uma das funções implica no conhecimento das demais.

### 4.1.3 Estimador de Kaplan-Meier

Devido a presença de censuras, o principal componente da análise envolvendo dados de sobrevivência é a função de sobrevivência. Sendo assim, é necessário encontrar uma estimativa para tal função, para a partir dela estimar estatísticas de interesse, que



usualmente são o tempo médio ou mediano, percentis ou certas frações de falha em tempos fixos. Na literatura existem alguns estimadores não paramétricos para essa função, ou seja, que independem da distribuição de probabilidade dos dados, dentre eles o estimador de Kaplan-Meier.

O estimador de Kaplan-Meier é o mais utilizado para estimar a função de sobrevivência em estudos, que muitas vezes é difícil de ser calculada devido a presença das censuras. Este estimador foi proposto por Kaplan e Meier em 1958 e também é chamado de estimador limite-produto. Consiste em uma adaptação da forma empírica de  $S(t)$ , formulada a partir do princípio de ausência de censura, definida como:

$$\hat{S}(t) = \frac{\text{n}^\circ \text{ de observações que não falharam até o tempo } t}{\text{n}^\circ \text{ total de observações no estudo}}. \quad (4.1.9)$$

Seguindo a essência da equação 4.1.9, a expressão do estimador de Kaplan-Meier pode então ser apresentada como:

$$\hat{S}(t) = \prod_{j:t_j < t} \left( \frac{n_j - d_j}{n_j} \right) = \prod_{j:t_j < t} \left( 1 - \frac{d_j}{n_j} \right), \quad (4.1.10)$$

em que:

- $t_1 < t_2 \dots < t_k$ , são os  $k$  tempos distintos e ordenados de falha,
- $d_j$  é o número de falhas em  $t_j$ ,  $j = 1, \dots, k$ , e
- $n_j$  o número de indivíduos sob risco em  $t_j$ , ou seja, indivíduos que não falharam ou foram censurados até o instante anterior a  $t_j$

Com relação as suas propriedades, sabe-se que é um estimador não viciado para grandes amostras, é fracamente consistente, converge assintoticamente para uma distribuição normal e é um estimador de máxima verossimilhança de  $S(t)$  (COLOSIMO; GILOLO, 2006).

#### 4.1.4 Modelo de Regressão de Cox

As técnicas não paramétricas não permitem a inclusão de covariáveis na análise, apesar de serem importantes para descrever dados de sobrevivência pela sua facilidade de aplicação. Em análise e sobrevivência existem ainda duas classes de modelos propostos na

literatura, os paramétricos e os semi-paramétricos. Os paramétricos, também chamados de tempo de vida acelerado, são mais eficientes, porém menos flexíveis. Os modelos semi-paramétricos, também conhecidos como modelos de regressão de Cox, são mais flexíveis e permitem a incorporação de covariáveis dependentes no tempo.

O modelo de regressão de Cox (Cox, 1972), proposto em 1972, possibilita que as covariáveis de interesse sejam consideradas na análise dos tempos de falha, e é formulado pela seguinte função:

$$h(t|\mathbf{x}) = h_0(t) \exp(\mathbf{x}^T \boldsymbol{\beta}), \quad (4.1.11)$$

em que  $h_0(t)$  é a função de risco base, ou seja, a função de risco de um indivíduo com vetor de covariáveis nulo,  $\boldsymbol{\beta}$  é o vetor de dimensão  $p$  dos coeficientes de regressão desconhecidos e  $\mathbf{x}$  é o vetor de dimensão  $p$  de covariáveis observadas para o indivíduo  $i$ ,  $i = 1, \dots, n$ .

O modelo de Cox é chamado semi-paramétrico, pois é composto por  $h_0$ , que não assume nenhuma forma paramétrica, e pela parte paramétrica que são as covariáveis que atuam de forma multiplicativa por meio de uma  $g(x, \beta) = \exp(\beta x_i)$ . A razão das funções de risco para dois indivíduos,  $i$  e  $j$ , é dada por:

$$\frac{h(t|\mathbf{x}_i)}{h(t|\mathbf{x}_j)} = \frac{h_0(t) \exp(\beta' \mathbf{x}_i)}{h_0(t) \exp(\beta' \mathbf{x}_j)} = \frac{\exp(\beta' \mathbf{x}_i)}{\exp(\beta' \mathbf{x}_j)}, \quad (4.1.12)$$

não dependendo do tempo  $t$ . Portanto, supõe-se que as taxas de falhas sejam proporcionais. A função de sobrevivência para o modelo de Cox, considerando um indivíduo com vetor de covariável  $\mathbf{x}$  é dada por:

$$S(t|\mathbf{x}) = (S_0(t))^{\exp(\beta' \mathbf{x})}, \quad (4.1.13)$$

no qual,  $S_0(t) = \exp[-\int_0^t h_0(u) du]$  é a função de sobrevivência e  $\boldsymbol{\beta}$  é o vetor de covariáveis.

#### 4.1.5 Estimação

Além do tempo de sobrevivência e da variável indicadora de censura, é possível avaliar outras variáveis características da população. Essas variáveis são conhecidas como

variáveis explicativas. Eventualmente estudos de sobrevivência buscam entender a relação entre o tempo de falha e essas variáveis explicativas. Para tanto, pode ser ajustado um modelo de regressão que represente essa relação.

Existem alguns métodos de estimação conhecidos na literatura, sendo o método de máxima verossimilhança uma opção apropriada para dados de sobrevivência, pois este método incorpora as censuras, é relativamente fácil de ser entendido e possui propriedades ótimas para grandes amostras.

O método de máxima verossimilhança se baseia nos resultados obtidos pela amostra, buscando identificar qual é a distribuição entre todas possíveis, considerando os valores dos parâmetros, com maior possibilidade e ter gerado a amostra obtida.

Dessa forma, o objetivo é encontrar o vetor de parâmetros  $\boldsymbol{\theta}$  que maximize a função de verossimilhança  $L(\boldsymbol{\theta})$ , dada por:

$$L(\boldsymbol{\theta}) = \prod_{i=1}^n f(t_i; \boldsymbol{\theta}), \quad (4.1.14)$$

nos quais  $t_1, \dots, t_n$  representam os tempos observados de uma amostra de tamanho  $n$  de uma população.

A função de verossimilhança  $L(\boldsymbol{\theta})$  mostra que a contribuição de cada observação não censurada é sua função de densidade. E, observação censurada à direita contribui informando que o tempo de falha é maior que o tempo observado, ou seja, contribuem com sua função de sobrevivência. Dessa forma, tem-se a seguinte expressão:

$$L(\boldsymbol{\theta}) \propto \prod_{i=1}^n [f(t_i; \boldsymbol{\theta})]^{\delta_i} [S(t_i; \boldsymbol{\theta})]^{1-\delta_i}, \quad (4.1.15)$$

onde  $\boldsymbol{\theta}$  é o vetor de parâmetros do modelo a ser estimado.

Contudo, é conveniente utilizar o logaritmo da função de verossimilhança. Os estimadores de máxima verossimilhança são os valores de  $\boldsymbol{\theta}$  que maximizam  $L(\boldsymbol{\theta})$  ou equivalentemente  $\log(L(\boldsymbol{\theta}))$ . Podem ser encontrados resolvendo o sistema de equações: (COLOSIMO; GIOLO, 2006)

$$U(\boldsymbol{\theta}) = \frac{\sigma \log L(\boldsymbol{\theta})}{\sigma \boldsymbol{\theta}} = 0 \quad (4.1.16)$$

Portanto, o método de máxima verossimilhança é utilizado para obter estimativas pontuais para os parâmetros do modelo. Este método permite ainda a construção de intervalos de confiança para os parâmetros, a partir das propriedades para grandes amostras.

A escolha do modelo a ser utilizado é uma etapa extremamente importante da análise para que os resultados sejam fidedignos. O método de máxima verossimilhança só pode ser aplicado após a definição do modelo probabilístico para os dados em estudo. Essa definição pode ser feita através de métodos gráficos ou por meio de testes de hipóteses.

Um dos métodos gráficos utilizado consiste na comparação da função de sobrevivência do modelo proposto com o estimador de Kaplan-Meier. O intuito é selecionar o modelo cuja curva se aproxima melhor da curva do estimador de Kaplan-Meier. Outro método gráfico utilizado consiste na linearização da função de sobrevivência. A ideia é a construção de gráficos que sejam aproximadamente lineares caso o modelo proposto seja apropriado.

Outra forma de discriminar modelos é através de testes de hipóteses, com a hipótese nula de que o modelo de interesse é adequado versus a hipótese alternativa de que o modelo não é adequado. Este teste é realizado utilizando a estatística da razão de verossimilhanças em modelos encaixados (Cox e Hinkley, 1977). Essa estatística é calculada a partir dos seguintes valores,  $(\log L(\hat{\theta}_G))$  e  $(\log L(\hat{\theta}_M))$ , que são respectivamente, o modelo generalizado e obtenção do valor do logaritmo de sua função de verossimilhança e o modelo de interesse e obtenção do valor do logaritmo de sua função de verossimilhança, desse modo, a estatística é definida como:

$$TRV = -2 \left[ \frac{(\log L(\hat{\theta}_M))}{(\log L(\hat{\theta}_G))} \right] = 2[(\log L(\hat{\theta}_G)) - (\log L(\hat{\theta}_M))], \quad (4.1.17)$$

e segue aproximadamente uma distribuição qui-quadrado com graus de liberdade igual a diferença do número de parâmetros  $(\hat{\theta}_G)$  e  $(\hat{\theta}_M)$  dos modelos sendo comparados.

A presença do componente não-paramétrico no modelo de Cox torna inviável o uso do método de máxima verossimilhança para estimar o vetor de coeficientes  $\beta$ . Assim, Cox propôs em um artigo em 1975, o método de máxima verossimilhança parcial.

Para definir a função de verossimilhança parcial, deve-se considerar uma amostra de  $n$  indivíduos com  $k \leq n$  falhas distintas nos tempos  $t_1 \leq \dots \leq t_k$ . Supondo ainda que a probabilidade condicional da  $i$ -ésima observação vir a falhar no tempo  $t_i$ , conhecendo quais observações estão sob risco em  $t_i$  é (COLOSIMO; GIOLO, 2006):

$$\frac{h_i(t)}{\sum_{j \in R(t_i)} h_j(t)} = \frac{h_0(t) \exp\{\mathbf{x}'_i \boldsymbol{\beta}\}}{\sum_{j \in R(t_i)} h_0(t) \exp\{\mathbf{x}'_j \boldsymbol{\beta}\}} = \frac{\exp\{\mathbf{x}'_i \boldsymbol{\beta}\}}{\sum_{j \in R(t_i)} \exp\{\mathbf{x}'_j \boldsymbol{\beta}\}}, \quad (4.1.18)$$

em que  $R(t_i)$  é o conjunto dos índices das observações sob risco no tempo  $t_i$ .

A função de verossimilhança a ser utilizada para fazer inferências no modelo é formada pelos produtos de todos os termos representados em 4.1.18 associados aos tempos distintos de falha, ou seja (COLOSIMO; GIOLO, 2006),

$$L(\boldsymbol{\beta}) = \prod_{i=1}^k \frac{\exp\{\mathbf{x}'_i \boldsymbol{\beta}\}}{\sum_{j \in R(t_i)} \exp\{\mathbf{x}'_j \boldsymbol{\beta}\}} = \prod_{i=1}^n \left( \frac{\exp\{\mathbf{x}'_i \boldsymbol{\beta}\}}{\sum_{j \in R(t_i)} \exp\{\mathbf{x}'_j \boldsymbol{\beta}\}} \right)^{\delta_i}, \quad (4.1.19)$$

com  $\delta_i$  o indicador de falha. Os valores de  $\boldsymbol{\beta}$  que maximizam  $L(\boldsymbol{\beta})$ , a função de verossimilhança parcial são obtidos resolvendo os sistemas de equações,  $U(\boldsymbol{\beta}) = 0$ , em que  $U(\boldsymbol{\beta})$  é o vetor escore de primeiras derivadas da função  $l(\boldsymbol{\beta}) = \log(L(\boldsymbol{\beta}))$ . Portanto, (COLOSIMO; GIOLO, 2006)

$$U(\boldsymbol{\beta}) = \sum_{i=1}^n \delta_i \left[ \mathbf{x}_i - \frac{\sum_{j \in R(t_i)} \mathbf{x}_j \exp\{\mathbf{x}'_j \hat{\boldsymbol{\beta}}\}}{\sum_{j \in R(t_i)} \exp\{\mathbf{x}'_j \hat{\boldsymbol{\beta}}\}} \right] = 0. \quad (4.1.20)$$

A função de verossimilhança parcial 4.1.20 assume que os tempos de sobrevivência são contínuos e não pressupõe possibilidade de empates nos valores observados.

A função 4.1.20 deve ser modificada para incorporar as observações empatadas, quando estão presentes. Considerando o vetor formado pela soma das correspondentes  $p$  covariáveis para os indivíduos que falham no mesmo tempo  $t_i$ , com  $i = 1, \dots, k$  e  $d_i$  sendo o número de falhas neste mesmo tempo. A aproximação para 4.1.20 que incorpora os empates considera a seguinte função de verossimilhança parcial:

$$L(\boldsymbol{\beta}) = \prod_1^n \frac{\exp(\mathbf{s}'_i \boldsymbol{\beta})}{\left[ \sum_{j \in R_i} \exp(\boldsymbol{\beta}' X_j) \right]^{\delta_i}}, \quad (4.1.21)$$

em que  $R_i$  é o conjunto dos indivíduos ainda em risco no tempo  $t$ . Essa função é adequada quando o número de observações empatadas em qualquer tempo não é grande.

#### 4.1.6 Avaliação do Ajuste do Modelo de Cox

##### Método gráfico

A partir da expressão do modelo de Cox, pode-se observar que o efeito das covariáveis é de acelerar ou desacelerar a função de risco. Tomando a razão das taxas de falhas de dois indivíduos ( $iej$ ) que possuem os mesmos valores para as covariáveis com exceção da  $l$ -ésima, tem-se:

$$\frac{h_i(t)}{h_j(t)} = \exp \{ \beta_l (\mathbf{x}_{il} - \mathbf{x}_{jl}) \}, \quad (4.1.22)$$

que pode ser interpretado como a razão de riscos no tempo  $t$ , constante em todo o acompanhamento.

Entretanto, funções relacionadas com  $h_0(t)$  são importantes no modelo de Cox, como a função de taxa de falha acumulada,  $H_0(t) = \int_0^t h_0(u) du$ , a função de sobrevivência  $S_0(t) = \exp(-H_0(t))$  e a função de sobrevivência  $S_0(t) = \exp(-H_0(t))$ .

Os estimadores para estas funções são de natureza não-paramétrica. Uma estimativa para  $H_0(t)$ , proposta por Breslow(1972), é uma função escada com saltos nos tempos de falha e expressa por:

$$\hat{H}_0(t_i) = \sum_{j:t_j \leq t} \frac{d_j}{\sum_{l \in R_j} \exp \{ \mathbf{x}'_l \hat{\beta} \}}, \quad (4.1.23)$$

com  $d_j$  sendo o número de falhas no tempo  $j$ . Logo  $S_0(t)$  e  $S(t)$  podem ser estimadas por:

$$\hat{S}_0(t) = \exp \{ -\hat{H}_0(t) \} \quad (4.1.24)$$

e

$$\hat{S}(t) = \left[ \hat{S}_0(t) \right]^{\exp \{ \mathbf{x} \hat{\beta} \}}, \quad (4.1.25)$$

ambas funções escada decrescente ao longo do tempo.

O modelo de regressão de Cox é bastante flexível devido à presença do componente não-paramétrico e possui como suposição que os riscos são proporcionais, dessa forma, são necessários métodos para avaliar a adequação desse modelo.

Para verificar a suposição de riscos proporcionais no modelo de Cox, métodos gráficos podem ser utilizados. Um deles consiste em dividir os dado em  $m$  estratos, de acordo com alguma variável e em seguida deve-se estimar  $\hat{H}_0(t)$  para cada estrato da covariável em questão. A técnica consiste então em comparar as curvas do logaritmo de  $\hat{H}_0(t)$  versus  $t$ , ou  $\log(t)$ , se a suposição for pertinente as curvas apresentarão diferenças aproximadamente constantes ao longo do tempo.

### Resíduos

Além da suposição dos riscos proporcionais, há interesse em examinar outros aspectos do modelo de Cox, dentre eles, verificar o ajuste global do modelo ajustado, verificar a melhor forma para explicar a influência de uma dada variável na sobrevivência, na presença das demais variáveis explicativas, verificar a presença de outliers e analisar a influência de cada indivíduo no modelo ajustado. A fim de examinar estes aspectos, técnicas de diagnóstico também podem ser utilizadas para o modelo de Cox, baseadas nos mesmos resíduos para os modelos paramétricos, tais como os resíduos de Cox-Snell, Martingale e deviance.

Os resíduos de Cox-Snell (1968), são úteis para examinar o ajuste global do modelo. Para o modelo de Cox são definidos por:

$$\hat{e}_i = \hat{H}_0(t_i) \exp \left\{ \sum_{k=1}^p \mathbf{x}_{ip} \hat{\beta}_k \right\}, i = 1, \dots, n. \quad (4.1.26)$$

Se o modelo for adequado os resíduos devem seguir uma distribuição exponencial padrão (Lawless, 2003). Para que o modelo exponencial seja adequado, o gráfico de  $\hat{e}_i$  versus  $\hat{H}(\hat{e}_i)$  deve ser aproximadamente uma reta.

Os resíduos de martingale resultam de uma modificação dos resíduos de Cox-Snell. Assim, quando os dados apresentarem censuras à direita e todas as covariáveis não forem dependentes do tempo, os resíduos de martingale para o modelo de Cox são definidos por:(COLOSIMO; GIOLO, 2006)

$$\hat{m}_i = z = \delta_i - \hat{H}_0(t_i) \exp \left\{ \sum_{k=1}^p \mathbf{x}_{ip} \hat{\beta}_k \right\} = \delta_i - \hat{e}_i, i = 1, \dots, n. \quad (4.1.27)$$

Os resíduos de martingale são utilizados para verificar a forma funcional das

covariáveis.

Os resíduos deviance, detectam a presença de outliers, devem apresentar comportamento aleatório em torno de zero caso o modelo seja apropriado. Os resíduos deviance no modelo de Cox são definidos por:

$$\hat{d}_i = \text{sinal}(\hat{m}_i) [-2(\hat{m}_i + \delta_i \log(\delta_i - \hat{m}_i))]^{1/2}. \quad (4.1.28)$$

## 4.2 Medida de inovações relacionadas a adoção de políticas públicas

Estudiosos da ciências políticas ao longo do tempo estudaram o questionamento seminal de Walker (1969), de que alguns estado americanos seriam mais inovadores do que outros, de várias maneiras. No entanto, com o tempo a literatura passou da questão original da inovação como traço geral para o estudo da inovação em uma política específica. Essa mudança gerou debates acerca do índice de inovação política, preocupações com relação à comparabilidade entre políticas e ao longo do tempo foram explanadas.

Essas questões começaram a ser investigadas pela introdução da análise de sobrevivência (BERRY; BERRY, 1990), que abordou essas preocupações dentro de uma única política.

Com base no trabalho de Rogers (1962) sobre a difusão da tecnologia agrícola entre os agricultores, Walker definiu a inovação política como "um programa ou política que é novo para os Estados que a adotam, não importa a idade do programa ou quantos outros estados podem tê-lo adotado" (WALKER, 1969). Com base nessa definição, ele procurou estudar por que alguns estados eram mais inovadores do que outros, no sentido de que um Estado inovador adotaria, pelo menos em média, novas políticas mais cedo do que estados menos inovadores. Para responder a essa pergunta, desenvolveu um índice de inovação política utilizando diversas políticas diferentes, que foi então analisado para determinar os padrões espaciais, políticos e demográficos de inovação entre os estados.

Em resposta à proposição de Walker, Gray (1973) levantou uma série de preocupações sobre a maneira de estudar a inovação como uma característica inerente dos Estados americanos.

Apesar das tentativas de revigorar a literatura, a questão da inovação política como um traço dos Estados americanos foi largamente deixada de lado. Os estudiosos deixaram de estudar a inovação política como um traço geral para estudar os padrões de difusão em políticas únicas ou, em particular, explicar padrões cumulativos de adoção.

A introdução da análise de sobrevivência nesse campo (BERRY; BERRY, 1990),



permitiu que os estudiosos avaliassem simultaneamente os determinantes internos e externos da inovação política.

O score de inovação de Walker formou então a base para muitos estudos de inovação em todos os estados americanos (Boushey, 2010; Canon e Baum, 1981; Savage, 1978). O score de Walker avalia o ano de adoção do estado  $i$ , comparando o primeiro e último estado a aderir à política  $k$ . Esse score é definido por:

$$W_{ik} = \frac{Y_k^{MAX} - Y_{ik}}{Y_k^{MAX} - Y_k^{MIN}}, \quad (4.2.1)$$

em que  $Y_k^{MAX}$  é o ano de adesão do último estado adotante da política,  $Y_k^{MIN}$  é o ano de adesão do primeiro estado que adotou a política e  $Y_{ik}$  o ano de adesão da política  $k$  para o estado  $i$ . Essa medida avalia a velocidade de adesão para cada uma das  $k$  políticas de cada estado  $i$ , variando entre 0 e 1. Sendo assim, o primeiro estado a adotar uma política recebe uma pontuação de um, enquanto o último estado recebe uma pontuação de zero.

Para gerar uma medida geral de inovação para cada estado, Walker então faz uma média das pontuações do estado de cada política como:

$$\hat{W}_i = \frac{1}{K} \sum_{k=1}^k W_{ik}. \quad (4.2.2)$$

A média acima possibilita a comparação da inovação entre os estados. No entanto, as pontuações de Walker sofreram algumas críticas, segundo Boehmke e Skinne as medidas atribuí uma pontuação igual a zero para os estados não adotantes da política no tempo observado assim como o último estado a adotar e não leva em consideração as mudanças da inovação ao longo do tempo.

Com o intuito de solucionar as críticas colocadas, estudos subsequentes atribuíram aos não adotantes da política um tempo de adoção de um ano após o último estado observado. Dessa forma, é possível diferenciar o último observado dos censurados (Boushey 2010; Canon e Baum 1981; Savage 1978), essa abordagem dá pelo menos ao último adotante uma pontuação maior do que os não adotantes mas ainda sofre com a atribuição de uma data de adoção improvável para os não adotantes.

Para lidar com o segundo desafio, referente a crítica de Gray (1973) de que as pontuações semelhantes ignoram as mudanças na inovação ao longo do tempo, uma abordagem frequentemente adotada na literatura é atribuir pontuações ao período em que o

10<sup>o</sup> estado o adotou (BOEHMKE; SKINNER, 2012), embora seja um corte arbitrário, qualquer decisão desse tipo criará problemas.

Para resolver os vários problemas com medidas de inovação, Boehmke e Skinner (2012) definem a medida de inovação em um dado ponto no tempo,  $t$ , para o estado  $i$  como:

$$R_{it} = \frac{\sum_{i=1}^{K_{it}} Y_{ikt}}{K_{it}}, \quad (4.2.3)$$

em que  $Y_{ikt}$  assume valor igual a zero para estados em que não aderiram a política no ano  $t$ , valor igual a um para os que aderiram e é tratada como missing nos anos subquentes e  $K_{it}$  representa o número de políticas adotadas por pelo menos um estado - mas não pelo estado  $i$  - no ano  $t$  (ou seja, o número de políticas que o estado  $i$  poderia potencialmente adotar no ano  $t$ ).

Essa medida pode ser calculada em períodos de tempo arbitrários para produzir a taxa de inovação para o estado  $i$ , dada por:

$$\bar{R}_{ij} = \frac{\sum_{t=T_0}^T \sum_{i=1}^{K_{it}} Y_{ikt}}{\sum_{t=T_0}^T K_{it}}. \quad (4.2.4)$$

Segundo Boehmke e Skinner (2012), essa medida proposta por eles traz diversas vantagens, aborda a censura à direita, facilita a construção da medida para diversos intervalos de tempo. Por outro lado, mesmo abordando o problema da censura a direita essa medida ainda não contempla casos em que estados podem nunca aderir a uma política, pois a medida supõe que os estados censurados têm chance, ainda que pequena, de adotar a política. Além disso, traz um problema ao não considerar a presença de estados imunes a determinadas políticas (casos raros).

Os autores colocaram duas maneiras de lidar com essa segunda preocupação (BOEHMKE; SKINNER, 2012). Primeiro, os pesquisadores devem inspecionar estados imunes a certas políticas para desconsiderá-las como políticas potenciais. Em segundo, pode-se considerar a abordagem probabilística e assumir que alguma proporção de estados censurados à direita são imunes e, em vez de considerá-los como possíveis adotantes, eles contribuiriam com  $\pi_{ij}$ , a probabilidade estado de idade  $i$  corre o risco de adotar a política.

Dessa forma, a taxa de inovação proposta por Boehmke e Skinner (2012) aborda uma série de preocupações relacionadas ao score proposto por Walker, porém ainda deixa diversas lacunas a serem preenchidas, algumas delas seriam, a possibilidade de avaliar

simultaneamente uma lista de fatores que influenciam o tempo de adoção das inovações políticas entre os estados mais e menos inovadores e uma maneira de tratar as censuras a direita presente em dados de difusão de políticas de forma adequada.

## 5 Metodologia

### 5.1 Material

Com o objetivo de propor uma medida de inovação que preencha as lacunas da taxa de inovação proposta por Boehmke e Skinner (2012) serão utilizados dados de adoção dos estados americanos relacionados a três políticas da área de educação. Os dados que serão avaliados são parte da base de dados do artigo Boehmke; Skinner, 2012.

As três políticas educacionais americanas selecionadas para o estudo foram *Charter Schools*, *High School Exit Exams* e *School Choice*. A primeira diz respeito as *Charters Schools* que é uma escola que recebe financiamento do governo, mas opera independentemente do sistema escolar estadual estabelecido em que está localizada. É independente no sentido de que opera de acordo com o princípio básico da autonomia para prestação de contas, que está livre das regras, mas responsável pelos resultados. Críticos das *Charter Schools* afirmam que a operação privada das escolas charter com falta de responsabilidade pública as torna mais parecidas com instituições privadas subsidiadas pelo público. *High School Exit Exams* é um exame de saída do ensino médio como o objetivo de manter os padrões de graduação nas escolas públicas de ensino médio. Nos estados adotantes dessa política, os exames de saída são obrigatórios para todos os alunos, devendo portanto, ser aprovado nestes para obter diploma do ensino médio. School choice, nos Estados Unidos, é um termo usado para se referir às alternativas ao sistema de educação pública, como o Cheques-Ensino e o Ensino doméstico.

A partir do banco de dados foi possível extrair a informação do ano de adoção dos 50 estados americanos para cada uma das políticas. A partir dessa informação foram criadas a variável de tempo de sobrevivência, que indica os anos desde a primeira adoção da política até o ano de adoção do estado ou a censura deste, e a variável indicadora de falha ou censura, que informa se o tempo é de falha ( $\delta = 1$ ) ou censura ( $\delta = 0$ ).

Inicialmente o banco de dados contava com observações de 50 estados americanos, porém foram excluídos o Hawaii e Alaska, conforme recomendações da literatura, por serem unidades não contíguas e apresentarem características distintas dos demais estados americanos. Por fim, as análises subsequentes compreenderão informações de 48 estados, sendo eles: Alabama, Arizona, Arkansas, California, Colorado, Connecticut, Delaware, Florida, Georgia, Idaho, Illinois, Indiana, Iowa, Kansas, Kentucky, Louisiana, Maine, Maryland, Massachusetts, Michigan, Minnesota, Mississippi, Missouri, Montana, Nebraska, Nevada, New Hampshire, New Jersey, New Mexico, New York, North Carolina, North Dakota, Ohio, Oklahoma, Oregon, Pennsylvania, Rhode Island, South Carolina, South Dakota, Tennessee, Texas, Utah, Vermont, Virginia, Washington, West Virginia,

Wiscosin e Wyomig.

O conjunto de fatores que influenciam o tempo de adoção das inovações políticas entre os estados mais e menos inovadores serão selecionados dentre os candidatos, Número da palavra "Educação" nas constituições estaduais, Composição partidária das câmaras legislativas estaduais, Medida estadual da ideologia do cidadão, Medida estadual da ideologia do governo, Profissionalismo legislativo do estado e a População do estado.

As informações do Número da palavra "Educação" nas constituições estaduais e do Profissionalismo legislativo foram fornecidas pelo Professor Denilson Bandeira Coêlho, associado ao Instituto de Ciência Política da Universidade de Brasília, estudioso referência nesta temática.

A composição partidária das câmaras legislativas estaduais podem ser ambas as câmaras legislativas com maioria republicana, ambas as câmaras legislativas com maioria democrata ou podem não ter maioria de nenhum dos partidos políticos em ambas as câmaras legislativas. O nebraska é uma legislatura unicameral apartidária e por este motivo será retirado das análises envolvendo essa variável. Essa informação foi retirada do site da *National Conference of State Legislatures*, uma associação de funcionários públicos apartidários composta por legisladores estaduais em exercício dos estados, territórios e comunidades dos Estados Unidos. Sendo o ano de referência para as medidas o ano de adoção de cada uma das políticas para cada estado ou o ano imediatamente anterior a adoção com informação disponível.

As medidas estaduais da ideologia do cidadão e do governo (1960-2017) provém do *site* do Richard C. Fording, professor de Ciência Política na Universidade do Alabama. Seus principais interesses de ensino e pesquisa incluem políticas públicas, raça e política, política estatal, movimentos sociais e metodologia quantitativa. O ano considerado para as medidas foi o ano de adoção de cada uma das políticas para cada estado ou o ano imediatamente anterior a adoção com informação disponível.

O Departamento do Censo dos Estados Unidos, mais especificamente, a divisão de população, é a fonte para as estimativas intercensitárias da população residente. Foram consideradas as estimativas para o ano de 2000.

Explicações para a inovação da política geral requerem uma amostra suficientemente grande de políticas com uma seleção adequada de políticas de um amplo conjunto de categorias. O recorte do banco de dados original de Boehmke e Skinner fornece o ponto de partida, oferecendo uma amostra de políticas que simplifica e propicia a comparação com os scores já propostos na literatura.

## 5.2 Método

Esta pesquisa busca investigar quais fatores influenciam a decisão de adoção de políticas e procura mensurar o grau de inovação de estados baseado na adoção de políticas. Portanto, este trabalho propõe uma taxa que irá medir o grau de inovação dos estados segundo as adesões às políticas. Este é o mesmo objetivo das taxas propostas na literatura, o score de Walker e o *Adoption rate*, proposto por Boehmmke e Skinner (2012). No entanto estas medidas já existentes na literatura da ciência política não consideram diversos fatores como a censura e a influência de variáveis explicativas.

A variável dependente comumente avaliada em pesquisas de difusão é o tempo até a adoção de determinada política (COÊLHO; CAVALCANTE; TURGEON, 2016). Como nos modelos de análise de sobrevivência em que um dos dois componentes da resposta é o tempo entre o início do estudo, a ser fixado, e a ocorrência do evento de interesse, essa técnica pode ser utilizada em estudos com essa temática, neste caso o evento de interesse será a adoção da política. Esta variável costuma ser chamada como tempo de "falha", no contexto do estudo tem-se o tempo de adoção da política.

O segundo componente da resposta em dados de sobrevivência é a censura, que é a observação parcial da resposta, que foi interrompida por algum motivo, o que não permite a avaliação completa do tempo de falha. Essa variável censura, indica se a informação foi censurada ou não, ou seja, se o elemento não experimentou a falha ou falhou no período observado. Sem a presença de censura, as técnicas estatísticas clássicas, como análise de regressão e planejamento de experimentos, poderiam ser utilizadas na análise deste tipo de dados (COLOSIMO; GIOLO, 2006). Os casos censurados neste estudo são aqueles que não adotaram a política durante o período de análise.

Estudar o comportamento das variáveis explicativas também é essencial, para então propor um modelo adequado que consiga esclarecer os principais fatores que melhor explicam o padrão de adoção da política.

Em vista disso, neste estudo é proposto o uso do modelo de riscos proporcionais de Cox que incorpora a censura e avalia a interferência de cada uma das covariáveis na adoção estadual para cada uma das políticas. A partir disso, a medida de inovação proposta é dada pela soma dos riscos de falha para cada política em um mesmo tempo. Sendo assim, quanto maior o valor da medida proposta, maior será o risco do estado adotar as políticas, logo mais inovador o estado será.

Supondo que a função de risco para a política *Charter Schools* seja dada por  $h_1(t|\mathbf{x})$ ,  $h_2(t|\mathbf{x})$  para política *High School Exit Exams e School Choice* e  $h_3(t|\mathbf{x})$  para política *School Choice*, assumindo que o risco de falhar em cada política seja independente, o Escore de risco global proposto neste trabalho é definido como:

$$h_{eg}(t|\mathbf{x}) = h_1(t|\mathbf{x}) + h_2(t|\mathbf{x}) + h_3(t|\mathbf{x}),$$

em que  $h_j(t|\mathbf{x})$  é a função de risco do modelo de riscos proporcionais de Cox definida na equação 4.1.11 e  $j=1,2,3$  representa cada uma das políticas. Esse índice pode ser generalizado para mais políticas ao supor  $j = 1,2,\dots,k$ , sendo  $k$  o número de políticas independentes, logo o Escore de risco global é definido por:

$$h_{eg}(t|\mathbf{x}) = h_1(t|\mathbf{x}) + h_2(t|\mathbf{x}) + \dots + h_k(t|\mathbf{x}),$$

em que  $h_j(t|\mathbf{x})$  é a função de risco do modelo de riscos proporcionais de Cox.

Os resultados da medida proposta serão comparados com os cálculos para o score de Walker e *Adoption rate*. Em suma, a análise de sobrevivência é uma importante técnica para estudar o que explica a difusão de política, especialmente por considerar o fator tempo e a possibilidade de censura.

## 6 Resultados e Análises

### 6.1 Análise Descritiva

Primordialmente, deve-se realizar uma análise exploratória do conjunto de dados que seguidamente será utilizado na modelagem do tempo até a adoção de cada uma das políticas em avaliação neste estudo.

#### 6.1.1 *Charter Schools*

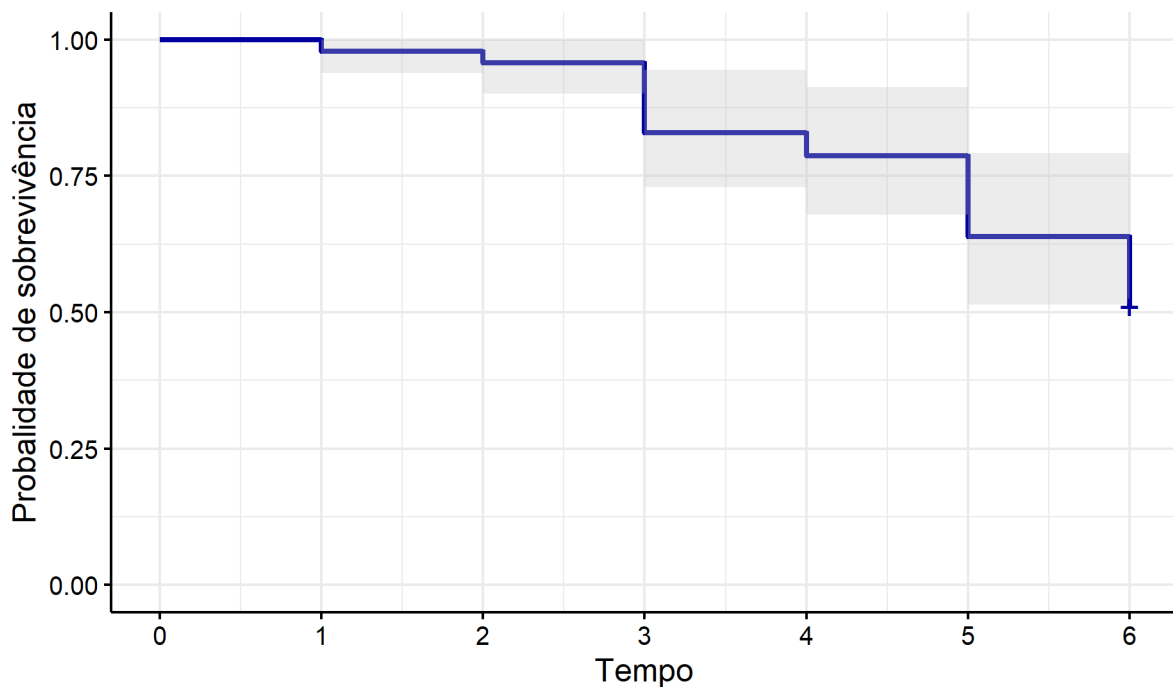


Figura 1: Gráfico da função de sobrevivência para a política *Charter Schools*.

Dos 48 estados americanos 23 falharam, ou seja, adotaram a política *Charter Schools* e 25 foram censurados, não adotaram a política no período observado.

No primeiro e segundo tempo a probabilidade de sobrevivência apresenta uma pequena queda, com apenas 1 falha no primeiro ano e 1 no segundo ano, a partir do tempo três esta probabilidade revela uma queda maior, contabilizando 6 falhas no tempo três.

A probabilidade de sobreviver ao longo do tempo se aproxima de 0,52, devido presença de censuras no último tempo do estudo.



## Composição partidária das câmaras legislativas estaduais

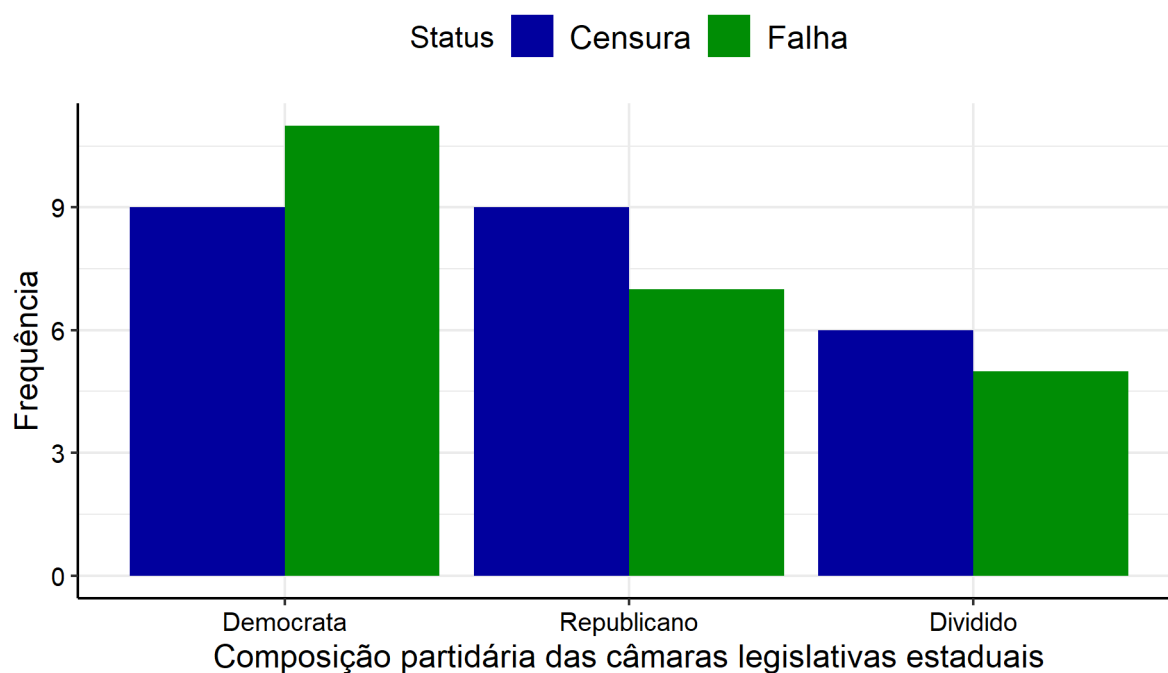


Figura 2: Gráfico de barras da Composição partidária das câmaras legislativas estaduais pelo Status da adoção da política *Charter Schools*.

Examinando o gráfico 2 da variável Composição partidária das câmaras legislativas estaduais pelo status da adoção da política *Charter Schools*, nota-se que a maioria dos estados em que ambas as câmaras legislativas têm maioria democrata falharam, portanto adotaram a política. Enquanto a maioria dos estados em que ambas as câmaras legislativas têm maioria republicana ou quando não há maioria de nenhum dos partidos foram censurados, não adotaram a política.

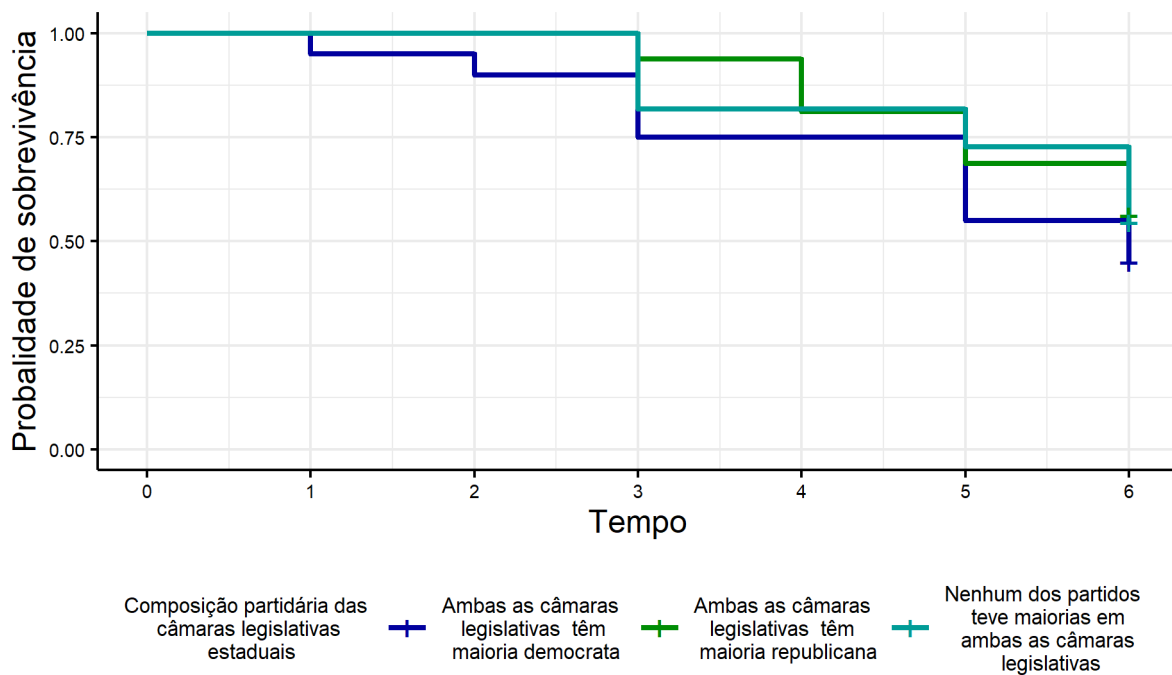


Figura 3: Gráfico da função de sobrevivência pela Composição partidária das câmaras legislativas estaduais.

A fim de avaliar as funções de sobrevivência entre a composição partidária das câmaras legislativas estaduais tem-se o gráfico acima que corrobora com as indicações anteriores. Nele é perceptível que a probabilidade de sobreviver (não adotar a política) é maior para os estados em que as câmaras legislativas têm maioria republicana ou naqueles estados em que nenhum dos partidos é maioria nas câmaras. Por outro lado, a probabilidade de sobreviver, é menor nos estados em que as câmaras legislativas têm maioria democrata.

## Número de vezes que a palavra "Educação" aparece na constituição estadual

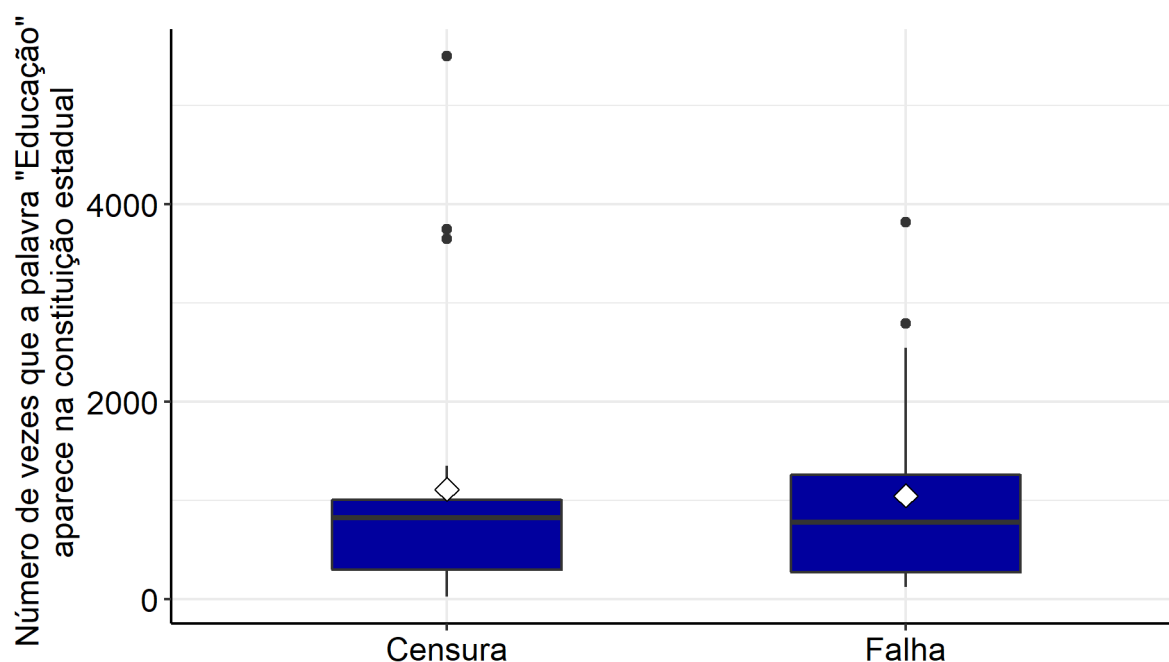


Figura 4: *Box plot* do Número de vezes que a palavra "Educação" aparece na constituição estadual pelo Status da adoção da política *Charter Schools*.

Por meio do gráfico acima é possível verificar a distribuição do número de vezes que a palavra "Educação" aparece nas constituições estaduais pela variável indicadora de falha, adoção da política, ou censura, não adoção. Observa-se que a média e a mediana do número da palavra "Educação" nas constituições federais estão próximas entre os estados que falharam e aqueles que foram censurados.

Entretanto, para os estados censurados o número de vezes que "Educação" aparece na constituição federal está variando menos e estes valores estão concentrados em números ligeiramente menores do que nos estados que falharam, ou seja, que adotaram a política. Nota-se ainda que há três estados censurados e dois estados que falharam com valores altos para a frequência da palavra "Educação".

### Medida estadual da ideologia do cidadão

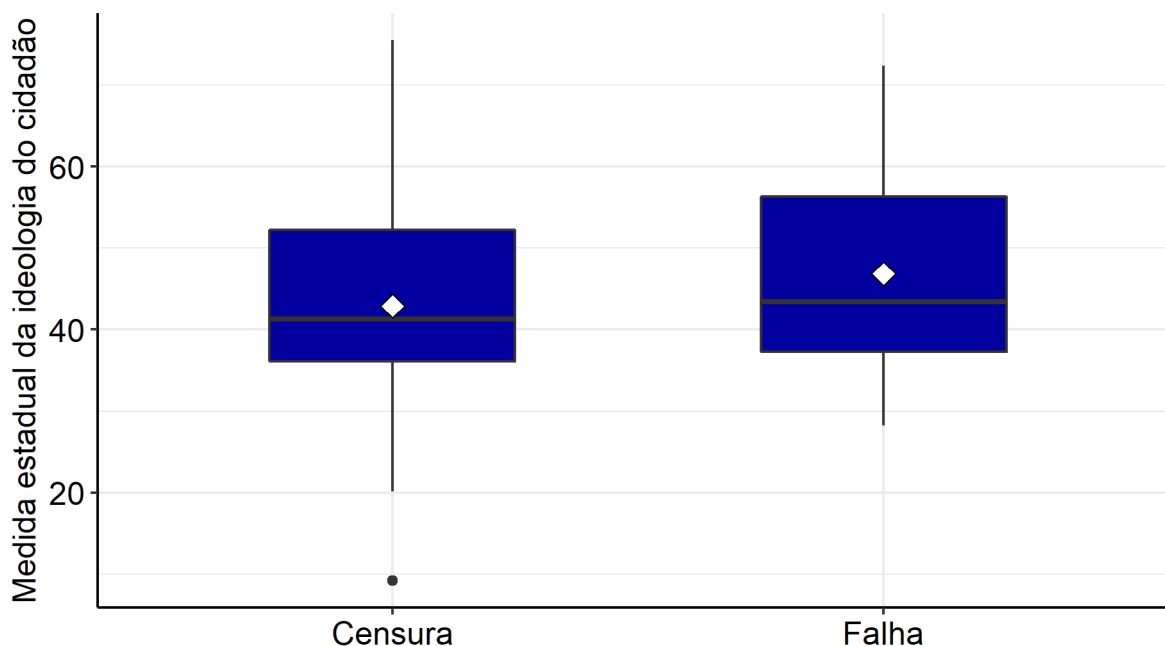


Figura 5: *Box plot* da Medida estadual da ideologia do cidadão pelo Status da adoção da política *Charter Schools*.

Segundo o *Box plot* da Medida estadual da ideologia do cidadão pelo Status da adoção da política *Charter Schools*, os estados com maior média para essa medida são os adotantes da política, com média de 46,84 contra uma média de 42,88 para os não adotantes. O mesmo ocorre com a mediana, que é ligeiramente maior para os adotantes (43,45) do que para os não adotantes (41,26).

Nota-se ainda que não há grandes diferenças na distribuição da medida estadual da ideologia do cidadão pela adoção ou não da política *Charter Schools*.

Dentre os estados censurados, há uma observação discrepante, com a medida estadual de ideologia do cidadão menor do que as demais, constituindo um mínimo de 9,25 para este grupo ao passo que o mínimo para os estados que falharam foi de 28,24.

## Medida estadual da ideologia do governo

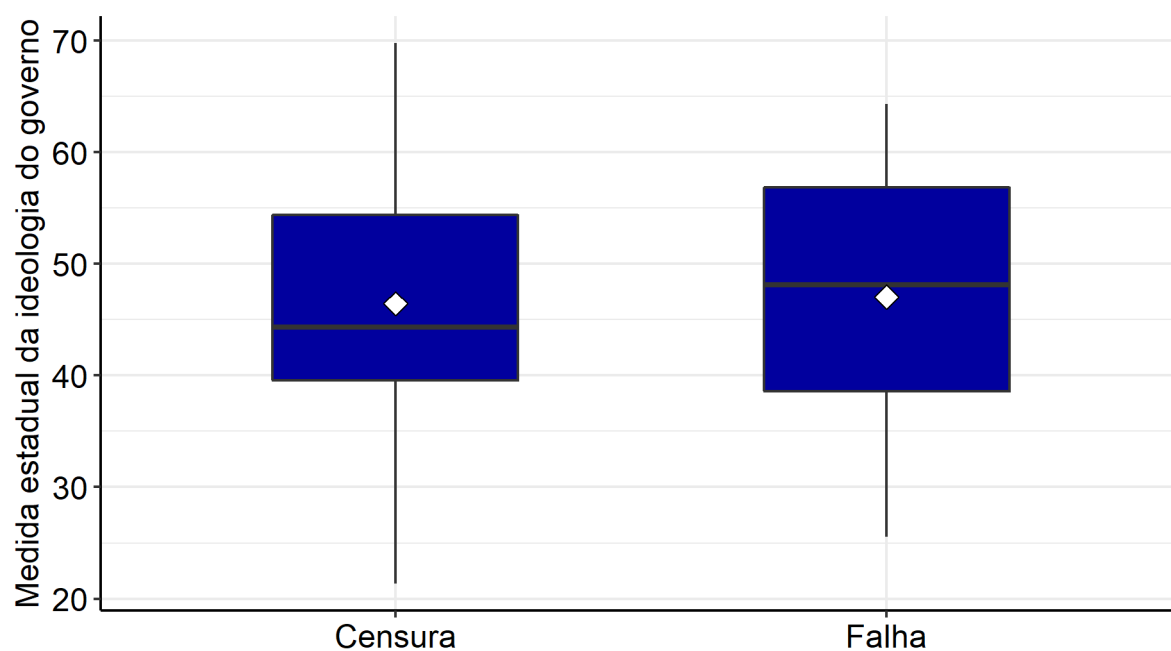


Figura 6: *Box plot* da Medida estadual da ideologia do governo pelo Status da adoção da política *Charter Schools*.

Constata-se por meio do gráfico 6 que a média da medida estadual da ideologia do governo é próxima entre os estados que adotaram a política *Charter Schools* (46,66) e os que não adotaram (47,03). Tal como a mediana com valores de 44,40 e 48,10, respectivamente. Além disso, a variância desta medida é maior nos estados que não adotaram a política, 146,46 e 122,89 para os adotantes.

## Profissionalismo legislativo do estado

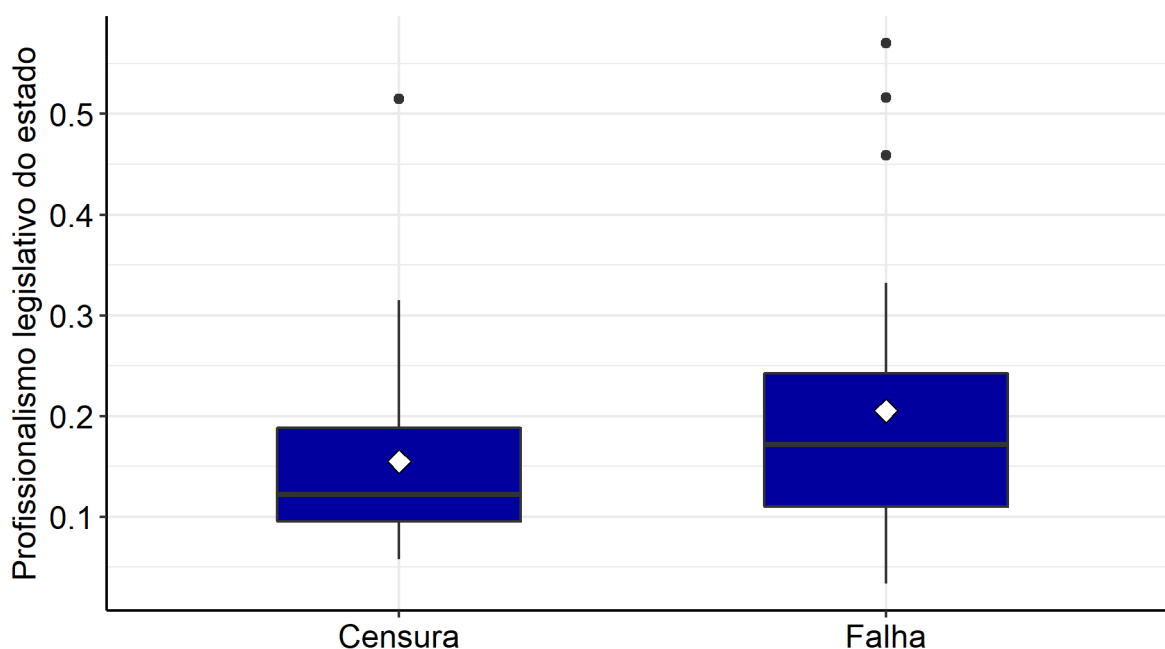


Figura 7: *Box plot* do Profissionalismo legislativo do estado pelo Status da adoção da política *Charter Schools*.

Segundo o gráfico que ilustra o comportamento da variável Profissionalismo legislativo estadual pelo Status da adoção da política *Charter Schools*, tem-se que os estados com maior média para essa medida são os adotantes da política, com média de 0,2054 contra uma média de 0,1558 para os não adotantes. O mesmo ocorre com a mediana, que é levemente maior para os adotantes (0,20) do que para os não adotantes (0,1270).

Ademais a distribuição do profissionalismo legislativo estadual para as observações de falha se apresenta mais dispersa do que nas observações de censura.

Dentre os estados censurados, há uma observação discrepante, com o profissionalismo legislativo do estado maior do que as demais e três dentre os estados que falharam.

## População

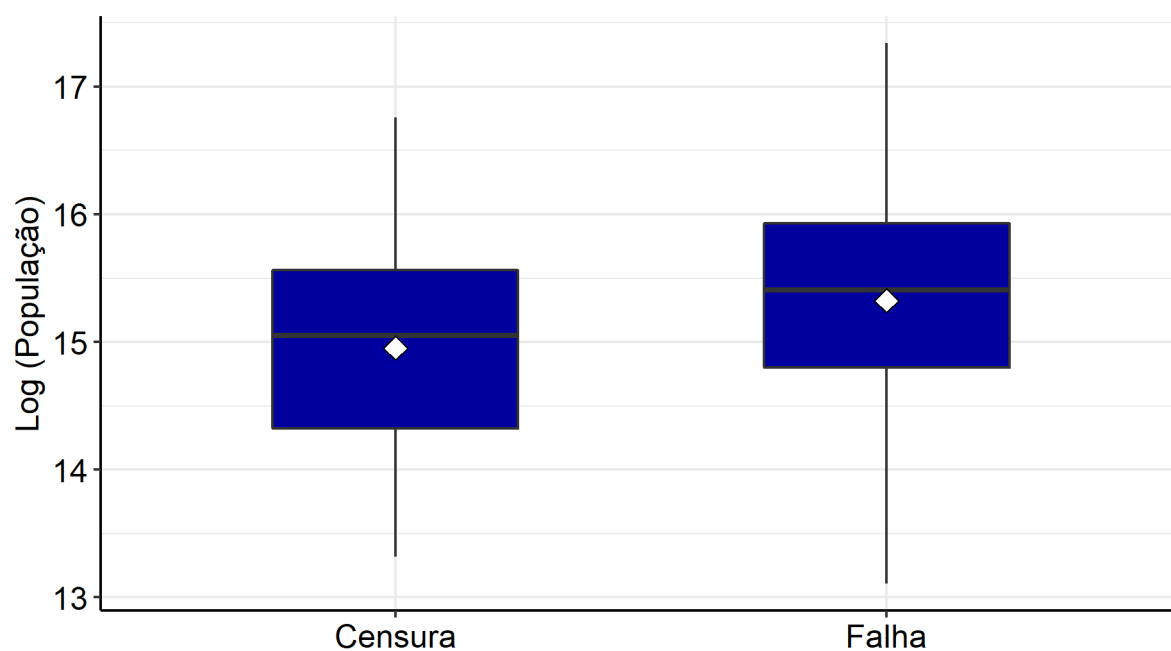


Figura 8: *Box plot* do Log (População) pelo Status da adoção da política *Charter Schools*.

Quanto a população do estado pelo Status da adoção da política *Charter Schools*, os estados adotantes da política possuem média populacional relativamente maior, o mesmo ocorre com a mediana.

Nota-se ainda que a distribuição da população estadual varia mais nos estados que falham.

### 6.1.2 High School Exit Exams

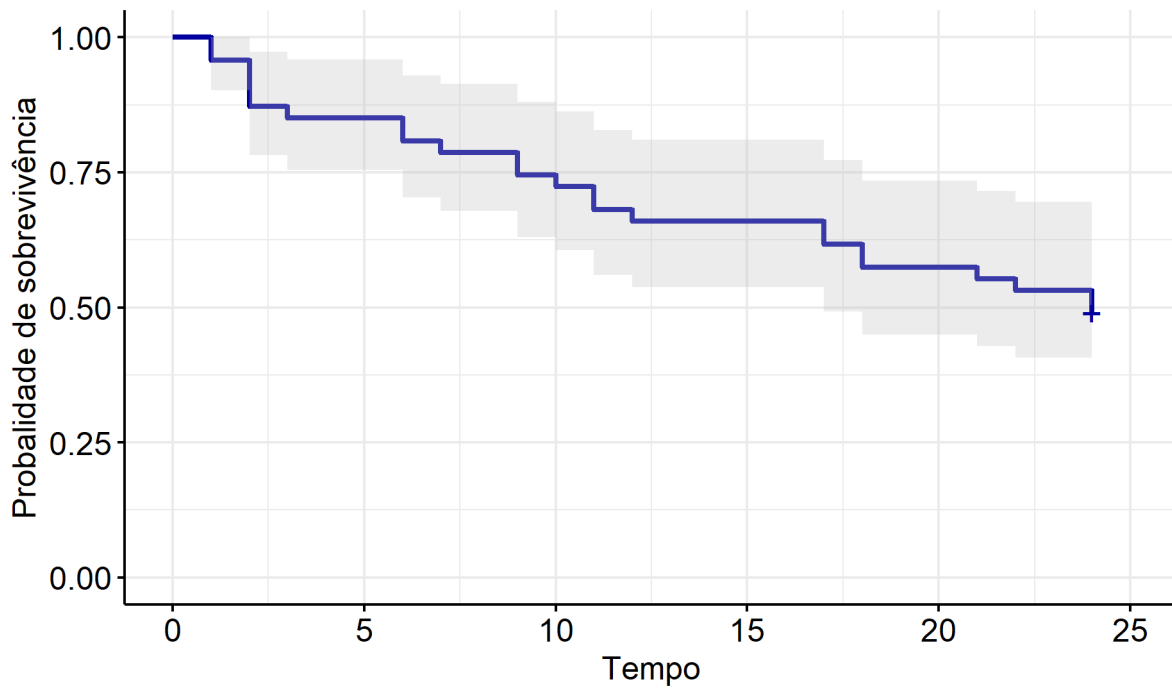


Figura 9: Gráfico da função de sobrevivência para a política *High School Exit Exams*.

Avaliando as adoções da política educacional *High School Exit Exams*, 24 estados falharam, isto é, adotaram a política e 24 foram censurados, não adotaram a política.

No geral a probabilidade de sobrevivência decai desde o tempo inicial até o tempo 24, ao final do período observado a probabilidade de sobreviver se aproxima de 0,50, em virtude das censuras no último tempo do estudo.



## Composição partidária das câmaras legislativas estaduais

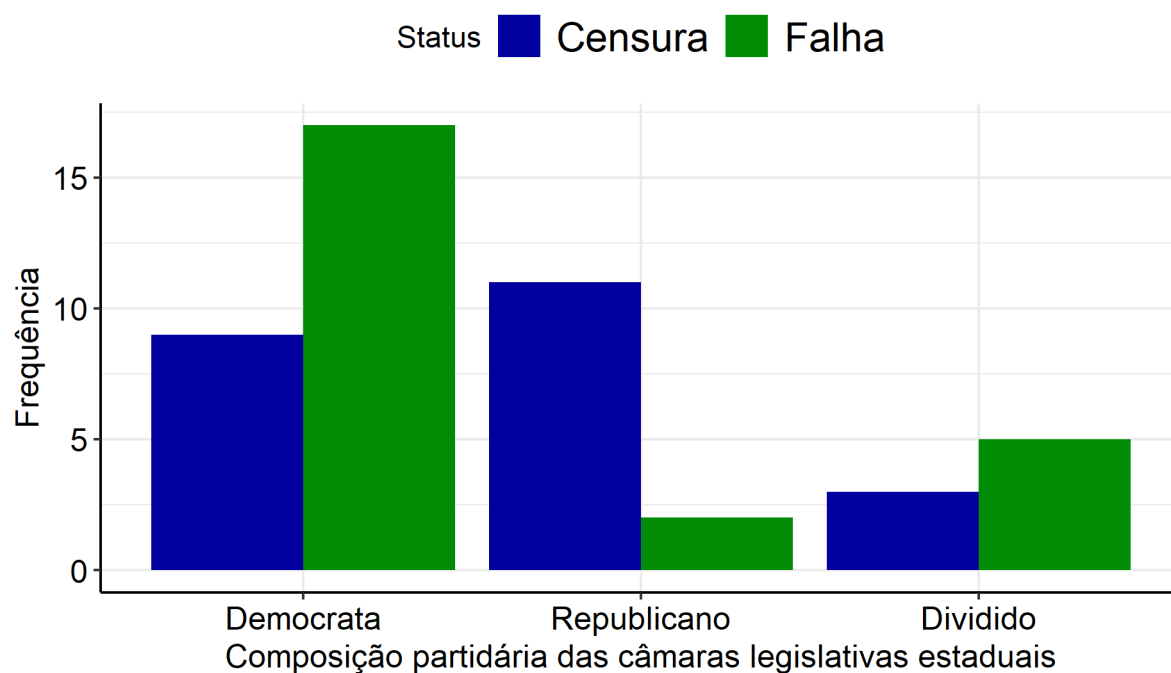


Figura 10: Gráfico de barras da Composição partidária das câmaras legislativas estaduais pelo Status da adoção da política *High School Exit Exams*.

Examinando o gráfico 10 da variável Composição partidária das câmaras legislativas estaduais pelo status da adoção da política *High School Exit Exams*, nota-se que a maioria dos estados em que ambas as câmaras legislativas têm maioria democrata ou quando não há maioria de nenhum dos partidos falharam, portanto adotaram a política. Enquanto a maioria dos estados em que ambas as câmaras legislativas têm maioria republicana foram censurados, não adotaram a política.

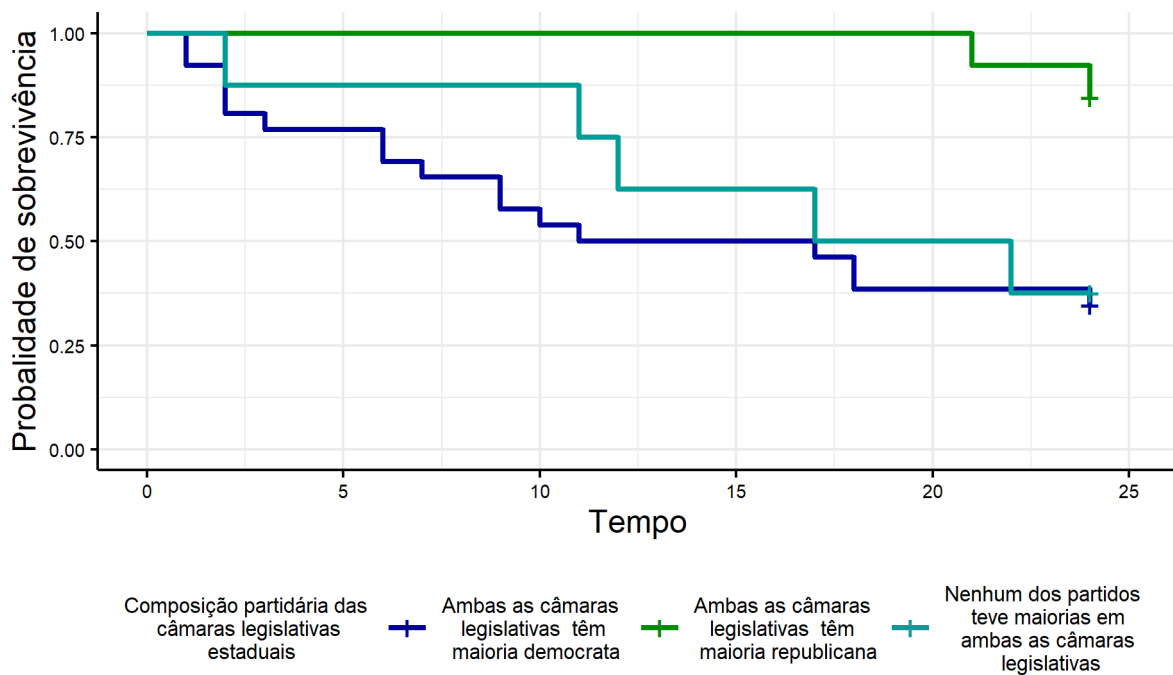


Figura 11: Gráfico da função de sobrevivência pela Composição partidária das câmaras legislativas estaduais.

Afim de avaliar as funções de sobrevivência entre a composição partidária das câmaras legislativas estaduais tem-se o gráfico acima que reforça as evidências anteriores.

Tem-se que a probabilidade de sobreviver, ou seja, de não adotar a política *High School Exit Exams*, é maior para os estados em que ambas as câmaras legislativas têm maioria republicana. Por outro lado, a probabilidade de sobreviver é menor nos estados em que ambas as câmaras legislativas têm maioria democrata e naqueles em que nenhum dos partidos teve maiorias em ambas as câmaras legislativas, ou seja, estes estados foram os que mais adotaram a política.

## Número de vezes que a palavra "Educação" aparece na constituição estadual

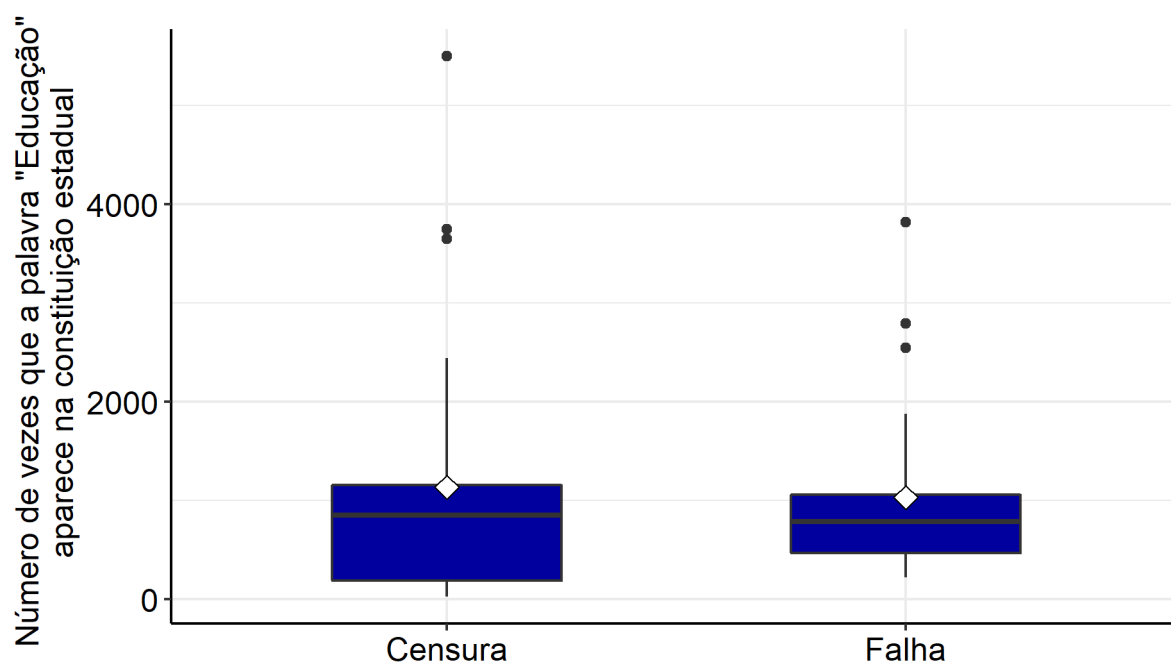


Figura 12: *Box plot* do Número de vezes que a palavra "Educação" aparece na constituição estadual pelo Status da adoção da política *High School Exit Exams*.

A figura acima, revela a distribuição do número de vezes que a palavra "Educação" aparece nas constituições estaduais nos estados americanos que adotaram ou não a política *High School Exit Exams*. A média e a mediana em ambos os grupos de estados são próximas para esta variável. Já a variância, todavia, é maior nos estados censurados.

Nota-se ainda que há três estados censurados e três estados que falharam com valores altos para a frequência da palavra "Educação".

## Medida estadual da ideologia do cidadão

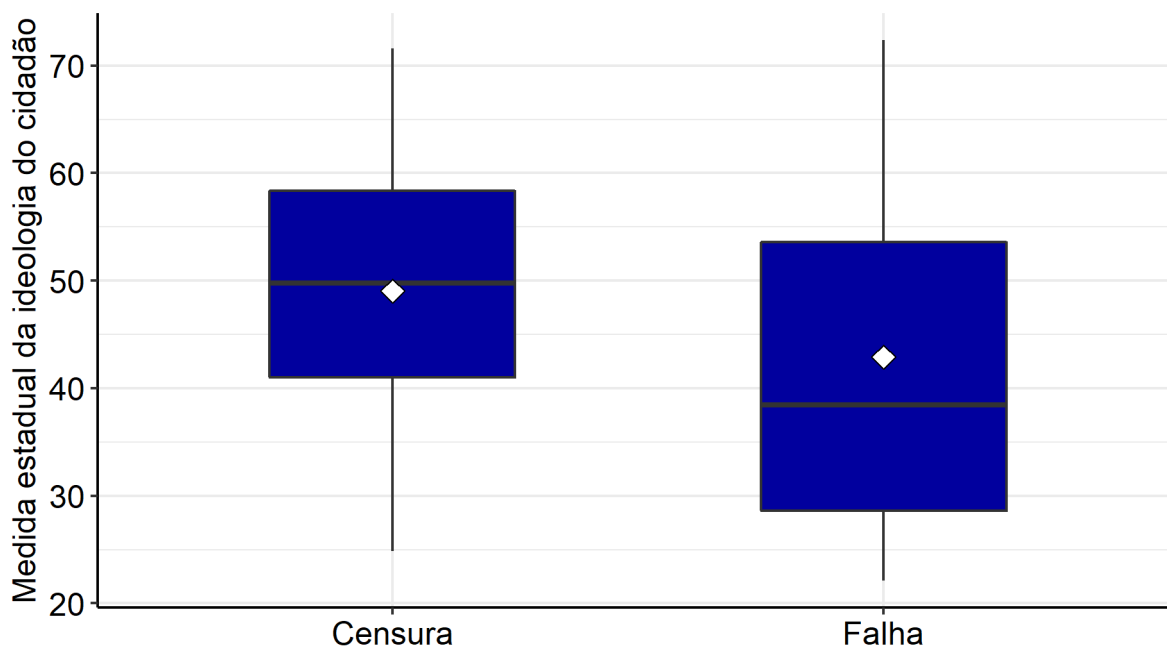


Figura 13: *Box plot* da Medida estadual da ideologia do cidadão pelo Status da adoção da política *High School Exit Exams*.

Segundo o *Box plot* da Medida estadual da ideologia do cidadão pelo Status da adoção da política *High School Exit Exams*, os estados com maior média para essa medida são os não adotantes da política, com média de 48,41 contra uma média de 42,90 para os adotantes, o oposto do comportamento na primeira política em análise. O mesmo ocorre com a mediana, que é ligeiramente maior para os não adotantes (49,06) do que para os adotantes (38,45).

## Medida estadual da ideologia do governo

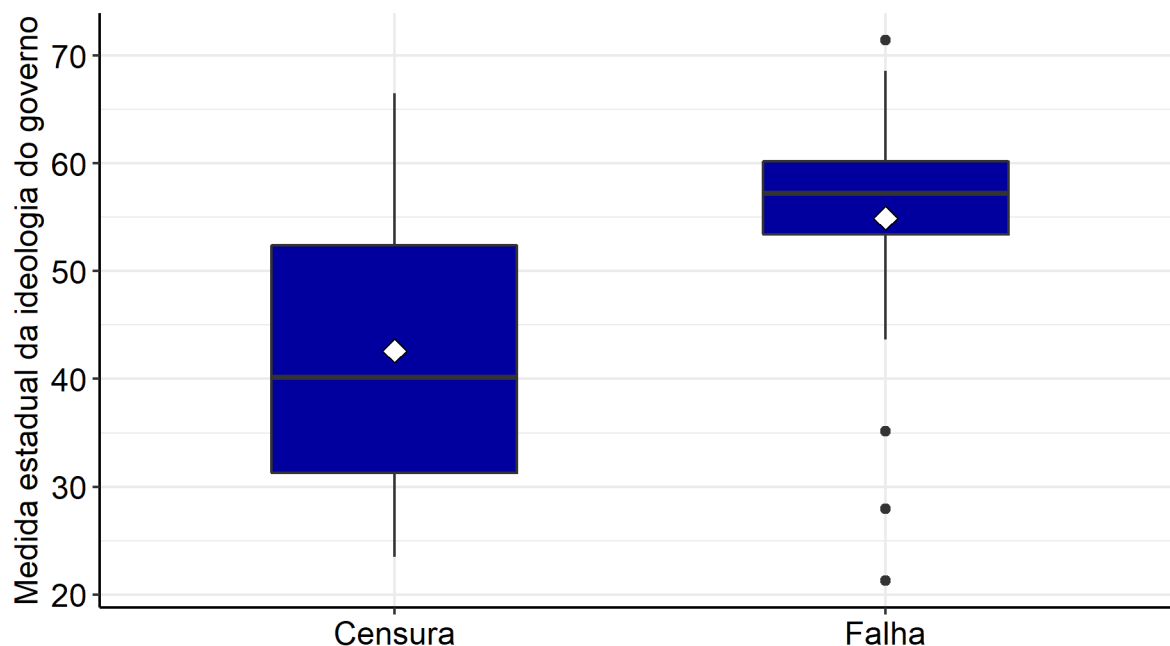


Figura 14: *Box plot* da Medida estadual da ideologia do governo pelo Status da adoção da política *High School Exit Exams*.

Considerando o gráfico 14, constata-se que a média da medida estadual da ideologia do governo é maior nos estados que adotaram a política *High School Exit Exams* (54,92) do que nos que não adotaram (42,30). Tal como a mediana com valores de 57,21 e 39,94, respectivamente.

Contudo a variância desta medida é maior nos estados que não adotaram a política, 172,06 e 148,08 para os adotantes. Sendo a distribuição no grupo de falha mais concentrada em valores em valores mais altos da medida.

Ainda assim dentre os estado que falharam há 4 observações distantes das demais, *outliers*, sendo 1 com valor alto da medida e 3 para valores pequenos.

### Profissionalismo legislativo do estado

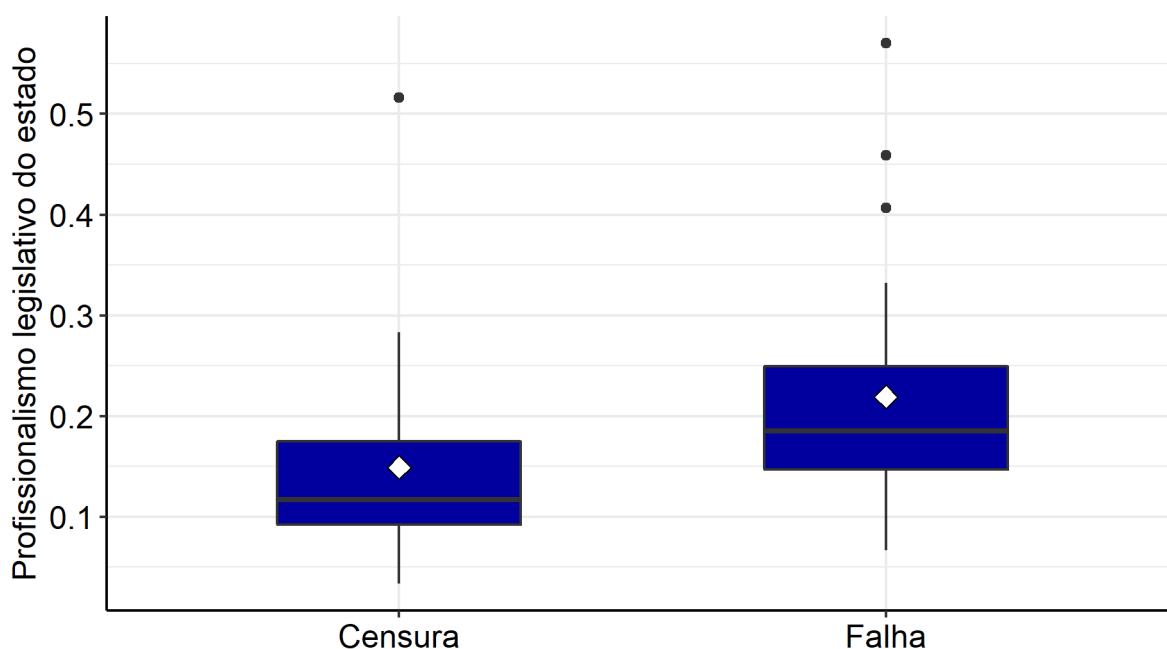


Figura 15: *Box plot* do Profissionalismo legislativo do estado pelo Status da adoção da política *High School Exit Exams*.

Em conformidade com o gráfico acima, que demonstra o comportamento da variável profissionalismo legislativo estadual pelo Status da adoção da política *High School Exit Exams*, os estados com maior média e mediana para essa variável são aqueles que implementaram a política.

Nota-se ainda que não há grandes diferenças na distribuição do profissionalismo legislativo estadual pela adoção ou não da política educacional em questão.

Dentre os estados censurados, há uma observação discrepante, com o profissionalismo legislativo do estado maior do que as demais e três dentre os estados que falharam.

## População

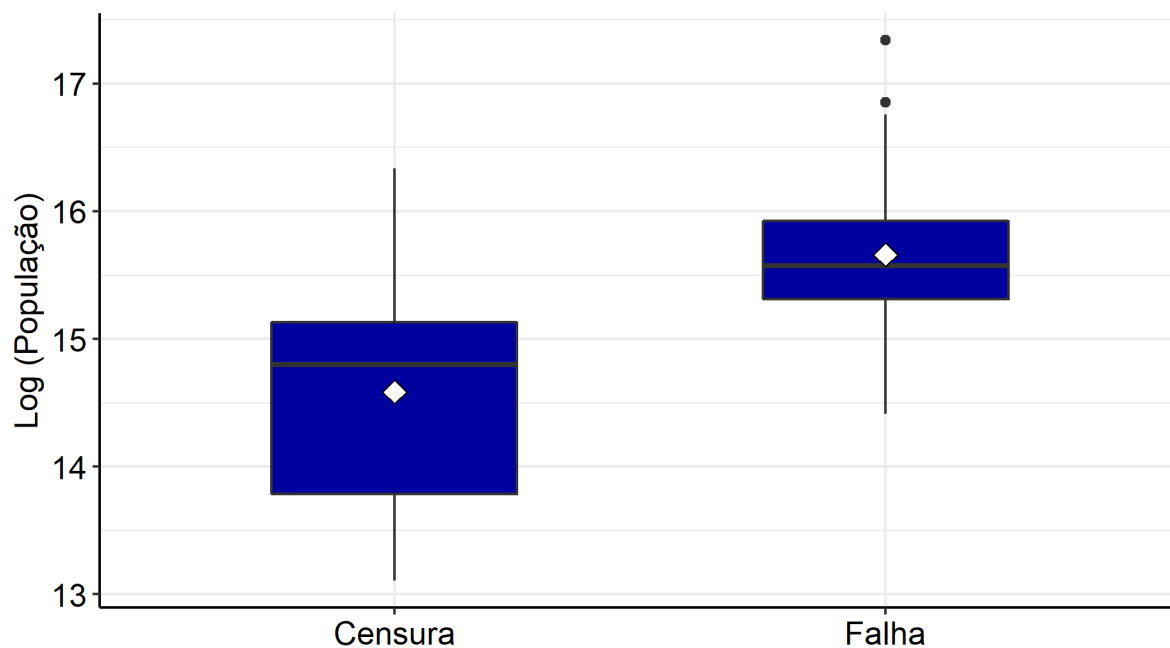


Figura 16: *Box plot* do Log (População) pelo Status da adoção da política *High School Exit Exams*.

Com relação a população do estado pelo Status da adoção da política *High School Exit Exams*, os estados adotantes da política possuem média populacional relativamente maior, o mesmo ocorre com a mediana.

Nota-se ainda que a distribuição da população estadual varia mais nos estados que falham.

Dentre os estados censurados, há uma observação discrepante, com população maior do que os demais, já nos estados que falharam há três observações discrepantes, com valores altos populacionais.

### 6.1.3 *School Choice*

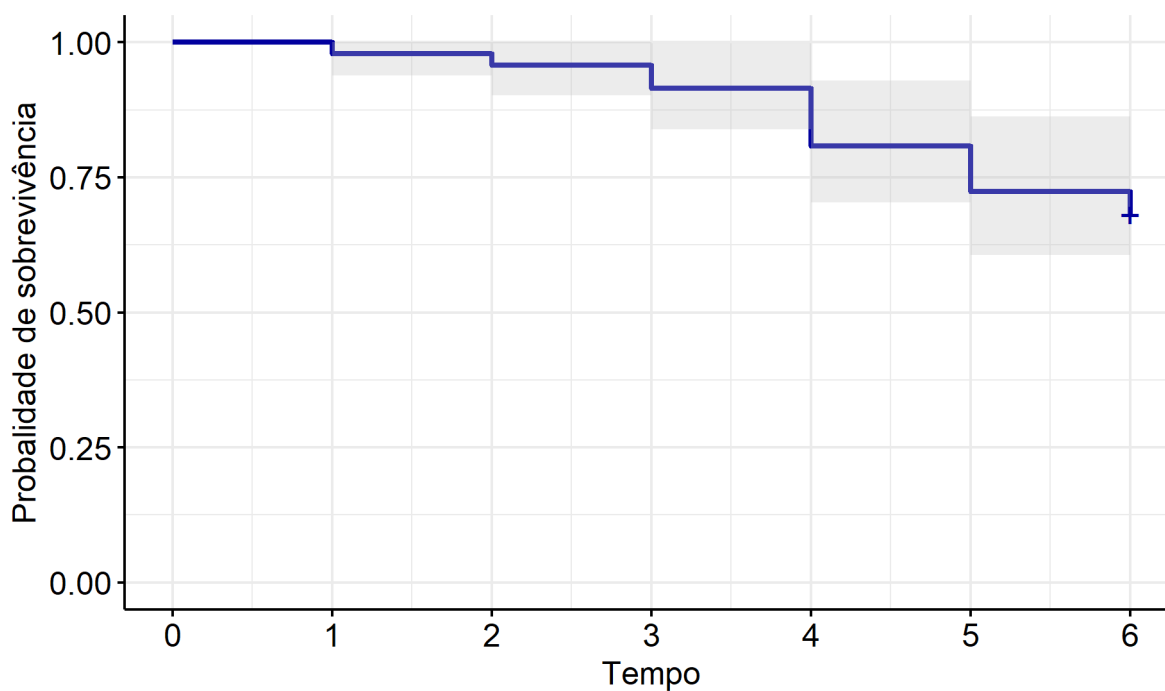


Figura 17: Gráfico da função de sobrevivência para a política *School Choice*.

Dos 48 estados americanos apenas 16 falharam, ou seja, adotaram a política *School Choice* e 32 foram censurados, não adotaram a política.

A probabilidade de sobrevivência apresenta uma pequena queda nos primeiros tempos, a partir dos tempos quatro e cinco esta probabilidade revela quedas maiores. Totalizando ao final do estudo uma probabilidade de sobrevivência de 0,67, devido as censuras contabilizadas.



## Composição partidária das câmaras legislativas estaduais

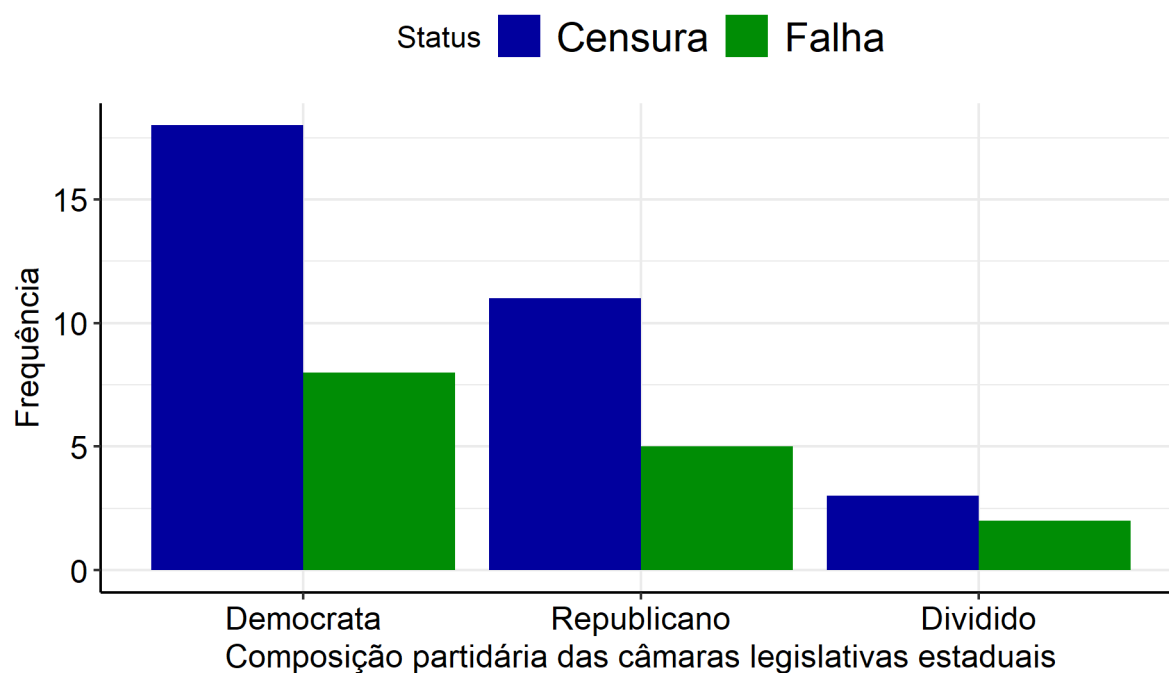


Figura 18: Gráfico de barras da Composição partidária das câmaras legislativas estaduais pelo Status da adoção da política *School Choice*.

Examinando o gráfico 18 da variável Composição partidária das câmaras legislativas estaduais pelo status da adoção da política *School Choice*, nota-se que independente da composição partidária a maioria dos estados nos grupos não adotaram a política.

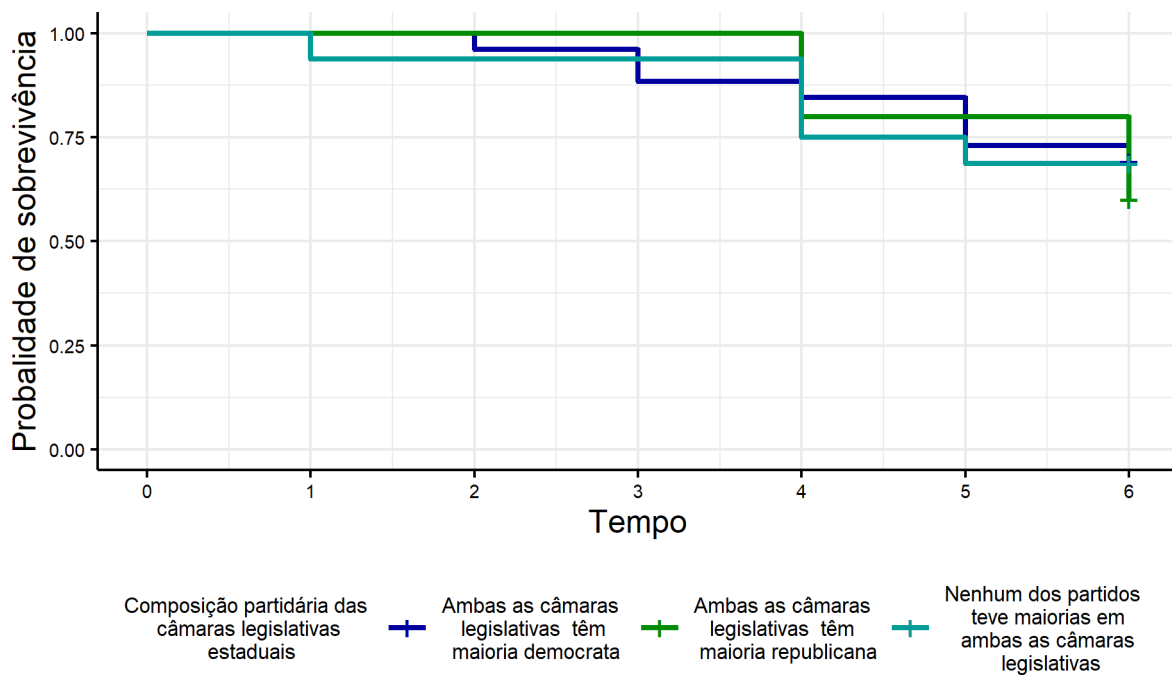


Figura 19: Gráfico da função de sobrevivência pela Composição partidária das câmaras legislativas estaduais.

Avaliando as funções de sobrevivência entre a composição partidária das câmaras legislativas estaduais tem-se o gráfico acima que indica que a probabilidade de sobreviver (não adotar a política) ao longo do tempo tem comportamento semelhante independente da composição partidária das câmaras legislativas.

## Número de vezes que a palavra "Educação" aparece na constituição estadual

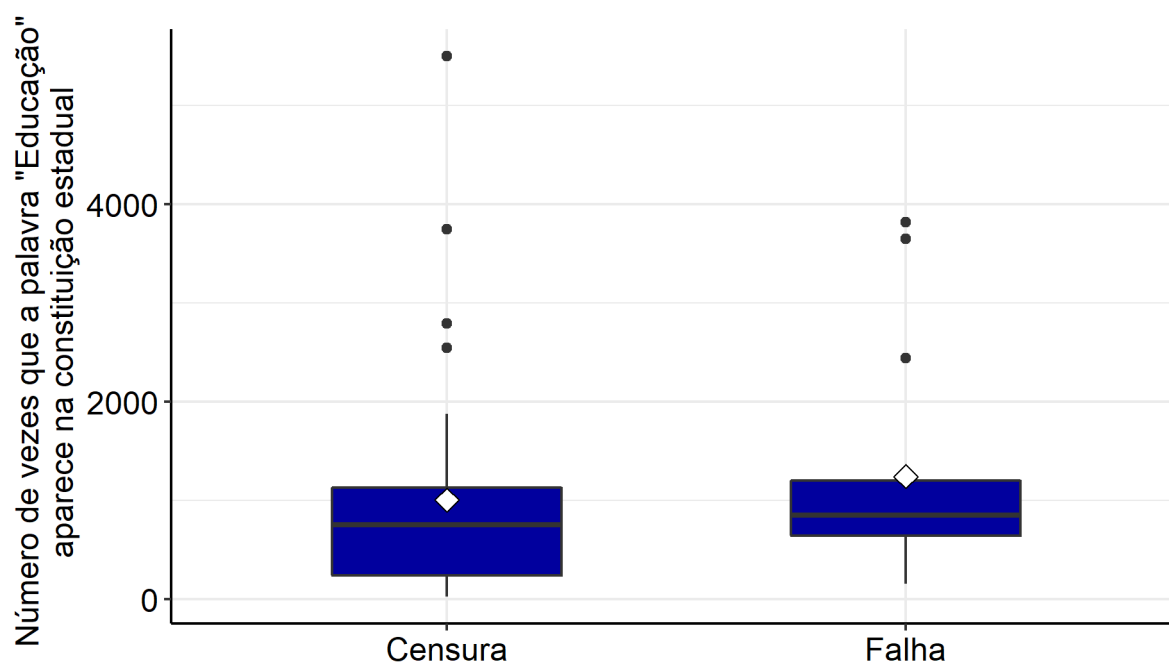


Figura 20: *Box plot* do Número de vezes que a palavra "Educação" aparece na constituição estadual pelo Status da adoção da política *School Choice*.

Por meio do gráfico acima é possível verificar a distribuição do número de vezes que a palavra "Educação" aparece nas constituições estaduais pela variável indicadora de falha, adoção da política, ou censura, não adoção. Observa-se que a média e a mediana do número da palavra "Educação" nas constituições federais estão próximas entre os estados que falharam e aqueles que foram censurados.

Entretanto, para os estados censurados o número de vezes que "Educação" aparece na constituição federal parece variar menos do que nos estados que falharam. Nota-se ainda que há três estados censurados e dois estados que falharam com valores altos para a frequência da palavra "Educação".

## Medida estadual da ideologia do cidadão

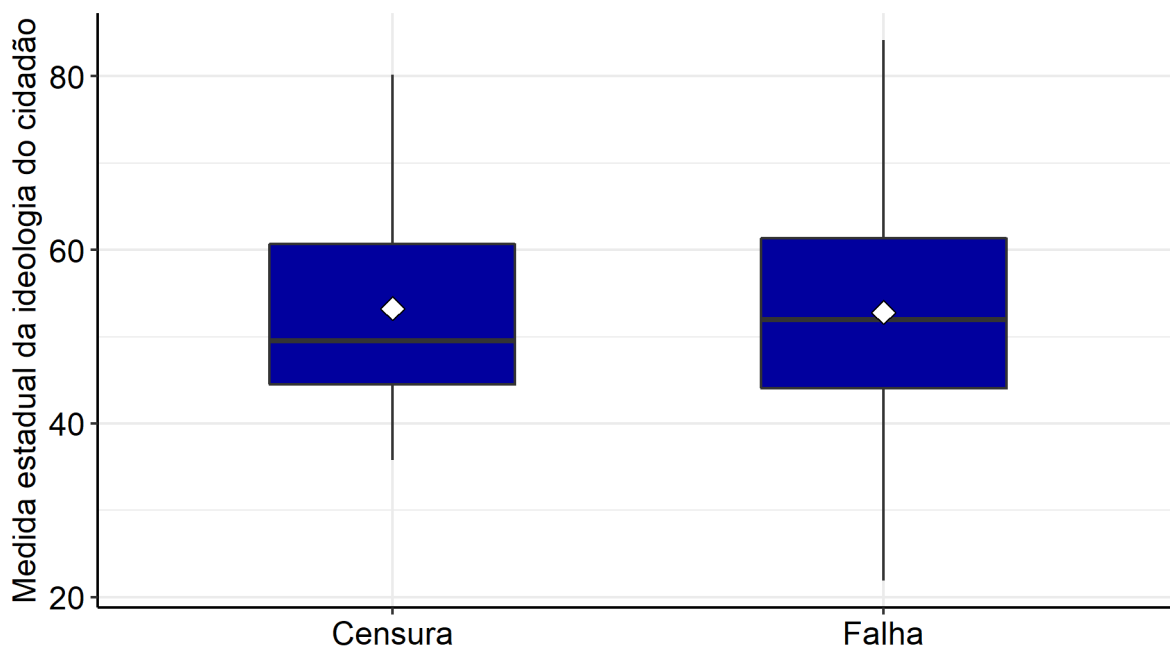


Figura 21: *Box plot* da Medida estadual da ideologia do cidadão pelo Status da adoção da política *School Choice*.

Segundo o *Box plot* da Medida estadual da ideologia do cidadão pelo Status da adoção da política *School Choice*, não há diferença significativa na distribuição da medida pela adoção ou não da política. Os valores para a média, mediana, 1º e 3º quartil são próximos entre os estados adotantes e não adotantes.

## Medida estadual da ideologia do governo

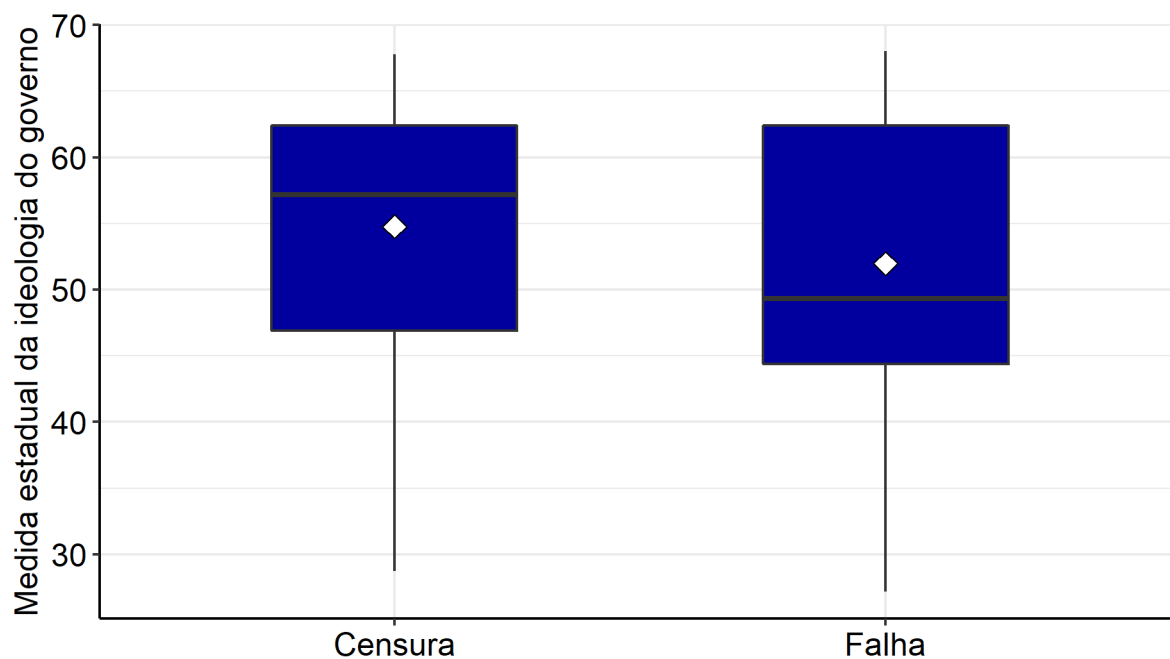


Figura 22: *Box plot* da Medida estadual da ideologia do governo pelo Status da adoção da política *School Choice*.

Constata-se por meio do gráfico 6 que a média da medida estadual da ideologia do governo é próxima entre os estados que adotaram a política *School Choice* (51,52) e os que não adotaram (54,76). Tal como a mediana com valores de 49,27 e 57,14, respectivamente. Além disso, a variância desta medida é maior nos estados que adotaram a política, 134,44 e 86,66 para os não adotantes.

### Profissionalismo legislativo do estado

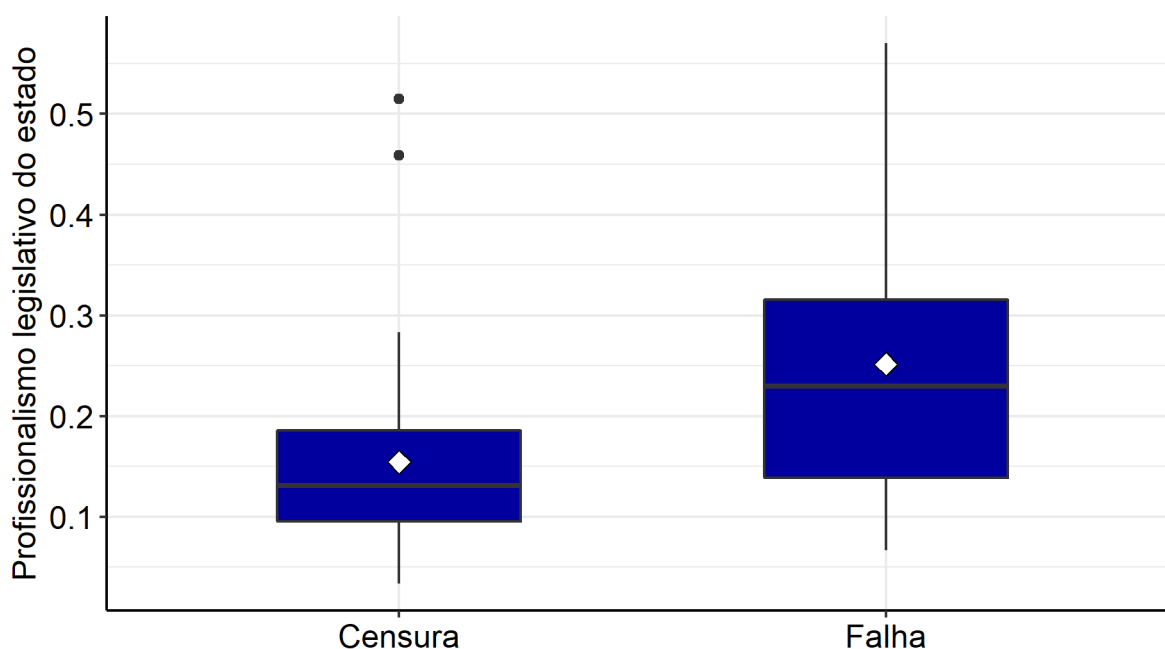


Figura 23: *Box plot* do Profissionalismo legislativo do estado pelo Status da adoção da política *School Choice*.

Segundo o gráfico que ilustra o comportamento da variável Profissionalismo legislativo estadual pelo Status da adoção da política *School Choice*, tem-se que os estados com maior média para essa medida são os adotantes da política, com média de 0,2467 contra uma média de 0,1545 para os não adotantes. O mesmo ocorre com a mediana, que é levemente maior para os adotantes (0,2275) do que para os não adotantes (0,13100).

Ademais a distribuição do profissionalismo legislativo estadual para as observações de falha se apresenta mais dispersa e mais presente em valores maiores do que nas observações de censura.

Dentre os estados censurados, há duas observação discrepantes, com o profissionalismo legislativo do estado maior do que as demais.

## População

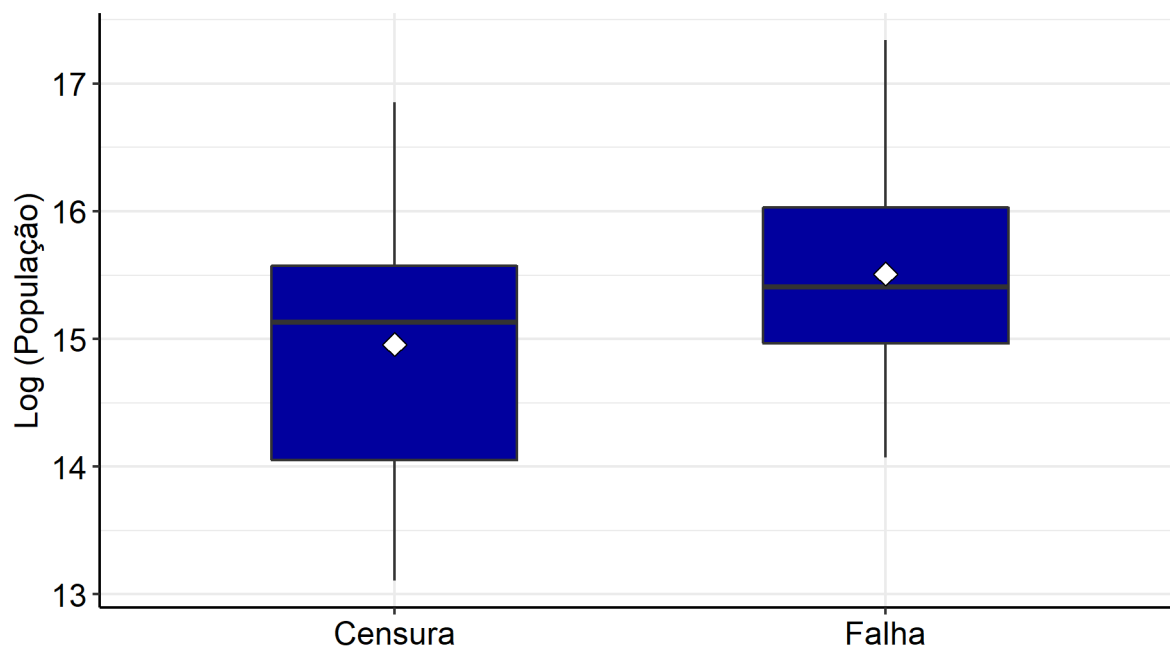


Figura 24: *Box plot* do Log (População) pelo Status da adoção da política *School Choice*.

Quanto a população do estado pelo Status da adoção da política *School Choice*, os estados adotantes da política possuem média populacional relativamente maior, o mesmo ocorre com a mediana.

Nota-se ainda que a distribuição da população estadual varia mais nos estados que falharam.

## 6.2 Modelo de Riscos Proporcionais de Cox

Os dados deste estudo referem-se à informações de 48 estados americanos e para cada um deles é considerado o tempo de adesão/censura a cada uma das três políticas educacionais dos Estados Unidos, *Charter Schools*, *High School Exit Exams* e *School Choice*. Além disso, para cada estado as seguintes características foram avaliadas, a composição partidária das câmaras legislativas, o número de vezes que a palavra "Educação" aparece na constituição estadual, a medida da ideologia do cidadão, a medida da ideologia do governo, o profissionalismo legislativo e a população.

Na seção anterior foi realizada a análise descritiva e exploratória das variáveis explicativas e em seguida será utilizado o modelo de regressão de Cox, para modelar o tempo de adesão as políticas em função das covariáveis registradas no estudo. O modelo

de regressão de Cox permite a análise de dados provenientes de estudos de tempo de vida em que a resposta é o tempo até a ocorrência de um evento de interesse, ajustando por covariáveis (Colosimo e Giolo, 2006).

O objetivo é verificar quais as covariáveis influenciam de maneira significativa o tempo de sobrevivência do indivíduo. Portanto, ajustando o modelo de Cox para análise desses dados, primeiramente modelos individuais para cada uma das variáveis explicativas, encontram-se os seguintes resultados apresentados na tabela a seguir.

Tabela 1: Modelos de Cox para a política *Charter Schools*.

<i>Charter Schools</i>	Estimativa	Erro-Padrão	Valor-p
Composição partidária das câmaras legislativas estaduais			
Ambas têm maioria republicana	-0,39	0,48	0,424
Nenhum dos partidos teve maiorias	-0,34	0,54	0,528
Número de vezes que a palavra "Educação"			
aparece na constituição estadual	2,49 e-05	1,76 e-04	0,887
Medida estadual da ideologia do cidadão	0,02	0,02	0,285
Medida estadual da ideologia do governo	0,003	0,02	0,887
Profissionalismo legislativo	3,19	1,60	0,046 *
População	6,41 e-08	3,27 e-08	0,050 *

Examinando os p-valores da tabela acima, observa-se que aqueles menores do que o nível de significância de 10% são os referentes as variáveis profissionalismo legislativo e população do estado, desse modo tais covariáveis influenciam individualmente no tempo até a adesão estadual da política *Charter Schools*. Para as demais covariáveis não há indicativo de influência no tempo de adesão da política em questão.

Vale ressaltar que para o modelo com a composição partidária das câmaras legislativas estaduais foi desconsiderado o estado Nebraska por ser uma legislatura unicameral apartidária.

Tabela 2: Modelos de Cox para a política *High School Exit Exams*.

<i>High School Exit Exams</i>	Estimativa	Erro-Padrão	Valor-p
Composição partidária das câmaras legislativas estaduais			
Nenhum dos partidos teve maiorias	-0,24	0,51	0,632
Número de vezes que a palavra "Educação"			
aparece na constituição estadual	-0,0001	0,0002	0,543
Medida estadual da ideologia do cidadão	-0,02	0,01	0,125
Medida estadual da ideologia do governo	0,05	0,017	0,001*
Profissionalismo legislativo	2,19	1,35	0,105
População	5,11 e-08	2,03 e-08	0,012*



Para a política *High School Exit Exams* há indícios das variáveis medida estadual da ideologia do governo e a população interferirem no tempo de adesão do estado, segundo o p-valor dos modelos individuais. Além do mais, os p-valores para a medida estadual da ideologia do cidadão e para o profissionalismo legislativo são próximos do nível de significância (10%).

Tabela 3: Modelos de Cox para a política *School Choice*.

<i>School Choice</i>	Estimativa	Erro-Padrão	Valor-p
Composição partidária das câmaras legislativas estaduais	0.23	0,79	0,770
Ambas têm maioria republicana	0.06	0,57	0,921
Nenhum dos partidos teve maiorias			
Número de vezes que a palavra "Educação"			
aparece na constituição estadual	0,00009	0,0002	0,647
Medida estadual da ideologia do cidadão	-0,001	0,02	0,944
Medida estadual da ideologia do governo	-0,02	0,02	0,414
Profissionalismo legislativo	3,10	1,48	0,036 *
População	2,64 e-08	3,10 e-08	0,394

Avaliando as variáveis explicativas de forma individual para a adesão da política *School Choice* apenas o profissionalismo legislativo apontou ser significativo com p-valor menor que o nível de significância.

O próximo passo para identificar quais covariáveis influenciam de forma significativa no tempo de sobrevivência do estado é construir um modelo com todas as covariáveis e retirar aquelas com os maiores p-valores, uma de cada vez, afim de se chegar a um modelo final. Em cada passo deste processo de seleção será verificado o nível de significância das covariáveis separadamente. O passo a passo deste procedimento está descrito na tabela a seguir.

Tabela 4: Seleção de covariáveis para o modelo de regressão de Cox para a política *Charter Schools*.

Passos	Modelo	Estimativa	Erro padrão	Valor-p
Passo 1	Câmaras legislativas com maioria republicana	-7,86 e-01	6,30 e-01	0,212
	Câmaras legislativas sem maioria	-9,18 e-01	6,69 e-01	0,170
	Nº "Educação" na constituição estadual	8,45 e-05	1,88 e-04	0,652
	Medida estadual da ideologia do cidadão	2,12 e-02	2,13 e-02	0,320
	Medida estadual da ideologia do governo	-2,08 e-02	2,81 e-02	0,459
	Profissionalismo legislativo	2,23 e+00	3,16 e+00	0,480
	População	2,44 e-08	5,48 e-08	0,656
Passo 2	Câmaras legislativas com maioria republicana	-8,09 e-01	6,38 e-01	0,205
	Câmaras legislativas sem maioria	-9,58 e-01	6,65 e-01	0,150
	Nº "Educação" na constituição estadual	7,87 e-05	1,86 e-04	0,672
	Medida estadual da ideologia do cidadão	1,94 e-02	2,09 e-02	0,354
	Medida estadual da ideologia do governo	-2,10 e-02	9,79 e-01	0,462
	Profissionalismo legislativo	3,31 e+00	2,01 e+00	0,101
Passo 3	Câmaras legislativas com maioria republicana	-0,80	0,64	0,215
	Câmaras legislativas sem maioria	-1,0	0,66	0,130
	Medida estadual da ideologia do cidadão	0,02	0,02	0,414
	Medida estadual da ideologia do governo	-0,02	0,03	0,477
	Profissionalismo legislativo	3,43	2,03	0,091*
Passo 4	Câmaras legislativas com maioria republicana	-0,50	0,49	0,307
	Câmaras legislativas sem maioria	-0,83	0,62	0,176
	Medida estadual da ideologia do cidadão	0,01	0,02	0,633
	Profissionalismo legislativo	3,84	1,96	0,051*
Passo 5	Câmaras legislativas com maioria republicana	-0,50	0,49	0,311
	Câmaras legislativas sem maioria	-0,80	0,61	0,191
	Profissionalismo legislativo	4,17	1,85	0,024*
Passo 6	Profissionalismo legislativo	3,19	1,59	0,048*

Logo, permaneceu no modelo final para a adoção da política *Charter Schools* apenas a covariável profissionalismo legislativo, revelando ser a única influenciando a implementação da política nos estados. O coeficiente do modelo indica que quanto maior o profissionalismo legislativo, maior é o risco do estado falhar, ou seja, o risco de adotar a política. Sendo assim, o aumento em uma unidade do profissionalismo legislativo, acarreta em um aumento no risco de 23,28.

Tabela 5: Seleção de covariáveis para o modelo de regressão Cox para a política *High School Exit Exams*.

<b>Passos</b>	<b>Modelo</b>	<b>Estimativa</b>	<b>Erro padrão</b>	<b>Valor-p</b>
Passo 1	Câmaras legislativas com maioria republicana	-7,80 e-01	1,03 e+00	0,450
	Câmaras legislativas sem maioria	6,93 e-03	5,88 e-0	0,991
	Nº "Educação" na constituição estadual	-4,37 e-04	2,96 e-04	0,140
	Medida estadual da ideologia do cidadão	-6,10 e-02	1,92 e-02	0,001*
	Medida estadual da ideologia do governo	5,30 e-02	3,03 e-02	0,081*
	Profissionalismo legislativo	4,28 e+00	2,80 e+00	0,126
	População	9,76 e-09	5,01 e-08	0,846
Passo 2	Nº "Educação" na constituição estadual	-5,21 e-04	3,19 e-04	0,102
	Medida estadual da ideologia do cidadão	-6,31 e-02	1,85 e-02	0,001*
	Medida estadual da ideologia do governo	7,10 e-02	2,20 e-02	0,001*
	Profissionalismo legislativo	4,16 e+00	2,88 e+00	0,148
	População	1,41 e-08	5,24 e-08	0,788
Passo 3	Nº "Educação" na constituição estadual	-4,81 e-04	2,71 e-04	0,076*
	Medida estadual da ideologia do cidadão	-6,39 e-02	1,82 e-02	0,0005*
	Medida estadual da ideologia do governo	7,32 e-02	2,04 e-02	0,0003*
	Profissionalismo legislativo	4,67 e+00	2,12 e+00	0,027*

Após o processo de seleção, o modelo de Cox resultante para o tempo de sobrevivência para a adesão a política *High School Exit Exams*, incluiu o seguinte conjunto de variáveis explicativas, Nº de vezes que a palavra "Educação" aparece nas constituições estaduais, Medida estadual da ideologia do cidadão, Medida estadual da ideologia do governo e Profissionalismo legislativo.

A Tabela acima mostra que valores mais altos da medida estadual da ideologia do governo e do profissionalismo legislativo aumentam o risco de adesão a política. O inverso acontece com as covariáveis Nº que a palavra "Educação" aparece nas constituições estaduais e a Medida estadual da ideologia do cidadão.

Dessa forma, o modelo indica que o aumento da presença de uma palavra "Educação" na constituição estadual diminui em aproximadamente 0,99 o risco do estado falhar. O aumento em uma unidade da Medida estadual da ideologia do cidadão diminui o risco em 0,94. Por outro lado, o aumento em uma unidade da Medida estadual da ideologia do governo aumenta em 1,08 o risco de um determinado estado falhar. Com relação ao profissionalismo legislativo, se aumentado uma unidade o risco de falha do estado aumenta 106,45.

Tabela 6: Seleção de covariáveis para o modelo de regressão Cox para a política *School Choice*.

Passos	Modelo	Estimativa	Erro padrão	Valor-p
Passo 1	Câmaras legislativas com maioria republicana	2,84 e-01	1,04 e+0	0,784
	Câmaras legislativas sem maioria	-3,57 e-02	5,95 e-01	0,952
	Nº "Educação" na constituição estadual	2,45 e-05	2,31 e-04	0,915
	Medida estadual da ideologia do cidadão	-1,28 e-02	3,05 e-02	0,675
	Medida estadual da ideologia do governo	-4,19 e-03	3,97 e-02	0,916
	Profissionalismo legislativo	5,22 e+0	2,62 e+00	0,047*
	População	-4,15 e-08	5,04 e-08	0,411
Passo 2	Nº "Educação" na constituição estadual	4,87 e-05	2,10 e-04	0,8166
	Medida estadual da ideologia do cidadão	-1,49 e-02	2,92 e-02	0,6106
	Medida estadual da ideologia do governo	-1,23 e-02	3,14 e-02	0,6950
	Profissionalismo legislativo	5,16 e+00	2,44 e+00	0,034*
	População	-4,78 e-08	4,93 e-08	0,333
Passo 3	Medida estadual da ideologia do cidadão	-1,62 e-02	2,87 e-02	0,572
	Medida estadual da ideologia do governo	-1,22 e-02	3,13 e-02	0,698
	Profissionalismo legislativo	5,16 e+00	2,43 e+00	0,034*
	População	-4,55 e-0	4,89 e-08	0,352
Passo 4	Medida estadual da ideologia do cidadão	-2,22 e-02	2,40 e-02	0,354
	Profissionalismo legislativo	5,51e+00	2,28 e+00	0,016*
	População	-5,19 e-08	4,65 e-08	0,264
Passo 5	Profissionalismo legislativo	4,59 e+00	2,12 e+0	0,031*
	População	-4,16 e-08	4,70 e-08	0,376
Passo 6	Profissionalismo legislativo	3,10	1,48	0,036

Considerando os passos da implementação estratégica da seleção de covariáveis, apresentados acima. Após o processo e seleção para o modelo de Cox resultante para o tempo de sobrevivência até a adesão da política *School Choice* teve como covariável apenas o Profissionalismo legislativo. O modelo encontrado indica que quanto maior o profissionalismo legislativo, maior é o risco do estado de adotar a política. Portanto, o aumento em uma unidade do profissionalismo legislativo, provoca um aumento no risco de de 22,16.

### 6.3 Diagnóstico do Modelo

Com os modelos ajustados para a adesão de cada uma das três políticas educacionais é necessário avaliar a suposição de riscos proporcionais do modelo de Cox e a qualidade do ajuste.

A suposição de riscos proporcionais deve ser atendida para que o modelo de Cox

possa ser considerado adequado aos dados desse estudo. Afim de avaliar esta suposição será utilizado o teste de hipótese proposto por Cox (1979) que compreende as seguintes hipóteses:

H0: As funções de riscos são proporcionais.

H1: As funções de riscos não são proporcionais.

A seguir, são apresentados as estatísticas do teste e p-valores para cada um dos modelos.

Tabela 7: Resultados do teste para os riscos proporcionais de Cox.

Modelo	Estimativa	Valor-p
<i>Charter School (Global)</i>	1,35	0,24
<i>High School Exit Exams (Global)</i>	7,49	0,112
Nº "Educação" na constituição estadual	2,84	0,092
Medida estadual da ideologia do cidadão	0,45	0,503
Medida estadual da ideologia do governo	0,22	0,638
Profissionalismo legislativo	2,02	0,155
<i>School Choice (Global)</i>	1,14	0,29

A partir dos p-valores acima, assume-se que os modelos finais possuem riscos proporcionais, inclusive cada covariável individualmente, exceto a variável que indica o número de vezes que a palavra "Educação" aparece nas constituições estaduais no modelo para o política *High School Exit Exams*, pois possui p-valor ligeiramente menor que o nível de significância de 10% adotado no estudo. Dessa forma, é possível constatar que os modelos atendem a suposição de riscos proporcionais.

Além da suposição de riscos proporcionais, há interesse em examinar outro aspecto do modelo de Cox. Sendo ele, verificar o ajuste global do modelo ajustado. Com esta finalidade será utilizado os resíduos de Cox-Snell, segundo Lawless (2003), os resíduos de Cox-Snell vêm de uma população homogênea e devem seguir uma distribuição exponencial com média 1. A Figura a seguir compara graficamente a função de sobrevivência dos resíduos calculados segundo o modelo de Cox, quanto mais próximo a função de sobrevivência dos resíduos se aproxima da exponencial, melhor é o ajuste.

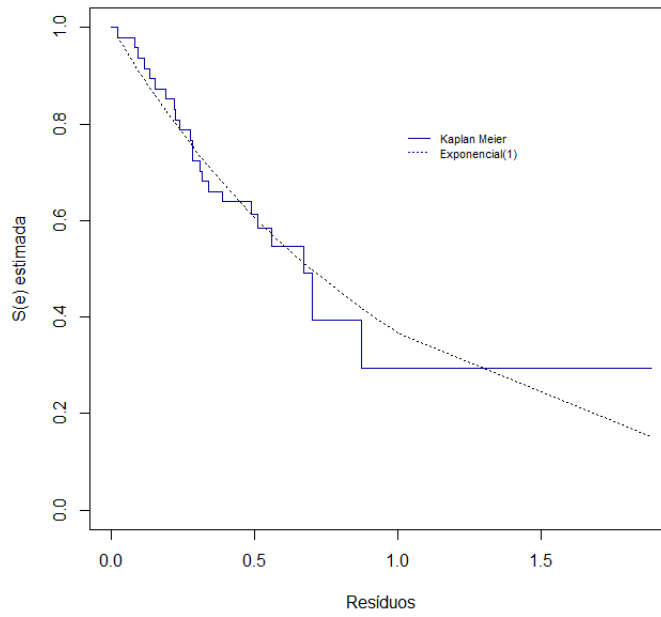


Figura 25: Resíduos de Cox-Snell do modelo *Charter Schools*.

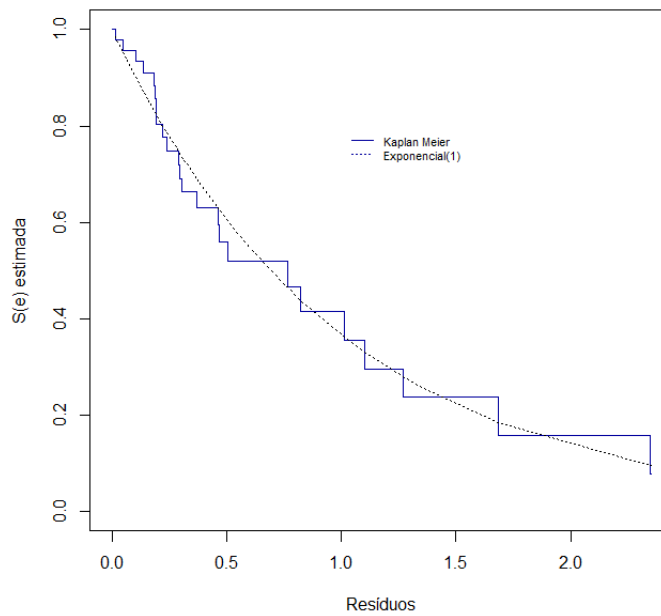


Figura 26: Resíduos de Cox-Snell do modelo *High School Exit Exams*.

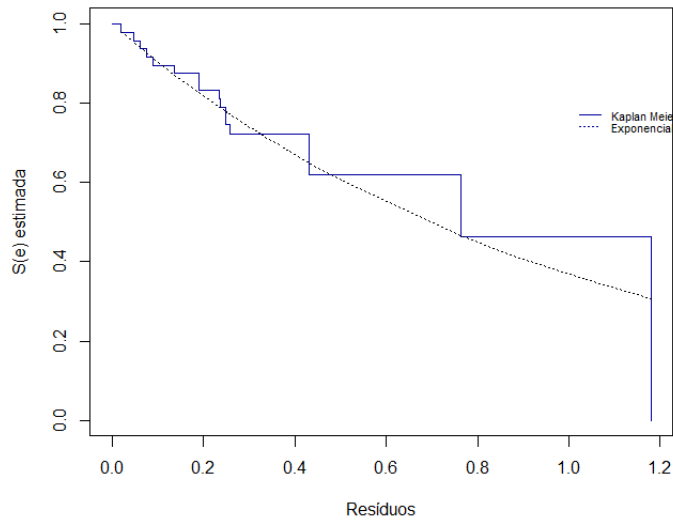


Figura 27: Resíduos de Cox-Snell do modelo *School Choice*.

Considerando as figuras acima o ajuste global dos modelos parecem adequados aos dados, visto que os resíduos seguem uma forma semelhante a exponencial padrão.

#### 6.4 Comparação da Medida de Inovação com o Score de Walker e a Taxa de adoção

Avaliada a adequação dos modelos finais para cada uma das políticas prossegue-se a análise para determinar a medida de inovação estadual proposta neste trabalho. Dessa forma, a medida será a soma dos riscos do estado falhar nos tempos selecionados em cada uma das três políticas, sendo o risco do estado falhar em cada política calculado por meio da seguinte expressão:

$$h(t|\mathbf{x}) = h_0(t) \exp(\mathbf{x}^T \boldsymbol{\beta}).$$

Na tabela a seguir encontram-se os resultados da medida de inovação proposta no tempo 1 e no tempo 6 e os Scores de Walker e *Adoption Rate* para cada um dos 48 estados americanos considerando as três políticas da área da educação em estudo. Ressalta-se que o risco no tempo 1 para cada estado é definido pela soma do estado falhar no tempo 1 em cada uma das três políticas em questão, o mesmo procedimento foi realizado para o tempo 6.

Tabela 8: Medida de Inovação, Score de Walker, Adoption Rate.

Estado	Risco (t=1)	Risco (t=6)	Score de Walker	Adoption Rate
Alabama	0,0845	0,2958	0,3275	0,0833
Arizona	0,0405	0,2775	0,0857	0,0833
Arkansas	0,0455	0,2409	0,1768	0,0833
California	0,2113	1,0804	0,2667	0,1304
Colorado	0,0469	0,2923	0,3333	0,0833
Connecticut	0,0482	0,2873	0,0000	0,0400
Delaware	0,0602	0,2934	0,0667	0,0400
Florida	0,1862	0,5934	0,3333	0,0833
Georgia	0,1488	0,4337	0,4029	0,0833
Idaho	0,0400	0,2346	0,4377	0,0400
Illinois	0,0596	0,3550	0,1333	0,0833
Indiana	0,0529	0,2556	0,1333	0,0400
Iowa	0,0601	0,3054	0,2000	0,0400
Kansas	0,0318	0,1981	0,1333	0,0400
Kentucky	0,1077	0,3479	0,0000	-
Louisiana	0,1013	0,3655	0,2551	0,0833
Maine	0,0447	0,2365	0,0000	-
Maryland	0,0951	0,3814	0,3188	0,0400
Massachusetts	0,0754	0,4630	0,3536	0,1304
Michigan	0,1296	0,8202	0,2667	0,0833
Minnesota	0,0507	0,2942	0,7681	0,1304
Mississippi	0,1631	0,4708	0,2464	0,0400
Missouri	0,0856	0,3697	0,0000	-
Montana	0,0303	0,1980	0,0000	-
Nebraska	0,0496	0,2863	0,2000	0,0400
Nevada	0,1280	0,4249	0,3188	0,0400
New Hampshire	0,0577	0,2323	0,0667	0,0400
New Jersey	0,0780	0,4492	0,6000	0,1304
New México	0,0466	0,2195	0,3884	0,0833
New York	0,1878	0,9263	0,3188	0,0400
North Carolina	0,2276	0,6024	0,3188	0,0833
North Dakota	0,0262	0,1838	0,0000	-
Ohio	0,1339	0,5559	0,3072	0,0833
Oklahoma	0,0434	0,2850	0,0000	-
Oregon	0,0380	0,2533	0,0667	0,0400
Pennsylvania	0,0631	0,3936	0,0000	-
Rhode Island	0,0353	0,2266	0,0667	0,0400
South Carolina	0,2533	0,6421	0,2174	0,0833
South Dakota	0,0285	0,1911	0,0000	-
Tennessee	0,0667	0,2869	0,2609	0,0400
Texas	0,1027	0,4131	0,2841	0,0833
Utah	0,0331	0,2005	0,0000	0,0833
Vermont	0,0480	0,2521	0,0000	-
Virginia	0,1106	0,3861	0,3043	0,0400
Washington	0,1063	0,4106	0,2203	0,0833
West Virginia	0,0351	0,2277	0,0000	-
Wisconsin	0,1110	0,6896	0,2290	0,6886
Wyoming	0,0312	0,1927	0,0667	0,0400

Na tabela acima foram destacados os 10 estados mais inovadores, ou seja, aqueles com os maiores valores para cada uma das medidas. Para o Score de Walker foram ressaltados 12 estados, pois Maryland, Nevada, New York e North Carolina apresentam



o mesmo valor para o coeficiente. O mesmo ocorre com a *Adoption Rate*, em que foram evidenciados 21 estados, uma vez que 17 dos estados possuem o mesmo valor para a taxa.

Ao comparar os 10 estados mais inovadores entre os risco no tempo 1 e no tempo 6, apenas 2 estados diferem entre eles. No tempo 1 os estados da Georgia e Nevada aparecem no ranking e no tempo 6 Massachusetts e Wisconsin. Os 8 estados em comum para essa medida proposta no tempos 1 e 6 são California, Florida, Michigan, Mississippi, New York, North Carolina, Ohio e South Carolina.

Constrastando o Score inicialmente proposto por Walker com o risco no tempo 1 proposto nesse trabalho, nota-se que o risco no tempo 1 inclui como estados inovadores California, Michigan, Mississippi, Ohio e South Carolina que não são considerados inovadores segundo o Score de Walker. Porém, os estados, Florida, Georgia, Nevada, New York e South Carolina são considerados inovadores segundo as duas medidas.

Realizando a mesma comparação para o risco no tempo 6, constata-se que este apresenta os 6 estados inovadores California, Michigan, Mississippi, Ohio, South Carolina e Wisconsin, que não são considerados segundo o Score de Walker. Os estados em comum entre estes scores são Florida, Massachusetts, New York e North Carolina.

Confrontando o risco no tempo 1 com a *Adoption Rate*, proposta por Boehmmke e Skinner (2012), evidencia-se que há 7 estados dos 10 compartilhados entre as taxas e 3 díspares. Sendo que a California, Florida, Georgia, Michigan, North Carolina, Ohio e South Carolina são os estados compartilhado e Mississippi, Nevada e New York os díspares.

Para o risco no tempo 6 com a *Adoption Rate*, verifica-se 8 estados inovadores em comum e 2 divergentes. Em que a California, Florida, Massachusetts, Michigan,, North Carolina, Ohio, South Carolina e Wisconsin são os estados inovadores em comum e Mississippi e New York os divergentes.

Desse modo, 45% em média dos estados considerados mais inovadores para o grau de inovação proposto coincidem com aqueles mais inovadores em conformidade com o Score de Walker. Para a *Adoption Rate* esse percentual médio é de 75%. Portanto, os resultados obtidos indicam que as conclusões da taxa de inovação poroposta coincide com os indicadores já existentes na literatura da ciência política.

## 7 Conclusão

À luz do campo de estudos da difusão de políticas públicas principiado pelo interesse em entender de que forma ocorre a difusão de políticas entre estados e sucedido pela busca em identificar os motivos da adoção de políticas públicas por grupo de estados, este trabalho buscou contribuir com avanços para este campo.

Nesse contexto, baseado no Score de Walker, proposto por Jack Walker (1969), e na *Adoption Rate*, apresentada por Boehmke e Skinner (2012), que visam medir o grau de inovação estadual fundamentado na adoção de políticas públicas, este estudo tentou em contemplar as limitações dessas medidas.

Dessa forma, a metodologia proposta emprega a Análise de Sobrevivência para integrar esta literatura ao possibilitar a avaliação simultânea de uma lista de fatores que influenciam o tempo e a adoção ou não das inovações políticas, compreende a informação de estados censurados e ainda contempla o Escore de risco global que se mostrou ser uma medida robusta para avaliar o nível de inovação estadual.

Embora este relatório permita analisar o fenômeno de falhas de políticas públicas em um espaço específico, os Estados Unidos, e apenas da área educacional, vale ressaltar que o modelo de Cox e o Escore de risco global propostos podem ser aplicados para políticas de várias áreas e em diversos contextos, inclusive para o cenário brasileiro.

Os resultados apresentados revelam que o Escore de risco global proposto demonstrou ser uma boa alternativa aos scores existentes na literatura, o Score de Walker e *Adoption Rate*. Por fim, este trabalho pretende servir de base e estímulo para que novos estudos implementando a metodologia apresentada sejam realizados por pesquisadores da área afim de que novas e brilhantes descobertas possam ser alcançadas.



## Referências

- BATISTA, M. A difusão da lei de acesso à informação nos municípios brasileiros: fatores internos e externos. Escola Nacional de Administração Pública (Enap), 2018.
- BATISTA, M.; DOMINGOS, A. Mais que boas intenções: técnicas quantitativas e qualitativas na avaliação de impacto de políticas públicas. *Revista Brasileira de Ciências Sociais*, SciELO Brasil, v. 32, n. 94, 2017.
- BERRY, F. S.; BERRY, W. D. State lottery adoptions as policy innovations: An event history analysis. *American political science review*, Cambridge University Press, v. 84, n. 2, p. 395–415, 1990.
- BOEHMKE, F. J.; SKINNER, P. State policy innovativeness revisited. *State Politics & Policy Quarterly*, Cambridge University Press, v. 12, n. 3, p. 303–329, 2012.
- CARVALHO, G. d. P. Difusão de programas e políticas públicas nos municípios brasileiros: fatores explicativos e implementação pró-forma em processos coercitivos conduzidos pela esfera federal. 2020.
- CAVALCANTI, A. C. R.; SOUZA, F. A. M. de. Um ensaio sobre a difusão de ideias na política pública. *Revista de Administração Educacional*, v. 1, n. 1, 2016.
- COÊLHO, D. B.; CAVALCANTE, P.; TURGEON, M. Mecanismos de difusão de políticas sociais no Brasil: uma análise do programa saúde da família. *Revista de Sociologia e Política*, SciELO Brasil, v. 24, p. 145–165, 2016.
- COLOSIMO, E. A.; GIOLO, S. R. *Análise de sobrevivência aplicada*. [S.l.]: Editora Blucher, 2006.
- COX, D. R.; OAKES, D. *Analysis of survival data*. [S.l.]: Chapman and Hall/CRC, 2018.
- DEARING, J. W.; COX, J. G. Diffusion of innovations theory, principles, and practice. *Health affairs*, v. 37, n. 2, p. 183–190, 2018.
- GRAHAM, E.; SHIPAN, C.; VOLDEN, C. *The diffusion of policy diffusion research (Unpublished)*. Ohio State University and University of Michigan. 2008.
- LINS, R.; FILHO, D. F.; ROCHA, E. Não temos tempo a perder: uma introdução à análise de sobrevivência. *Política Hoje*, v. 26, n. 1, p. 279–298, 2017.
- SCHMIDT, P.; WITTE, A. D. Predicting criminal recidivism using ‘split population’ survival time models. *Journal of Econometrics*, Elsevier, v. 40, n. 1, p. 141–159, 1989.
- WALKER, J. L. The diffusion of innovations among the American states. *American political science review*, Cambridge University Press, v. 63, n. 3, p. 880–899, 1969.