

UNIVERSIDADE FEDERAL DE VIÇOSA
INSTITUTO DE POLÍTICAS PÚBLICAS E DESENVOLVIMENTO SUSTENTÁVEL

**AVALIAÇÃO DOS IMPACTOS ECONÔMICOS E SOCIAIS DECORRENTES DA
APLICAÇÃO DOS RECURSOS DO FUNDO CONSTITUCIONAL DE
FINANCIAMENTO DO NORDESTE (FNE)**

**META 3 - PRODUTO 9: DESCRIÇÃO DAS ROTINAS E METODOLOGIAS DE
AVALIAÇÃO EMPREGADAS NO ESTUDO**

**Instituto de Políticas Públicas e
Desenvolvimento Sustentável**

Viçosa-MG
Agosto de 2023

SUMÁRIO

APRESENTAÇÃO	3
1. INTRODUÇÃO	4
2. PRODUTO 5 – AVALIAÇÃO DA EFICÁCIA DO FNE	5
2.1. ESTIMADOR DE DIFERENÇAS EM DIFERENÇAS	6
2.2. FUNÇÕES DOSE-RESPOSTA	14
2.3. REGRESSÃO QUANTÍLICA	19
2.4. FONTE E TRATAMENTO DOS DADOS	24
3. PRODUTO 6 – AVALIAÇÃO DA EFICIÊNCIA DO FNE	26
3.1. ÍNDICE DE MALMQUIST	31
3.2. TESTE DE CONVERGÊNCIA DAS MEDIDAS DE EFICIÊNCIA	33
3.3. MÉTODO DE DETECÇÃO DE OUTLIERS	35
3.4. TESTES NÃO PARAMÉTRICOS DE FRONTEIRAS DE EFICIÊNCIA	37
3.5. ANÁLISE DISCRIMINANTE	39
3.6. PROCEDIMENTOS E BASE DE DADOS	42
3.6.1. MODELO REGIONAL	42
3.6.2. MODELO SETORIAL	47
4. PRODUTO 7 – AVALIAÇÃO DA EFETIVIDADE DO FNE	50
4.1. ANÁLISE DESCRITIVA	51
4.2. ESTIMADOR DE DIFERENÇAS EM DIFERENÇAS	53
4.3. REGRESSÃO QUANTÍLICA	62
4.4. FONTE E TRATAMENTO DOS DADOS	67
5. PRODUTO 8 – AVALIAÇÃO DO RETORNO ECONÔMICO E SOCIAL DO FNE	69
5.1. ANÁLISE DE INSUMO-PRODUTO	69
5.2. MODELO DE EQUILÍBRIO GERAL COMPUTÁVEL	77
5.3. FONTE E TRATAMENTO DOS DADOS	80
5.4. ESTRATÉGIA DE AVALIAÇÃO	82
6. CONSIDERAÇÕES FINAIS	85
ANEXO 1. ROTINAS DE PROGRAMAÇÃO EMPREGADAS NO PRODUTO 5	86

ANEXO 2. ROTINAS DE PROGRAMAÇÃO EMPREGADAS NO PRODUTO 6	112
ANEXO 3. ROTINAS DE PROGRAMAÇÃO EMPREGADAS NO PRODUTO 7	120
ANEXO 4. ROTINAS DE PROGRAMAÇÃO EMPREGADAS NO PRODUTO 8	170



IPPDS
Instituto de Políticas Públicas e
Desenvolvimento Sustentável

APRESENTAÇÃO

O presente documento trata da entrega formal de documento técnico (Meta 3 - Produto 9: Descrição das Metodologias de Avaliação), relativo à execução do projeto intitulado “*Avaliação dos impactos econômicos e sociais decorrentes da aplicação dos recursos do Fundo Constitucional de Financiamento do Nordeste (FNE) - 1º ano do ciclo de avaliações*”, em conformidade com a Portaria Interministerial Nº 7/2020 e prioridades do programa 2217 - Desenvolvimento Regional, Territorial e Urbano. O projeto tem âmbito estendido e abarca vários níveis analíticos de avaliação do FNE, tendo por objetivo geral, segundo descrito em seu Plano de Trabalho (SEI SUDENE 0303722, pág. 1), “*avaliar a eficiência, eficácia, efetividade e retorno econômico e social da aplicação dos recursos do FNE*”.

Seguindo o estabelecido pelo Plano de Trabalho, o documento técnico possui caráter descritivo, fundamentando-se na estruturação detalhada da descrição das metodologias de avaliação e das rotinas de programação empregadas na avaliação de impacto conduzida no contexto do projeto. Considerando essa base indicativa, a equipe técnica buscou detalhar as metodologias utilizadas para o desenvolvimento dos relatórios técnicos referentes à *Meta 2: Avaliação de eficácia, eficiência, efetividade e retorno econômico do FNE*¹.

Estruturalmente, este documento organiza-se em cinco seções. Além dessa breve apresentação, a introdução sumariza os métodos aplicados para a avaliação econômica e social do FNE. A ela se segue a descrição pormenorizada das ferramentas estatísticas e econométricas utilizadas em cada um dos quatro produtos considerados: (i) Avaliação da Eficácia; (ii) Avaliação da Eficiência; (iii) Avaliação da Efetividade; e (iv) Avaliação do Retorno Econômico e Social. Por fim, apresentam-se as considerações finais.

¹ Em específico, foram considerados os Produtos 5, 6, 7 e 8.

1. INTRODUÇÃO

Conforme estabelecido no Plano de Trabalho, esse relatório versa sobre a estruturação detalhada da descrição das metodologias de avaliação empregadas na avaliação de impacto conduzida no contexto da Meta 2 do Projeto. Em específico, foram pormenorizadas as ferramentas metodológicas empregadas nos seguintes produtos: (5) Avaliação da Eficácia do FNE; (6) Avaliação da Eficiência do FNE; (7) Avaliação da Efetividade do FNE; e (8) Avaliação do Retorno Econômico e Social do FNE.

Reconhecendo que os produtos possuem escopos de análise diferentes e que, por esse motivo, diferentes metodologias foram empregadas, optou-se por apresentar os métodos utilizados em subseções separadas para cada produto. Além do detalhamento técnico dos modelos teóricos empregados, foram destacados os procedimentos utilizados para ajuste de dados e o nível de agregação dos dados para a condução da análise.

De modo específico, tem-se que as análises conduzidas nos Produtos 5 e 7 tiveram por base a utilização de técnicas de Avaliação de Impacto, com destaque para o estimador de Diferenças em Diferenças. Para o Produto 6, por seu turno, a análise foi pautada no emprego de abordagens não-paramétricas baseadas em programação matemática. Já para o Produto 8, foram empregados modelos de análise de equilíbrio geral no intuito de avaliar os efeitos econômicos e sociais do FNE.

Além da descrição pormenorizada das ferramentas metodológicas empregadas na avaliação de eficácia, eficiência, efetividade, e retorno econômico e social do FNE, também são apresentadas, nos Anexos, as rotinas de programação utilizadas para a condução das análises empíricas. Destaca-se ainda que abordagens diferentes foram empregadas a depender do tipo de avaliação conduzida, de modo que diferentes softwares tenham sido utilizados durante a condução do estudo.

2. PRODUTO 5 – AVALIAÇÃO DA EFICÁCIA DO FNE

Essa seção descreve o conjunto de metodologias utilizado para conduzir a avaliação da eficácia do acesso ao crédito concedido via FNE às empresas localizadas na área de atuação da Sudene, no período entre 2000 e 2018. Para avaliar a eficácia do acesso ao crédito (em nível de empresas) sobre faturamento, emprego, massa salarial e salário médio, propõe-se a utilização do procedimento de Diferenças em Diferenças (DD), o qual apresenta controle para heterogeneidade não observável.

O desenho tradicional do modelo DD inclui dois períodos de tempo onde nenhuma observação é tratada no primeiro período e parte delas é tratada no segundo período. No entanto, dada a disponibilidade de dados, considerou-se múltiplos períodos de tempo (2000 a 2018) e variação da designação do tratamento no tempo, ou seja, as empresas são expostas ao tratamento em momentos diferentes. O termo *staggered design* ou *staggered rollout design* é utilizado para descrever esse tipo de situação, onde o tratamento é recebido em períodos diferentes e uma vez exposta ao tratamento, a empresa passa ao estado tratado de forma permanente.

Atualmente, muitos estudos (e.g., Borusyak e Jaravel, 2018; De Chaisemartin e D'Haultfoeuille, 2020; Sun e Abraham, 2021; Goodman-Bacon, 2021; Athey e Imbens, 2022) têm levantado preocupações em relação à dinâmica e à heterogeneidade dos efeitos de tratamento nesse tipo de análise, que, caso presentes, levam a estimativas viesadas do efeito de tratamento. Para evitar tais problemas na análise da eficácia do FNE, foram utilizadas as contribuições de Wooldridge (2021), visando análise de média condicional, compreendendo todo o período analisado.

Ainda quanto à avaliação da eficácia do montante de crédito concedido às empresas sobre faturamento, emprego, massa salarial e salário médio, utilizou-se o arcabouço das funções dose-resposta, estimadas por Mínimos Quadrados Ordinários (MQO), levando em consideração problemas que causam vieses nas estimativas, como o viés de auto-seleção, má especificação da forma funcional e o fato de o tratamento ser contínuo e não-binário.

Além da análise de média condicional, é interessante também observar os efeitos da política de crédito do FNE ao longo da distribuição dos indicadores das empresas. Neste

sentido, a metodologia empregada por Carvalho e Firpo (2014), para analisar o impacto de determinadas políticas públicas sobre o desempenho escolar, considera a heterogeneidade das empresas e controla bem os possíveis problemas causados pela existência de variáveis não observadas que teriam o poder de afetar os resultados de eficácia da política desenvolvida pelo FNE.

Nesse contexto, optou-se por empregar o método de Regressão Quantílica, aplicado em um conjunto de dados de painel com efeitos fixos. Tal como desenvolvido por Machado e Santos Silva (2019), o procedimento contém informações das empresas, compreendendo períodos entre os anos de 2000 a 2018, desenvolvendo análise de DD e avaliando a eficácia da política nos diversos quantis de nível de atividade econômica.

Essa abordagem se torna relevante quando considerado o fato de que os efeitos da política de concessão de crédito do FNE podem ser heterogêneos nos diferentes quantis de distribuição das variáveis de resultado das empresas. Os resultados podem indicar se há diferenças na eficácia da política ao direcionar recursos para empresas no quantil inferior dos indicadores utilizados comparada com os do quantil superior, por exemplo.

2.1. ESTIMADOR DE DIFERENÇAS EM DIFERENÇAS

A avaliação de eficácia visa mensurar o chamado efeito de tratamento associado a uma intervenção ou programa/política pública sobre algum produto ou indicador socioeconômico de interesse. Geralmente, o tratamento é aplicado a algum grupo de interesse, que passa a ser chamado de grupo de tratamento, ao passo que o grupo que não recebe tratamento é chamado de grupo de controle.

É de interesse saber a eficácia dessa intervenção sobre o grupo de tratamento em relação ao grupo de controle. No presente estudo, a intervenção de interesse é o FNE, o qual objetiva relaxar a restrição de crédito das empresas assistidas e, com isso, obter ganhos de produtividade e geração de emprego e renda, desencadeando assim o crescimento e o desenvolvimento da economia regional. Nesse contexto, o grupo de tratamento é composto pelas firmas que receberam recursos provenientes do FNE e o grupo de controle pelas que não receberam².

² Por uma limitação de acesso aos dados, o grupo tratado é composto exclusivamente de beneficiários do FNE. Como este programa tem características e objetivos específicos, a interpretação é de comparação do impacto do

Quando se aborda a análise de impacto dos fundos constitucionais de financiamento, em nível de firmas, a literatura tradicionalmente busca identificar eficácia do FNE sobre o estoque de emprego, massa salarial e salário médio (e.g., Lima et al., 2013; Silva, Resende e Silveira Neto, 2006; Soares, Sousa e Pereira Neto, 2009; Sousa, Nottingham e Gonçalves, 2010; Gondim, Sousa e Pires, 2011; Resende, 2012, 2014c; Soares et al., 2013; Oliveira et al., 2017, 2018; Daniel e Braga, 2020).

O objetivo da análise de eficácia, proposta no presente estudo, é mensurar a diferença existente nos resultados (*outcomes*) entre os grupos de tratamento e de controle para verificar o efeito da intervenção. Os modos mais comuns de se aferir o efeito de tratamento são o efeito médio de tratamento (*Average Treatment Effect – ATE*) e o efeito médio de tratamento sobre os tratados (*Average Treatment Effect on the Treated – ATT*).

O efeito médio de tratamento (ATE) pode ser definido como a diferença média existente nos resultados entre os grupos tratado e de controle:

$$\delta_{ATE} = E(y_1 - y_0) = E(y_1|X, D = 1) - E(y_0|X, D = 0) \quad (1)$$

em que $E(\cdot)$ representa o operador de valor esperado ou a média; y_1 e y_0 representam as variáveis de interesse, que no caso das empresas são: estoque de emprego, massa salarial e salário médio, para os grupos de tratamento e de controle, respectivamente; D identifica a designação de tratamento, sendo 1 para os tratados e 0 para os não-tratados; X é o vetor de características das empresas, conforme definido adiante.

Segundo Khandker, Koolwal e Samad (2010), o problema do ATE é que os grupos de tratamento e de controle podem não ter a mesma prioridade de intervenção. Assim, a diferença média observada entre esses grupos pode não ser totalmente devida à intervenção sob investigação. O problema ocorre devido à dificuldade de se definir o contrafactual (o que teria ocorrido com os indivíduos tratados, caso não tivessem recebido o tratamento).

Nos estudos observacionais, ao contrário dos experimentos controlados, a participação no tratamento não é aleatória devido aos propósitos da intervenção e à questão da auto-seleção. Ao estimar o ATE via regressão linear, o fato de características observáveis e não observáveis

FNE em relação às empresas que não receberam benefícios do programa mesmo que tenham recebido de outras fontes.

estarem relacionadas à participação na intervenção ($D = 1$) introduzem correlação entre o termo de erro e a variável de tratamento (D). Assim, viola-se um dos pressupostos do Modelo Clássico de Regressão Linear (MCRL), causando viés e inconsistência nas estimativas do efeito médio de tratamento obtidas via MQO.

Para contornar esse problema, estudos observacionais geralmente mensuram o efeito de tratamento via ATT, o qual é definido como a diferença média da variável de resultado entre o grupo de tratamento e esse mesmo grupo caso não houvesse a intervenção:

$$\delta_{ATT} = E(y_1 - y_0 | D = 1) = E(y_1 | X, D = 1) - E(y_0 | X, D = 1) \quad (2)$$

em que o último termo à direita representa o contrafactual, o qual é comumente estimado via pareamento por escore de propensão (*Propensity Score Matching* – PSM).

No PSM, o contrafactual é construído com base no grupo que não participou da intervenção a partir do pareamento com aqueles que participaram. O procedimento consiste na estimação de um modelo de probabilidade (logit/probit) voltado para as chances de participar da intervenção em função de características observáveis. Obtém-se, assim, o escore de propensão (probabilidade ser exposto à intervenção) e, por meio de métodos específicos (e.g., kernel, vizinho mais próximo, estratificação, distância de Mahalanobis), ocorre o pareamento de modo a se obter o contrafactual.

A equação (3) define o ATT:

$$\delta_{ATT} = E(y_1 - y_0 | p(X), D = 1) = E(y_1 | p(X), D = 1) - E(y_0 | p(X), D = 0) \quad (3)$$

em que $p(X) = \text{prob}(D = 1 | X) = E(D | X)$ é a probabilidade ou escore de propensão da empresa receber recursos do FNE, dadas suas características.

Apesar de mais adequado que o ATE para uso em estudos observacionais, o ATT mensurado via PSM ignora a existência de heterogeneidade não observável, possivelmente levando ao viés de seleção. Dada a divergência de resultados apresentados na literatura quanto à presença de restrições de crédito, essa hipótese parece ser suficientemente forte para se sugerir uma abordagem mais adequada, como o método de Diferenças em Diferenças (DD) para dados longitudinais.

O método DD assume que a heterogeneidade que afeta a participação das firmas no FNE está presente e é invariante no tempo. Nesse caso, o ATT é estimado a partir das diferenças nas variáveis de resultado entre os grupos de tratamento e de controle, antes e depois da intervenção. O desenho tradicional do modelo DD inclui dois períodos de tempo ($t = 1, 2$). No primeiro período ($t = 1$), nenhuma empresa é tratada, enquanto no segundo período ($t = 2$), parte delas recebe o tratamento. Ao fazer comparações entre as empresas que tratadas e não tratadas, antes e depois da intervenção, o método DD produz uma estimativa mais consistente do contrafactual.

O ATT estimado pelo método DD é definido na equação (4):

$$\delta_{ATT} = E[(y_{1,t=2} - y_{1,t=1}) - (y_{0,t=2} - y_{0,t=1})|X, D = 1] \quad (4)$$

em que o primeiro termo à direita se refere à diferença nos resultados antes ($y_{1,t=1}$) e depois ($y_{1,t=2}$) da intervenção para o grupo de tratamento. Este termo pode estar viesado, devido a tendências no tempo como inflação, crescimento econômico, dentre outras. O segundo termo à direita se refere à diferença nos resultados antes ($y_{0,t=1}$) e depois ($y_{0,t=2}$) da intervenção para o grupo de controle. Esse último termo é inserido para remover a possibilidade de vies.

De forma geral, Wooldridge (2010), Callaway e Sant'Anna (2021) e Roth et al. (2022) sugerem que, para identificar o parâmetro de interesse δ_{ATT} , são necessárias duas hipóteses básicas: (i) a suposição de tendências paralelas (*parallel trends*); e (ii) a suposição de ausência de efeitos antecipatórios ao tratamento (*no anticipatory effect*).

Uma maior robustez associada ao método DD provém da utilização de uma estrutura de dados em painel. No entanto, de acordo com Borusyak e Jaravel (2018), De Chaisemartin e D'Haultfoeuille (2020), Sun e Abraham (2021), Callaway e Sant'Anna (2021), Goodman-Bacon (2021) e Athey e Imbens (2022), a utilização do modelo de efeitos fixos não é adequada para a estimação do ATT quando existem múltiplos períodos de tempo e quando a designação do tratamento varia no tempo, o que é observado na presente análise, pois as empresas são expostas à intervenção em diferentes momentos. O termo *staggered design* ou *staggered rollout design* é utilizado para descrever a situação em que o tratamento é recebido em períodos diferentes e uma vez exposto ao tratamento, o empreendimento passa ao estado tratado de forma permanente.

Conforme Wooldridge (2021), com vários períodos de tempo ($t = 1, 2 \dots T$) e diferentes momentos de tratamento ($G = \min\{t: D_t = 1\}$), existem diversos níveis de exposição ao tratamento. Esse fato causa dinâmica e heterogeneidade nos efeitos de tratamento, tornando o modelo de efeitos fixos inadequado, uma vez que inclui apenas um parâmetro para medir o ATT. Callaway e Sant’Anna (2021) definem os níveis de tratamento pelo período de entrada no estado tratado. Portanto, ($G = g$) define o grupo de tratados exposto ao tratamento a partir do período g , indicando o primeiro momento em que o grupo G foi exposto ao tratamento.

Caso a empresa nunca seja tratada, a notação adotada para esse grupo é ($G = \infty$). Nesse caso, surgem dois tipos de grupos de controle: os que ainda não foram tratados e os que nunca serão tratados. Dessa forma, o ATT pode ser definido no ano t para o grupo exposto ao tratamento em g :

$$\delta_{ATT}(g, t) = E[(y_{1,t} - y_{1,t=g-1} | X, G = g) - (y_{0,t} - y_{0,t=g-1} | X, G \in \mathcal{G})] \quad (5)$$

em que o primeiro termo à direita se refere à diferença nos resultados do grupo de tratamento no período (g), antes de ser tratado, ($y_{1,t=g-1}$) e no período de interesse (t) em que se quer estimar o ATT após o tratamento ($y_{1,t}$). O segundo termo à direita usa diferenças nos resultados do grupo de controle ($y_{0,t} - y_{0,t=g-1}$) para os mesmos períodos. O grupo de controle ($G \in \mathcal{G}$) pode ser formado por empresas que nunca serão tratadas ($G = \infty$) ou por empresas que ainda não foram tratadas ($G = g' > t$), como em Callaway e Sant’Anna (2021), ou ainda por ambos, como em Wooldridge (2021).

De Chaisemartin e D’Haultfoeuille (2020) e Callaway e Sant’Anna (2021) mostram que é comum na literatura a utilização de alguma variação do modelo de efeitos fixos, também conhecido como *Two-way Fixed Effects* (TWFE), para estimar o parâmetro δ_{ATT} do modelo DD, com vários períodos de tempo e variação no tempo de tratamento. Dos artigos empíricos publicados na *American Economic Review* (AER) entre 2010 e 2012, 20% utilizavam esse tipo de regressão (DE CHAISEMARTIN E D’HAULTFOEUILLE, 2020).

O problema estaria relacionado à dinâmica e à heterogeneidade dos efeitos de tratamento nesse tipo de análise, que, caso presentes, levariam a estimações viesadas dos efeitos promovidos pela intervenção. Goodman-Bacon (2021) e De Chaisemartin e D’Haultfoeuille

(2020) demonstram que a utilização do TWFE para estimar o ATT em modelos DD com múltiplos períodos e diferentes momentos de exposição ao tratamento resulta em um parâmetro que é a soma ponderada de todos os recortes de DD possíveis com dois períodos de tempo. A interpretação dos parâmetros estimados se torna complexa e os pesos podem assumir valores negativos, o que no limite, pode resultar na estimação de um impacto negativo ou não significativo, mesmo que o verdadeiro ATT seja positivo e estatisticamente significativo.

Segundo Wooldridge (2021), a conclusão de que o estimador TWFE gera resultados viesados na presença de efeitos de tratamentos dinâmicos e heterogêneos é o resultado da especificação utilizada e não do estimador em si. Para solucionar o problema, propõe-se uma especificação flexível do estimador de efeitos fixos que permite que o ATT varie de acordo com o grupo de tratamento (g), com o período (t) e também com as variáveis de controle (X). A equação (6) apresenta tal especificação:

$$y_{it} = \alpha_i + \sum_{s=q}^T \phi_s fs_t + \sum_{s=q}^T (fs_t X_i) \pi_s + \sum_{g=q}^T \sum_{s=g}^T \delta_{gs} (D_{it} d_{ig} fs_t) + \sum_{g=q}^T \sum_{s=g}^T (D_{it} d_{ig} fs_t \dot{X}_{ig}) \rho_{gs} + \varepsilon_{it} \quad (6)$$

em que α são os efeitos fixos para as empresas, ϕ são os efeitos fixos para o tempo, fs são variáveis binárias que indicam os anos a partir do primeiro período de tratamento ($s = q > 1$), supõe-se que haja pelo menos um período em que nenhuma empresa é tratada. O vetor de parâmetros π capta a tendência das variáveis contextuais X_i , invariantes no tempo, sobre o resultado de interesse ao longo dos anos do programa. O principal vetor de parâmetros, δ_{gs} , capta os efeitos de tratamento da intervenção sobre as empresas, o ATT varia por grupo de tratamento (g) e ao longo do tempo de tratamento (s), onde D assume valor 1, a partir do momento em que a empresa recebe financiamento do FNE e (d) identifica o grupo de tratamento (g) ao qual a empresa faz parte. Já o vetor de parâmetros ρ mostra o efeito das variáveis de controle sobre o efeito de tratamento (ATT), sendo $\dot{X}_{ig} = X_i - \bar{X}_g$, ou seja, o desvio com relação à média do grupo de tratamento (g). Por fim, ε é o termo de erro da regressão.

Segundo Bertrand et al. (2004) e Imbens e Wooldridge (2009), as estimações feitas pelo procedimento DD com efeitos fixos são tipicamente afetadas por problemas de autocorrelação, devido à natureza das variáveis dependentes, que, juntamente com a heterocedasticidade, tornam as estimativas dos erros-padrão inválidas. Nesse sentido, Cameron e Trivedi (2005 e 2009) sugerem que, especialmente para painéis curtos, seja feita a correção robusta por *cluster*, de modo que a verdadeira matriz de variância e covariância dos coeficientes seja estimada, corrigindo assim os problemas.

A solução apresentada por Wooldridge (2021) altera apenas a especificação do modelo TWFE que estima o ATT por meio do desenho de Diferenças em Diferenças. No entanto, vários estimadores alternativos que permitem maior dinâmica e heterogeneidade do ATT foram propostos recentemente (e.g., De Chaisemartin e D'Haultfoeuille, 2020; Sun e Abraham, 2021; Callaway e Sant'Anna, 2021; Athey e Imbens, 2022).

Dentre os estimadores propostos, o de Callaway e Sant'Anna (2021) possui maior aplicação para a avaliação dos impactos do FNE em relação aos demais, pois permite a inclusão de variáveis de controle e possui condições de identificação mais flexíveis. Callaway e Sant'Anna (2021) propõe um estimador (CS)³ que combina as técnicas de ponderação pelo inverso da probabilidade (*Inverse Probability Weighting* – IPW) com o ajuste de regressão.

Atualmente, a estimação dos efeitos de tratamento ATT pelo método DD são motivo de significativa discussão na literatura. Vários estimadores foram propostos para superar as limitações levantadas em relação à regressão de efeitos fixos, principalmente relacionadas à possível dinâmica e heterogeneidade do ATT, em situações de múltiplos períodos de tempo e variação no tempo de tratamento. Segundo Woodridge (2021), as limitações levantadas são originadas da utilização de especificação inadequada, afirmando ainda que a regressão de efeitos fixos possui propriedades bem conhecidas em painéis desbalanceados e é eficiente em relação ao estimador CS, pois usa toda a informação disponível – grupos de controle nunca tratados e ainda não tratados – para a estimação do ATT. Optou-se, nesse caso, por utilizar as contribuições de Wooldridge (2021) para estimar os efeitos de tratamento do FNE.

Um problema inerente do modelo DD se refere à pressuposição de tendência comum entre os grupos de controle e de tratamento. Esta pressuposição assegura que não seja necessária a existência de suporte comum da propensão de participação do tratamento para

³ Em referência aos autores Callaway e Sant'Anna (2021).

estimação do impacto de interesse. No entanto, ela exige que as tendências dos resultados dos grupos analisados sejam iguais caso não houvesse tratamento.

A pressuposição de tendência comum garante que o impacto encontrado seja realmente devido à exposição à intervenção (GERTLER et al., 2011). Não é possível verificar formalmente tal pressuposição, porém, estudos sugerem a realização de testes de robustez (Cunningham, 2021; Huntington-Klein, 2022). Um desses testes de robustez é a estimação de um *event study*, o qual inclui modelos onde o tratamento ocorreria em outros pontos do tempo e não quando realmente ocorre o tratamento, ou seja, incluem-se avanços (*leads*) e defasagens (*lags*) nos modelos para verificar a existência de causalidade entre o momento de exposição à intervenção e a resposta em termos de impacto.

O modelo analítico utilizado é definido pela Equação (6), que define os determinantes do resultado y na firma i no ano t , podendo ser avaliada com base no seguinte modelo econométrico:

$$\ln(y_{it}) = \alpha_i + \sum_{s=q}^T \phi_s fs_t + \sum_{s=q}^T (fs_t X_i) \pi_s + \sum_{g=q}^T \sum_{s=g}^T \delta_{gs} (D_{it} d_{ig} fs_t) + \sum_{g=q}^T \sum_{s=g}^T (D_{it} d_{ig} fs_t X_{ig}) \rho_{gs} + \varepsilon_{it} \quad (7)$$

A variável y_{it} representa os resultados da firma i , no ano t , aos quais tem-se interesse em verificar a eficácia do programa de concessão de crédito vinculado ao FNE, quais sejam: estoque de emprego (capta a geração de emprego), massa salarial (capta a geração de renda) e salário médio (capta a produtividade média do trabalho). Utiliza-se a transformação logarítmica na variável dependente devido a seus bons atributos econométricos, como a compressão das unidades de dispersão com relação à média, maior propensão à normalidade e interpretação de semi-elasticidade. O período t compreende os anos entre 2000 e 2018.

Os parâmetros ϕ_s e α_i representam os efeitos fixos invariantes no tempo para as firmas e os efeitos fixos invariantes nas firmas para os anos, respectivamente, que controlam a heterogeneidade não observada no tempo e entre firmas. O vetor δ_{gs} representa o efeito que a participação no programa D exerce sobre o resultado y . A participação no programa D é

dividida de acordo com o programa e a finalidade do crédito (capital de giro, custeio ou investimento), sendo variáveis dicotômicas. O efeito de tratamento pode ser calculado por grupo de tratamento e ao longo do tempo, permitindo verificar sua dinâmica.

Os vetores π e ρ representam o efeito que cada uma das variáveis observadas X exerce sobre o resultado y e sobre o impacto do tratamento δ ao longo do tempo. Finalmente, ε é o termo de erro da regressão (Equação 7). Para calcular o ATT dos modelos DD, serão utilizados os estimadores de mínimos quadrados ponderados com efeitos fixos e correção robusta por *cluster* para heterocedasticidade e autocorrelação.

2.2. FUNÇÕES DOSE-RESPOSTA

O interesse em avaliar a eficácia promovida por políticas de crédito subsidiado surge do argumento de que a concessão deste reduz ou elimina as restrições de liquidez das empresas, causando assim maior acúmulo de capital, maior produção e aumento da produtividade, os quais, por sua vez, gera empregos, renda e aumento do salário médio.

Da mesma forma, o volume de crédito recebido pode resultar no relaxamento ou eliminação das restrições de liquidez das empresas, promovendo efeitos divergentes sobre faturamento, geração de emprego e renda, e aumento da produtividade. Surge, assim, o conceito de resposta à dose de tratamento. Se for esse o caso, as políticas de crédito podem ser desenhadas de forma mais precisa, ou seja, com objetivo de alcançar a maior eficácia possível. O interesse não reside no efeito médio de tratamento, mas no efeito médio para cada nível de tratamento.

Diferentemente do método de Pareamento por Escore de Propensão (PSM) proposto por Rosenbaum e Rubin (1983), no qual o contrafactual é estimado com base em tratamento binário, a estimação das funções dose-resposta é feita a partir do Escore de Propensão Generalizado (GPS) proposto por Imbens (2000), para o caso no qual o tratamento é categórico ou possui vários níveis, e por Hirano e Imbens (2004), para o caso onde o tratamento é contínuo. Da mesma forma que o PSM, o objetivo do GPS é remover o viés provocado pelo problema de auto seleção ao tratamento, provocado pelo fato de que características observáveis e não observáveis estão relacionadas com o montante de crédito recebido pelas empresas.

Para evitar o problema, utiliza-se o GPS para estimar as funções de resposta à determinada dose do tratamento. Para o caso do FNE, utilizam-se as informações referentes às empresas que receberam crédito para estimar as funções de, geração de emprego, renda e salário médio em resposta aos diferentes volumes de crédito recebidos. Cada nível de crédito recebido tem como contrafactual os demais níveis recebidos por outras firmas com as mesmas características. Como o interesse reside na estimação da resposta média a exposição a determinada dose de tratamento, nesse caso o montante de crédito, apenas as observações tratadas devem compor a amostra.

De acordo com Imbens (2000), Hirano e Imbens (2004) e Bia e Mattei (2007 e 2008), a média da função dose-resposta que permite verificar se quantidades diferentes de tratamento geram efeitos com intensidades divergentes pode ser definida como:

$$\mu(c) = E\{y_i(c)\} \quad (8)$$

em que, na equação (8), $\mu(c)$ é a média do produto y para o conjunto de empresas i que receberam o montante de crédito c , $\{y_i(c)\}_{c \in C}$ é a função dose-resposta que determina a resposta dos produtos y_i (variações no estoque de emprego, massa salarial e salário médio) em função do volume de crédito recebido ($c \in C_i$) e $i = 1, \dots, N$ representa o número de firmas na amostra, as quais possuem características observáveis captadas pelo vetor X_i .

Segundo Hirano e Imbens (2004), a função de propensão, que descreve a probabilidade r de uma empresa com vetor de características X receber o nível de tratamento c : $r(c, x)$, é então definida como a densidade do nível de crédito dadas as características dos empreendimentos:

$$r(c, x) = f_{C|X}(c|x) \quad (9)$$

em que, na equação (9), $C = c$ é a variável nível de crédito, $X = x$ é o conjunto de características dos empreendimentos e $f_{C|X}(c|x)$ é a função densidade do nível de crédito. Assim, o GPS pode ser definido como: $R = r(C, X)$, reunindo os escores de propensão do conjunto de níveis de crédito e de características das empresas. Deve ser atendida a propriedade de balanceamento da mesma forma que o PSM, ou seja, para grupos de empresas que possuam o mesmo escore de propensão $r(c, x)$, a probabilidade de receber determinado nível de crédito $C = c$ não

depende de suas características (X) o que permitiria associar a diferença de efeito ao nível de crédito.

Também é possível mostrar, segundo Hirano e Imbens (2004), que a utilização do GPS remove o viés associado à heterogeneidade observável, em outras palavras, a utilização do GPS permite que a diferença de resultados possa ser relacionado ao tratamento, nesse caso ao crédito:

$$\omega(c, r) = E\{y(c)|r(c, X) = r\} = E(Y|C = c, R = r) \quad (10)$$

$$\mu(c) = E[\omega\{c, r(c, X)\}] \quad (11)$$

em que, nas equações (10) e (11), $\omega(c, r)$ é a função do valor esperado dos produtos (y) condicionado aos valores do crédito (c) e do escore de propensão (r). Maiores detalhes sobre o GPS podem ser encontrados em Imbens (2000), Hirano e Imbens (2004) e Bia e Mattei (2007 e 2008).

O procedimento de estimação das funções dose-resposta consiste em três etapas: a estimação do GPS $r(c, x)$, que posteriormente é utilizada para a estimação da função do valor esperado dos produtos (y) condicionados aos valores do crédito (c) e do escore de propensão (r): $\omega(c, r) = E(Y|C = c, R = r)$ e, por fim, o cálculo da média $\mu(c) = E[\omega\{c, r(c, X)\}]$ que fornece a informação dos efeitos de cada nível de tratamento sobre as variáveis de interesse.

A primeira etapa para a obtenção das funções dose-resposta é a estimação do GPS que demanda a estimação da distribuição condicional do montante de crédito c em função das características das firmas X_i , que deve distribuir-se normalmente segundo Bia e Mattei (2008):

$$g(C_i)|X_i \sim N\{h(\gamma, X_i), \sigma^2\} \quad (12)$$

em que, $g(C_i)$ representa a variável de tratamento (montante de crédito recebido) ou alguma transformação que atenda ao pressuposto de normalidade, $h(\gamma, X_i)$ é uma função linear das variáveis X_i que dependem dos parâmetros γ e σ^2 é a variância da distribuição. Os parâmetros γ e σ^2 são estimados por máxima verossimilhança e, caso a hipótese de normalidade seja validada, o GPS é estimado pela seguinte função:

$$\hat{R}_i = \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} \exp \left[-\frac{1}{2\hat{\sigma}^2} \{g(C_i) - h(\hat{Y}, X_i)\} \right] \quad (13)$$

Após estimado o GPS, verifica-se a validade da hipótese de balanceamento entre grupos de propensão. Nesse caso, um teste t de diferenças de médias é utilizado para avaliar se os grupos de propensão possuem características semelhantes. Caso a propriedade de balanceamento seja atendida, a segunda fase da estimação das funções dose-resposta é a estimação do valor esperado:

$$\omega(c, r) = E(Y|C = c, R = r) = \beta_0 + \beta_1 C_i + \beta_2 C_i^2 + \beta_3 \hat{R}_i + \beta_4 \hat{R}_i^2 + \beta_5 C_i \hat{R}_i \quad (14)$$

em que, $\omega(c, r)$ é a função do valor esperado dos produtos (Y) condicionado aos valores do crédito ($C = c$) e do escore de propensão ($R = r$), C_i é o valor de crédito recebido pela empresa i , \hat{R}_i é o valor estimado do GPS para a empresa i , $C_i \hat{R}_i$ é a interação entre crédito e GPS e os β 's são os parâmetros a estimar. Hirano e Imbens (2004) alertam que os parâmetros β não possuem interpretação convencional ou com significado direto para a análise. Contudo, a significância estatística pode ser interpretada como evidência de influência causada pela variável.

Por fim, para se obter resultados que tenham interpretação no sentido de eficácia econômica, calcula-se a média da função dose-resposta, apresentada na equação (15):

$$E\{\hat{Y}(c)\} = \frac{1}{N} \sum_{i=1}^N \hat{\omega} \{c, \hat{r}(c, X_i)\} \quad (15)$$

A partir da média da função dose-resposta, para cada nível de crédito (c), é possível observar o resultado (\hat{Y}) em termos de variação da geração de empregos e variações da massa salarial e do salário médio. Para determinado nível de crédito (c), resultados positivos (negativos) indicam que esse montante de crédito gera impactos positivos (negativos), já a magnitude do impacto é comparável em relação aos demais níveis de crédito considerados.

Para verificar a eficácia do volume de crédito recebido pela firma sobre a variação do estoque de emprego, massa salarial e salário médio (y 's), a estimação do GPS considera as características observáveis X 's.

Essas variáveis são normalmente encontradas nos trabalhos em nível de firmas, pela disponibilidade e relação com o fenômeno econômico estudado. Dentre os estudos que utilizam as variáveis ou algumas delas pode se citar Silva, Resende e Silveira Neto (2006), Almeida Jr., Resende e Silva (2007), Silva, Resende e Silveira Neto (2009), Lima et al. (2013), Gurgel et al. (2013), Braga et al. (2013), Resende (2014a), Resende (2014b), Resende, Cravos e Pires (2014), Oliveira, Menezes e Resende (2018), Resende, Silva e Silva Filho (2015) e Bia e Mattei (2007).

Caso o GPS não atenda às pressuposições de balanceamento e normalidade, Cerulli (2015) sugere estimação mais flexível das funções dose resposta utilizando o estimador de Mínimos Quadrados Ordinários a partir da seguinte regressão:

$$y_i = \mu_0 + D_i ATE + \lambda_0 X_i + D_i (X_i - \bar{X}) \lambda_1 + D_i [h(c_i) - \bar{h}] + \eta_i \quad (16)$$

em que μ_0 é uma constante, D_i é uma variável binária que assume valor 1 se a empresa recebeu crédito e 0 caso contrário, ATE é o efeito médio de tratamento, X_i é o vetor de características observáveis, λ_0 e λ_1 são vetores de parâmetros a ser estimados, $h(c_i)$ é uma função polinomial em relação ao volume de crédito recebido e η_i é o termo de erro.

Com os parâmetros estimados em (16) pode-se então obter de forma consistente as funções dose resposta. Supondo uma forma quadrática para a função $h(c_i)$ a função dose resposta seria obtida a partir de:

$$\widehat{ATE}(c_i) = D \left\{ \widehat{ATT} + \hat{a} \left(c_i - \frac{1}{N} \sum_{i=1}^N c_i \right) + \hat{b} \left(c_i^2 - \frac{1}{N} \sum_{i=1}^N c_i^2 \right) \right\} + (1 - D) \widehat{ATNT} \quad (17)$$

onde $\widehat{ATT}(c_i) = \widehat{ATE}(c_i)_{c_i > 0}$ e $\widehat{ATNT}(c_i) = \widehat{ATE}(c_i)_{c_i = 0}$.

Nos casos em que o GPS não atendeu as pressuposições necessárias para estimação das funções dose resposta foi utilizada essa forma alternativa. Para estimação das funções dose resposta são consideradas o porte das empresas e a finalidade do crédito (capital de giro e

custeio ou investimento). Os efeitos dose-resposta são calculados para cada produto (estoque de emprego, massa salarial e salário médio).

2.3. REGRESSÃO QUANTÍLICA

Como se sabe, os estimadores de MQO são estimadores de médias condicionais e, como tal, possibilitam interpretar o comportamento médio de um evento estudado, que no caso é o comportamento dos índices utilizados para avaliar a eficácia da política de concessão de créditos do FNE sobre as empresas. Sendo assim, as estimações por MQO não possibilitam investigar qual seria a eficácia dos empréstimos do FNE ao longo da distribuição do faturamento e/ou massa salarial das empresas.

Diante disso, nesta seção é proposta uma metodologia de avaliação da eficácia baseada na técnica de Diferenças em Diferenças (DD), já comentada na seção anterior. Contudo, a análise irá focar em efeitos diferenciados da política ao longo da distribuição dos indicadores das empresas, considerando um grupo de tratados (empresas beneficiadas com a política de crédito) e os não tratados em cada quantil da distribuição desses indicadores. Tal análise ao longo da distribuição é trabalhada tendo como base regressões quantílicas aplicadas em dados em painel.

Segundo Meyer et. al. (1995), os estimadores de regressão quantílica reduziriam drasticamente a importância de *outliers* e permitiriam relaxar as pressuposições sobre as formas funcionais utilizadas na análise. Canay (2011) também ressalta que os modelos de regressão quantílica possibilitam o controle do efeito de variáveis não observadas e de covariáveis heterogêneas, enquanto que Santos (2012) deixa claro que outra vantagem seria o fato de permitir o relaxamento da hipótese do modelo clássico de regressão linear populacional de que os erros se distribuem normalmente.

Considerando tais vantagens, Santos (2012), define o quantil de ordem τ de uma população ou amostra como sendo o valor m , tal que $100\tau\%$ dos valores são inferiores a ele, com $0 < \tau < 1$. Seja Y com f.d.p. acumulada F^4 , o quantil de ordem τ também pode ser

⁴ Admitindo uma população, pode-se especificar a função densidade de probabilidade (f.d.p.) acumulada para a variável aleatória Y (indicadores das empresas) como: $F(Y) = P(Y < y)$. Definindo a inversa da f.d.p. acumulada no ponto τ como: $F^{-1}(\tau) = \inf\{x: F(y) \geq \tau\}$

definido como sendo o valor m que minimiza $E(Y - m)$, ou seja, aquele que minimiza os valores que estão acima (fora) do valor que define o quantil.

Esta ideia, de acordo com Koenker (2005), conduz ao conceito de função de perda, $\rho_\tau(\mu)$, que teria como principal objetivo minimizar a perda esperada em cada quantil possibilitando, dessa forma, trabalhar com o conjunto completo de informações dos indicadores das empresas, considerando toda sua distribuição, mediante a ponderação pelos quantis.

Sendo assim, dada uma amostra de n observações dos indicadores de empresas (Y), o quantil amostral de ordem τ resolveria o problema de minimização apresentado na equação (18), apresentado por Koenker (2005), sendo abordado através de métodos de programação linear (simplex, por exemplo), uma vez que os erros, neste caso, não zeram na média:

$$\min_{q \in \mathbb{R}} \sum_{i=1}^n \rho_\tau(Y_i - q) \quad (18)$$

A intenção é especificar o quantil condicional de Y , dado X , sendo que X são as variáveis independentes do modelo e $\beta(\tau)$ é o vetor de parâmetros definido para cada quantil, como função linear nos parâmetros da forma:

$$Q_\tau(Y/X) = X\beta(\tau) \quad (19)$$

Bastaria definir um vetor $\hat{\beta}(\tau)$ que seja a solução do problema de minimização:

$$\min_{\hat{\beta} \in \mathbb{R}^p} \sum_{i=1}^n \rho_\tau(Y_i - X_i\hat{\beta}) \quad (20)$$

Minimizando a soma do erro ponderado pela perda gerada por cada quantil τ , tem-se:

$$Y_i = \beta_0(\tau) + \beta_1(\tau)X_{1i} + \dots + \beta_p(\tau)X_{pi} + u_i \quad (21)$$

em que, u_i são erros aleatórios identicamente distribuídos.

$$Q_{\tau}(Y/x) = \beta_0(\tau) + \beta_1(\tau)X_{1i} + \dots + \beta_p(\tau)X_{pi} \quad (22)$$

Deve-se destacar que β deve ser indexado a τ , pois um dos interesses, neste caso, é exatamente estudar se o vetor β assumiria diferentes valores para τ 's diferentes. Por considerar diferentes pontos de distribuição da variável condicional Y , esses modelos ajudam a obter uma visão mais completa da relação entre as variáveis estudadas.

De acordo com Canay (2011), a utilização de estimadores de regressão quantílica com um painel, permite, além das vantagens citadas na seção anterior, inserir efeitos fixos para controlar os efeitos de covariáveis não observadas. No presente trabalho, é utilizado o Método dos Momentos – Regressão Quantílica (MM-RQ) apresentado por Machado e Silva (2019). Este método para estimação dos parâmetros ao longo da distribuição da variável dependente permite que os efeitos atuem como deslocadores de localização, mas também possibilitam mensurar os efeitos individuais ao longo de toda a distribuição.

Machado e Silva (2019) indicam que, diante de determinadas pressuposições, a relação entre o fenômeno estudado e seus argumentos poderia ser definida como:

$$Y_{it} = \alpha_i + X'_{it}\beta + (\delta_i + Z'_{it}\gamma)U_{it} \quad (23)$$

Na equação (23), Y_{it} representa os indicadores de interesse das empresas. As variáveis explicativas estariam na matriz de dados X'_{it} . Z'_{it} é um vetor de tamanho k de maneira que seria construído através de transformações em X'_{it} e teria probabilidade igual a 1⁵. O termo de erro U_{it} seria independente e identicamente distribuído (i.i.d.), normalizado para atender as condições de momento impostas por Machado e Silva (2019). Diante disso, a equação (21) pode ser escrita como:

$$Q_Y(\tau|X_{it}) = (\alpha_i + \delta_i q(\tau)) + X'_{it}\beta + Z'_{it}\gamma q(\tau) \quad (24)$$

Na equação (24) o efeito fixo no quantil τ é definido como $\alpha_i(\tau) \equiv (\alpha_i + \delta_i q(\tau))$, também chamado de efeito distributivo em τ . Gouveia (2020) indica o impacto de efeitos fixos,

⁵ Machado e Silva (2019) apresentam as condições para a transformação no vetor Z .

que seriam distintos ao longo da distribuição condicional de Y . O tempo é representado pelo subscrito $t = 0, 1$ e as empresas são representadas pelo subscrito $i = 1, 2, \dots, n$.

Na análise de eficácia para empresas, respectivamente, são separados em grupos de tratados (que tiveram acesso ou condições diferenciadas de acesso ao crédito do FNE no tempo 1) e os não tratados (os que não tiveram acesso no tempo 0 permaneceram sem acesso no tempo 1). Entretanto, como mencionado anteriormente, a eficácia pode ter apresentado efeitos distintos ao longo da distribuição dos indicadores considerados. Por exemplo, os empréstimos podem ter sido mais eficazes para elevar o faturamento em empresas tratadas, quando comparadas às não tratadas, quando é trabalhado em empresas que estão concentradas na parte inicial da distribuição do estoque de emprego, ou seja, nas de menor porte. Isto poderia ser um indicador de que o retorno seria mais eficaz quando a política se concentrasse nesse público empresarial.

Autores como Meyer et. al. (1995) e Athey e Imbens (2006) apresentam estimadores de diferenças em diferenças (DD) quando se trabalha com regressões quantílicas. Carvalho e Firpo (2014) enfatizam que, quando se trabalha com DD em regressões quantílicas, as pressuposições seriam mais fortes. Seria necessária realizar a imposição de que, considerando as variáveis de controle, as variáveis não observáveis deveriam ter a mesma distribuição para os grupos de controle e de tratamento em qualquer período no tempo, independentemente do período de adoção ou não da política analisada.

Considerando a inserção da política no painel quantílico, a equação (24) seria definida como:

$$Q_Y(\tau|X_{it}) = (\alpha_i + \delta_i q(\tau)) + \theta_\tau D_{it} + X'_{it}\beta + Z'_{it}\gamma q(\tau) \quad (25)$$

D_{it} é uma variável *dummy* definida igual a 1 caso a empresa não tenha acessado o crédito do FNE no período $t=0$ e tenha tido acesso ao crédito no período $t=1$. D_{it} assumira valor zero caso a empresa não acesse o crédito no período $t=0$ e se mantenha sem o crédito no período $t=1$. O parâmetro θ_τ seria responsável por avaliar se houve impacto diferenciado nas empresas tratadas e de controle. Considerando a mediana ($\theta_{\tau=0,5}$), tem-se:

$$\theta_{\tau=0,5} = E[(Y_{1,t=1} - Y_{1,t=0}) - (Y_{0,t=1} - Y_{0,t=0}) | X'_{it}, Z'_{it}, D_{i,t=1} = 1] \quad (26)$$

Observa-se que o primeiro termo da equação (26) apresenta o diferencial nos índices das empresas que receberam os créditos do FNE no segundo momento (tratados) e o segundo termo é a diferença de desempenho do grupo de controle, ou seja, a diferença entre os períodos analisados dos índices das empresas que não receberam empréstimos do FNE. $\theta_{\tau=0,5} > 0$ seria o indicativo de que a política de crédito do FNE, concedido às empresas localizadas na mediana da distribuição dos indicadores utilizados foi eficaz.

Inicialmente, a análise de eficácia leva em consideração os efeitos da política de concessão de crédito sobre o estoque de emprego, a massa salarial e o salário médio das empresas, ou seja, os retornos diretos esperados com a implantação da política. Como citado anteriormente, no período inicial, nenhuma das empresas utilizadas na amostra devem ter acesso à política de créditos do FNE. No período posterior, a amostra é dividida entre empresas tratadas, que são as que tiveram acesso ao crédito, e empresas de controle, definidas como as que permaneceram sem ter acesso ao crédito.

O modelo para analisar a eficácia é apresentado a seguir pela equação (27):

$$Y_{it} = \alpha_{i\tau} + \theta_{\tau}D_{it} + \beta_{1\tau}TE_{it} + \beta_{2\tau}S_{it} + \beta_{3\tau}UF_{it} + \beta_{4\tau}SA_{it} + \beta_{5\tau}NQ_{it} + u_{it} \quad (27)$$

A variável dependente Y_{it} é composta pelos seguintes indicadores de resultados das empresas: estoque de empregados, necessário para captar a geração de emprego entre os períodos analisados; massa salarial, capaz de captar a geração de renda por parte da empresa; e salário médio, útil para analisar se a empresa foi capaz de elevar a produtividade. Os efeitos fixos $\alpha_{i\tau}$ são estimados para cada quantil τ . D_{it} é a variável de tratamento de modo que: $D_{it} = 1$ se a empresa teve acesso ao crédito no período 1; $D_{it} = 0$ se no período 1 a empresa não teve acesso ao crédito do FNE; $D_{it} = 0$ para todo $t = 0$.

Como apresentado, é possível afirmar que são estimados quatro modelos para análise de eficácia, sendo que cada modelo considera a distribuição de Y_{it} nos quantis $\tau = 0,10; 0,25; 0,50; 0,75; 0,90$. No primeiro bloco de modelos, leva-se em consideração $t = 2000, 2018$. No segundo bloco tem-se modelos com $t = 2000, 2010$, finalizando com $t = 2010, 2018$. θ_{τ} é responsável, como já citado, por indicar se o Programa de Financiamento foi eficaz e em qual ponto da distribuição ele teria sido mais ou menos eficaz ou não teria sido eficaz. Como

exemplo, ao encontrar $\theta_{\tau=0,10} > 0$, $\theta_{\tau=0,90} > 0$ e $\theta_{\tau=0,10} > \theta_{\tau=0,90}$, indicaria que, tratando-se de faturamento, a política de concessão de créditos do FNE seria eficaz para as empresas de menor e de maior porte (estoque de emprego), mas que teria sido mais eficaz nas empresas de menor porte.

Para finalizar, as variáveis de controle são apresentadas como seguem: TE_{it} denota o tamanho da empresa; S_{it} denota o setor de atividade da empresa; UF_{it} denota a unidade da federação a qual a empresa pertence; SA_{it} indica se a empresa pertence à região semiárida; NQ_{it} denota o nível de qualificação da empresa; e u_{it} denota o termo de erro.

2.4. FONTE E TRATAMENTO DOS DADOS

As bases de dados em nível de firmas foram disponibilizadas pela Sudene, cujas fontes originárias são o Banco do Nordeste (BNB) e o Ministério da Economia (ME), sob acordo de confidencialidade, devido à presença de informações sensíveis que possibilitariam a identificação dos beneficiários, como CPF (pessoas físicas) e CNPJ (pessoas jurídicas).

A base de dados fornecida pelo BNB compreende os contratos firmados entre o banco (aplicador dos recursos do FNE)⁶ e os beneficiários. A base de dados fornecida pelo ME corresponde à Relação Anual de Informações Sociais (RAIS), as quais devem ser entregues anualmente pelas empresas que se enquadrem nos critérios delimitados pelo Ministério. A RAIS está disponível, de modo desidentificado, nas bases de dados on-line do ME. A RAIS disponibiliza as informações em nível de funcionários das empresas.

A princípio, os dados da RAIS foram organizados em nível de empresa ao invés de em nível de indivíduo (funcionário). As variáveis foram convertidas em agregados por empresa, como idade média, proporção de funcionários por grau de instrução, remuneração média, massa salarial e estoque de emprego (número de funcionários). Devido à disponibilidade de dados, foi considerado o período entre 2000 e 2018. As variáveis de massa salarial e remuneração média foram corrigidas pelo Índice de Preços ao Consumidor Amplo (IPCA), passando a ser expressas em preços constantes de 2018

⁶ A partir das Portarias Ministeriais nº 23/2017 e nº 147/2018, do então Ministério da Integração Nacional, é possível que outros bancos operem com os Fundos Constitucionais.

Posteriormente, utilizando o mesmo recorte de tempo, 2000 a 2018, os dados de contratos do BNB também foram convertidos em dados agregados por empresas. Assim, foram utilizadas as informações das empresas beneficiárias para juntá-las às informações da RAIS, o que foi possível através da informação do CNPJ. A variável crédito concedido já estava deflacionada na base de dados fornecida pelo BNB, o índice utilizado foi o IPCA com base em dezembro de 2021. Apesar das variáveis em bases diferentes, não houve influência nas estimações, nos modelos DD e quantílica não se utiliza a variável volume de crédito, já no modelo de dose resposta essa variável é normalizada em relação ao maior valor.

Adicionalmente às bases de dados utilizadas nas estimações, para calcular indicadores socioeconômicos foi utilizada a Pesquisa Nacional por Amostra de Domicílios Contínua (Pnad Contínua) dos anos de 2019 e 2020, disponibilizada pelo IBGE. Os valores monetários foram deflacionados utilizando o IPCA base 2020.



IPPDS
Instituto de Políticas Públicas e
Desenvolvimento Sustentável

3. PRODUTO 6 – AVALIAÇÃO DA EFICIÊNCIA DO FNE

Medidas de eficiência podem ser calculadas pela obtenção de uma tecnologia eficiente, que, por sua vez, é geralmente representada por algum tipo de função fronteira. As fronteiras, por sua vez, podem ser estimadas por diferentes métodos, sendo que os dois mais utilizados são as fronteiras estocásticas e a análise envoltória de dados (DEA). As fronteiras estocásticas consistem em abordagens paramétricas, sendo estimadas por métodos econométricos, enquanto a técnica DEA é uma abordagem não paramétrica, que envolve programação matemática em sua estimação.

Os modelos DEA são baseados em uma amostra de dados observados para diferentes unidades produtoras, também conhecidas como DMUs. Na literatura relacionada com modelos DEA, uma unidade produtora é tratada como DMU (*decision making unit*), uma vez que desses modelos provém uma medida para avaliar a eficiência relativa de unidades tomadoras de decisão. Por unidade produtora entende-se qualquer sistema que transforme insumos em produtos. O objetivo é construir um conjunto de referência a partir dos próprios dados das DMUs, e então classificá-las em eficientes ou ineficientes, tendo como referencial essa superfície formada.

Uma pressuposição fundamental na técnica DEA é que, se uma DMU A é capaz de produzir $Y(A)$ unidades de produtos, utilizando-se $X(A)$ unidades de insumos, outras DMUs poderiam também fazer o mesmo, caso elas estejam operando eficientemente. Com isso, é possível identificar quais são as gestões de referência, conhecidos como *benchmarks*, que são as DMUs que adotam as melhores práticas. A partir das DMUs consideradas eficientes é possível construir a fronteira de possibilidade de produção. As DMUs que formam a fronteira de possibilidade de produção servirão de referência para aquelas consideradas ineficientes. A partir dessa informação torna-se possível projetar as DMUs ineficientes para a fronteira, levando-se em consideração seus *benchmarks*.

Considere, como exemplo, um conjunto de DMUs que utilizam dois insumos (X_1 e X_2), para produzir um único produto (Y). Tendo em vista a pressuposição de retornos constantes à escala, ou seja, admitindo-se que a função de produção seja homogênea de grau um nos insumos, a fronteira tecnológica pode ser representada por isoquantas unitárias, conforme ilustra a Figura 1.

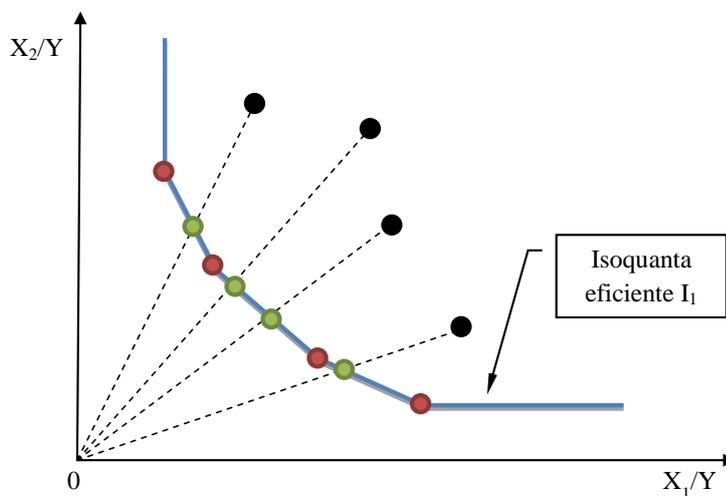


Figura 1. Construção da isoquanta eficiente e projeção das unidades ineficientes.

Fonte: Elaboração própria.

Na Figura 1, as DMUs representadas pelos pontos vermelhos são tecnicamente mais eficientes do que aquelas representadas pelos pontos pretos, uma vez que conseguem produzir a mesma quantidade utilizando proporcionalmente menos insumos. A isoquanta unitária eficiente I_1 é formada por estas DMUs, isto é, as unidades eficientes encontram-se sobre a isoquanta, enquanto as ineficientes estão acima. Não existe nenhuma combinação de insumos e produto melhor que as representadas pela isoquanta I_1 .

Para que as unidades ineficientes se tornem eficientes é preciso projetá-las na fronteira. Essa projeção pode ser feita por meio de uma contração radial, ou seja, a partir de raios que saem da origem do gráfico. Os pontos em verde representam as DMUs ineficientes projetadas na isoquanta eficiente. É importante destacar que os pontos projetados são combinações lineares das DMUs eficientes que compuseram a isoquanta. Em outras palavras, a projeção de uma DMU ineficiente é realizada levando-se em consideração seus *benchmarks* representados pelos pontos vermelhos.

Um problema envolvendo apenas dois insumos e um produto pode ser facilmente resolvido utilizando-se apenas gráficos semelhantes ao apresentado na Figura 1. Porém, quando se têm múltiplos insumos e/ou produtos, torna-se necessária a elaboração de problemas de programação linear.

Supondo que existam k insumos e m produtos para cada n DMUs, são construídas duas matrizes: a matriz X de insumos, de dimensões $(k \times n)$ e a matriz Y de produtos, de dimensões $(m \times n)$, representando os dados de todas as n DMUs. A partir desses dados são formulados os problemas de programação linear para calcular as medidas de eficiência, tendo como base a fronteira de eficiência.

Os modelos DEA podem assumir tanto retornos constantes como variáveis à escala de produção. O modelo considerando-se a pressuposição de rendimentos constantes à escala, também conhecido como CCR, em referência aos seus idealizadores Charnes, Cooper e Rhodes (1978), é dado pelo seguinte problema de programação matemática:

$$\begin{aligned} & \text{MIN}_{\theta, \lambda} \quad \theta, \\ & \text{sujeito a :} \\ & \quad -y_i + Y\lambda \geq 0, \\ & \quad \theta x_i - X\lambda \geq 0, \\ & \quad \lambda \geq 0, \end{aligned} \tag{28}$$

em que θ é um escalar, cujo valor será a medida de eficiência da i -ésima DMU. Caso o valor de θ seja igual a um, a DMU será eficiente; caso contrário será ineficiente. O parâmetro λ é um vetor $(n \times 1)$, cujos valores são calculados de forma a obter a solução ótima. Para uma DMU ineficiente, os valores de λ serão os pesos utilizados na combinação linear de outras DMUs eficientes, que influenciam a projeção da DMU ineficiente sobre a fronteira calculada. Isto significa que, para uma unidade ineficiente, existe pelo menos uma unidade eficiente, cujos pesos calculados fornecerão a DMU virtual da unidade ineficiente, mediante combinação linear. As unidades eficientes que, quando combinadas, fornecem a DMU virtual para a unidade ineficiente são conhecidas como pares ou *benchmarks* daquela DMU.

O problema de programação linear com retornos constantes pode ser modificado para atender à pressuposição de retornos variáveis, ou modelo BCC devido a Banker, Charnes e Cooper (1984), adicionando-se uma restrição de convexidade (COELLI et al., 2007). Dessa forma, o problema de programação linear com retornos constantes pode ser modificado para atender à pressuposição de retornos variáveis, adicionando-se a restrição de convexidade, em que N_I é um vetor $(n \times 1)$ de algarismos unitários:

$$\text{MIN}_{\theta, \lambda} \quad \theta,$$

sujeito a:

$$\begin{aligned} -y_i + Y\lambda &\geq 0, \\ \theta x_i - X\lambda &\geq 0, \\ N_1' \lambda &= 1, \\ \lambda &\geq 0. \end{aligned} \tag{29}$$

Essa abordagem forma uma superfície convexa de planos em interseção, a qual envolve os dados de forma mais compacta do que a superfície formada pelo modelo com retornos constantes. Isto quer dizer que o ponto projetado para cada DMU ineficiente será uma combinação convexa das DMUs eficientes observadas. As medidas de eficiência com retornos constantes são compostas pelas medidas de eficiência com retornos variáveis (a pura eficiência), bem como pela eficiência de escala, fazendo com que os valores obtidos para eficiência técnica, com a pressuposição de retornos variáveis, sejam maiores ou iguais aos obtidos com retornos constantes.

A eficiência de escala é obtida pela razão entre a eficiência com retornos constantes e a eficiência com retornos variáveis. Quando a medida for igual a um, a DMU estará operando com retornos constantes, entretanto, se os valores forem inferiores à unidade, não se sabe se a DMU estará realizando sua produção com retornos crescentes ou decrescentes. Para resolver tal limitação, admite-se a pressuposição de retornos não-crescentes, alterando a pressuposição de retornos variáveis no modelo DEA, substituindo-se a restrição $N_1' \lambda = 1$ em (04) por $N_1' \lambda \leq 1$.

Dessa forma, caso a medida de eficiência com retornos variáveis seja igual à medida de eficiência com retornos não-crescentes, a DMU estará operando com retornos decrescentes; caso contrário, estará operando com retornos crescentes. A hipótese de retornos não-decrescentes também pode ser admitida, bastando substituir a restrição $N_1' \lambda \leq 1$ por $N_1' \lambda \geq 1$.

O modelo até então apresentado busca identificar a ineficiência técnica das DMUs mediante redução proporcional na utilização dos insumos, ou seja, são modelos com orientação

a insumo. Entretanto, pode-se também obter medidas de eficiência técnica, baseadas no aumento do produto, também conhecido como orientação produto:

$$\begin{aligned}
 & \text{MAX}_{\varphi, \lambda} \quad \varphi, \\
 & \text{sujeito a :} \\
 & \quad -\varphi y_i + Y\lambda \geq 0, \\
 & \quad x_i - X\lambda \geq 0, \\
 & \quad \lambda \geq 0, \\
 & \quad N_1' \lambda = 1,
 \end{aligned} \tag{30}$$

em que $1 \leq \varphi < \infty$ corresponde ao aumento proporcional no produto considerado, mantendo-se constante a utilização dos insumos em questão. O parâmetro λ é um vetor ($n \times 1$), cujos valores são calculados de forma a obter a solução ótima. Para uma DMU eficiente, o valor de λ_k será igual a um, e todos os demais valores de λ serão zero, enquanto para uma DMU ineficiente, os valores serão os pesos utilizados na combinação linear de outras DMUs eficientes, que influenciam a projeção da ineficiente sobre a fronteira calculada.

Para cada unidade ineficiente, os modelos DEA fornecem seus respectivos *benchmarks*, determinados pela projeção dessas unidades na fronteira de eficiência. Essa projeção é feita de acordo com a orientação do modelo. Pode ser orientação a insumos quando se deseja minimizar os recursos, mantendo-se os valores dos produtos constantes, ou orientação a produtos quando se deseja maximizar os produtos sem diminuir os insumos.

Verifica-se que, como em qualquer técnica empírica, o modelo DEA é baseado em suposições, necessitando serem reconhecidas: (i) sendo determinístico, produz resultados que são particularmente sensíveis a erros de medida; (ii) DEA só mede a eficiência relativa da melhor prática entre um exemplo particular. Portanto, não é significativo comparar os escores de eficiência entre diferentes estudos, porque a melhor prática entre os estudos é desconhecida; e (iii) é sensível à especificação dos fatores e ao tamanho do grupo sob análise.

Existem vários outros modelos e pressuposições que podem ser incorporados na formulação dos problemas de programação utilizados pela DEA. Para descrições mais detalhadas da metodologia recomenda-se a consulta de livros-textos como, por exemplo, Cooper, Seiford e Tone (2004), Coelli et al. (2007) e Ferreira e Gomes (2020).

3.1. ÍNDICE DE MALMQUIST

Existindo dados de insumos e de produtos para mais de um período de tempo, pode-se calcular a mudança na produtividade total dos fatores ocorrida entre dois períodos. Essas mudanças são mensuradas utilizando-se números índices. Para verificar o comportamento da produtividade total dos fatores (PTF) ao longo do tempo optou-se por utilizar o índice de Malmquist. A opção pelo uso deste índice deveu-se ao fato de que ele pode ser decomposto em um índice de mudança na eficiência técnica e um índice de mudança tecnológica.

Em outras palavras, a análise de eficiência/produtividade pode ser desmembrada em duas partes: mudança da distância em relação à fronteira tecnológica, também chamada de mudança de eficiência ou efeito emparelhamento, e mudança da fronteira tecnológica ao longo do tempo, também chamada de progresso tecnológico ou efeito deslocamento. Neste trabalho, o efeito emparelhamento será utilizado para verificar as mudanças ocorridas na eficiência técnica ao longo do período avaliado.

De forma semelhante à análise de eficiência, os índices de mudança de Malmquist podem ser calculados utilizando orientações insumo ou produto. Contudo, para Krüger, Cantner e Hanusch (1998) e Ferreira e Gomes (2020), a orientação produto é uma pressuposição mais razoável em ambiente macroeconômico, pois está mais próxima do objetivo da política de crescimento que visa um maior produto possível, dada uma dotação de recursos.

A Figura 2 ilustra essa definição, considerando-se um modelo simples de um insumo e um produto.

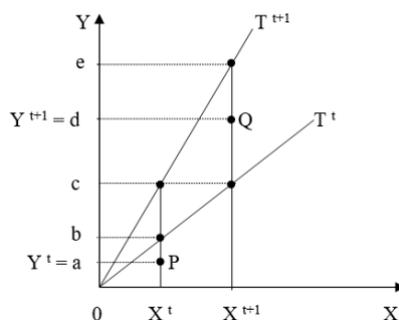


Figura 2. Representação do índice de produtividade de Malmquist

Fonte: Ferreira e Gomes (2020).

Na Figura 2, T e T^{t+1} representam as tecnologias de produção em dois períodos. Considere uma DMU produzindo no ponto P no período t e no ponto Q no período t+1. A mudança de eficiência é dada pela razão entre a eficiência técnica no período t+1 em relação ao período t, ou seja,

$$\text{Mudança de eficiência} = \frac{\frac{0d}{0e}}{\frac{0a}{0b}} \quad (31)$$

A mudança tecnológica é a média geométrica entre o deslocamento da tecnologia avaliada em X^{t+1} e o deslocamento da tecnologia avaliada em X^t, obtida de seguinte forma:

$$\text{Mudança tecnológica} = \left[\frac{0d}{0c} \frac{0a}{0b} \right]^{\frac{1}{2}} \times \left[\frac{0d}{0e} \frac{0a}{0c} \right] \quad (32)$$

O cálculo do índice de Malmquist entre os períodos t e t+1 é baseado em quatro funções distância: $d_0^t(x_t, y_t)$ representa o uso dos dados de insumos e produtos do período t com a tecnologia existente no período t; $d_0^{t+1}(x_{t+1}, y_{t+1})$ calculada com dados de t+1 e tecnologia de t+1; $d_0^{t+1}(x_t, y_t)$ calculada com dados de t e tecnologia de t+1; e $d_0^t(x_{t+1}, y_{t+1})$ calculada com dados de t+1 e tecnologia de t. O índice de mudança na produtividade de Malmquist com orientação produto, apresentado em Caves, Christensen e Diewert (1982), pode ser definido como:

$$M_0^{t,t+1} = \left[\frac{d_0^{t+1}(x_{t+1}, y_{t+1})}{d_0^t(x_t, y_t)} \right] \times \left[\frac{d_0^t(x_{t+1}, y_{t+1})}{d_0^{t+1}(x_{t+1}, y_{t+1})} \times \frac{d_0^t(x_t, y_t)}{d_0^{t+1}(x_t, y_t)} \right]^{\frac{1}{2}} \quad (33)$$

em que o primeiro termo do lado direito da equação mede a mudança de eficiência (efeito emparelhamento) e o segundo termo mede o progresso tecnológico (efeito deslocamento).

Supondo que existam k insumos e m produtos para n unidades tomadoras de decisão (*decision making units* - DMUs), são construídas duas matrizes: a matriz X de insumos, de dimensões $(k \times n)$ e a matriz Y de produtos, de dimensões $(m \times n)$, representando os dados de todas as DMUs. Para calcular os componentes do índice de Malmquist, é necessário resolver quatro problemas de programação linear do tipo:

$$\begin{aligned}
 [d_0^p(x_q, y_q)]^{-1} &= \text{MAX}_{\varphi, \lambda} \quad \varphi, \\
 \text{sujeito a: } & -\varphi y_{i,q} + Y_p \lambda_i \geq 0, \\
 & x_{i,q} - X_p \lambda_i \geq 0, \\
 & \lambda_1, \dots, \lambda_n \geq 0, \\
 & N_1' \lambda_i = 1,
 \end{aligned} \tag{34}$$

com $(p, q) \in \{(t, t), (t+1, t+1), (t, t+1), (t+1, t)\}$, em que y_i é um vetor $(m \times 1)$ de quantidades de produto da i -ésima DMU que está em análise; x_i é um vetor $(k \times 1)$ de quantidades de insumo da i -ésima DMU; Y é uma matriz $(n \times m)$ de produtos das n DMUs; X é uma matriz $(n \times k)$ de insumos das n DMUs; e λ_i é um vetor $(n \times 1)$ de pesos, cujos valores são calculados de forma a obter a solução ótima. Os valores obtidos para φ indicam a quantidade máxima de aumento em todos os produtos do período em análise (t ou $t+1$), com os insumos constantes requeridos para obter um ponto na função fronteira no período em que os dados são considerados (t ou $t+1$).

Os problemas de programação linear apresentados em (09) são solucionados pela técnica conhecida como análise envoltória de dados (DEA), que constrói fronteiras de eficiência considerando-se o banco de dados de insumos e produtos. Na DEA, a eficiência de cada unidade (DMU) é avaliada relativamente às demais. Por exemplo, se forem considerados $p=t$ e $q=t$, a equação (09) representa um modelo com retornos variáveis e orientação produto que calcula a eficiência para a i -ésima DMU com dados e tecnologia do período t .

3.2. TESTE DE CONVERGÊNCIA DAS MEDIDAS DE EFICIÊNCIA

De modo geral, pode-se dizer que existem dois tipos de convergência: β -convergência e σ -convergência. A β -convergência caracteriza-se por uma relação negativa entre o valor

inicial da variável e sua taxa de crescimento, o que implica que DMUs com menores valores iniciais de eficiência tendem a ter taxas maiores de mudanças na eficiência, do que aquelas com maiores valores. O conceito de β -convergência pode ser dividido em dois outros conceitos: β -convergência absoluta e β -convergência condicional.

A β -convergência absoluta indica que existe um único estado estacionário para o qual todas DMUs tendem. O estado estacionário será atingido quando ocorrer a completa equidade nas variáveis. Assim, quanto mais longe uma DMU estiver do estado estacionário, maior será sua taxa de crescimento. Em outras palavras, os indicadores das DMUs com menores valores de eficiência tendem a crescer mais rapidamente, uma vez que o valor original é baixo e o estado estacionário é constante para todos.

Caso não se constate a existência de β -convergência absoluta, pode ser que as DMUs ou grupos de DMUs tenham diferentes estados estacionários. Com isso, a taxa de crescimento de cada DMU será tanto maior quanto mais afastado ela estiver de seu próprio estado estacionário. Esse conceito refere-se à β -convergência condicional, uma vez que a taxa de crescimento da variável é positivamente relacionada com a distância que a separa de seu próprio estado estacionário. Nesse caso, não será válido dizer que as medidas de eficiência das DMUs com menores valores crescerão mais rápido. Para testar sua ocorrência, deve-se verificar a existência de relação negativa entre crescimento de variável e o valor inicial, quando se inserem na equação outras variáveis.

Para a verificação da existência de β -convergência nas medidas de eficiência, será utilizado o teste apresentado em Barro e Sala-I-Martin (1992). A hipótese de β -convergência é tradicionalmente testada por um modelo de regressão linear simples, pelo qual se estima a taxa de crescimento da variável em relação ao seu valor inicial. A equação básica deste teste é expressa por:

$$\frac{1}{T} \ln \left(\frac{Y_{it}}{Y_{i,0}} \right) = \beta_1 + \beta_2 \ln (Y_{i,0}) + \mu_i \quad (35)$$

em que $Y_{i,0}$, e $Y_{i,T}$ representam os indicadores de eficiência dos períodos inicial e final, respectivamente; e μ_i é o erro aleatório. O lado esquerdo da equação corresponde à taxa de crescimento da variável. Uma correlação negativa entre a taxa de crescimento e o valor inicial

(β_2 significativo e menor que zero) indica que está ocorrendo β -convergência absoluta, cuja velocidade de convergência é obtida através da expressão $\beta_2 = 1 - \frac{e^{-\beta t}}{t}$.

Neste trabalho, para testar a existência de convergência das medidas de eficiência, a taxa de mudança na eficiência técnica será obtida pela decomposição do índice de Malmquist, utilizando-se o efeito emparelhamento, ao passo que o nível de eficiência inicial será calculado pelo modelo DEA, ambos procedimentos conduzidos com retornos variáveis e orientação produto, conforme apresentado anteriormente.

Conforme mencionado, mesmo não havendo β -convergência absoluta, pode ser que a inclusão de outras variáveis na análise permita a existência de β -convergência condicional, alterando o modelo original da seguinte forma:

$$\frac{1}{T} \ln \left(\frac{Y_{it}}{Y_{i,0}} \right) = \beta_1 + \beta_2 \ln (Y_{i,0}) + \delta X + \mu_i \quad (36)$$

em que X representa um vetor de variáveis regionais relativas a outras características relacionadas às DMUs em estudo. Geralmente, essas variáveis são incluídas com seu valor no período inicial da amostra.

Contudo, é importante destacar que a ocorrência de β -convergência condicional não significa que as desigualdades regionais estão reduzindo ou que tendem a desaparecer ao longo do tempo. Ao contrário, significa que as DMUs tendem para uma situação de equilíbrio no longo prazo em que, por apresentarem diferentes estados estacionários, suas disparidades persistirão. Assim, o cálculo da β -convergência condicional será realizado caso não se verifique a β -convergência absoluta.

Já a σ -convergência consiste simplesmente em uma redução da dispersão (desvio padrão ou coeficiente de variação) dos valores atribuídos às variáveis entre as DMUs ao longo do tempo. Isso significa que os valores de eficiência das diferentes DMUs tendem a se aproximar de sua média, em algum momento futuro.

3.3. MÉTODO DE DETECÇÃO DE OUTLIERS

Dado o fato de que um problema crítico do método DEA é a grande sensibilidade na presença de *outliers* e erros amostrais, o presente trabalho utiliza-se a metodologia desenvolvida por Sousa e Stosic (2003) para detectar a presença desses possíveis *outliers* que poderiam afetar a fronteira de eficiência. O estudo de Sousa e Stosic (2003) idealizou uma combinação de duas metodologias de reamostragem, de modo a proceder com uma análise específica para o DEA. A partir dos métodos *jackknife* (determinístico) e *bootstrap* (estocástico), os autores deram origem ao procedimento denominado “*jackstrap*”. Em um primeiro momento, o *jackknife* é utilizado por meio de um algoritmo, que mensura a influência de cada DMU no cálculo das eficiências, isto é, cada DMU é removida isoladamente da amostra para que as eficiências sejam então calculadas sem a sua presença. Em um segundo instante, é utilizado o método *bootstrap* de reamostragem estocástica, levando em consideração a informação das influências obtidas pelo *jackknife*.

O estimador obtido desta maneira é denominado *leverage*, e possibilita uma análise automática da amostra, dispensando uma análise manual que, além de imprecisa, é inviável em grandes amostras. Formalmente, o *leverage* de Sousa-Stosic pode ser definido como o desvio padrão das medidas de eficiência antes e depois da remoção de cada DMU do conjunto amostral. Assim, o *leverage* da j -ésima DMU pode ser definido como:

$$\ell_j = \sqrt{\sum_{k=1; k \neq j}^K (\theta_{kj}^* - \theta_k)^2 / K - 1} \quad (37)$$

em que o índice k são as DMUs, variando de 1 até K , o índice j representa a DMU removida e θ são os indicadores de eficiência. Assim, $\{\theta_k | k = 1, \dots, K\}$ representa o conjunto de eficiências originais, sem alteração na amostra, e $\{\theta_{kj}^* | k = 1, \dots, K; k \neq j\}$ representa o conjunto de eficiências recalculado após a remoção individual de cada DMU.

Presume-se que as DMUs caracterizadas como *outliers* possuam um *leverage* consideravelmente acima da média global. Desta maneira, caso ℓ_j esteja muito acima dessa média, há a suspeita de que a DMU em questão seja uma *outlier*. Quando a DMU j está localizada dentro da fronteira eficiente, ocorre que $\theta_{kj}^* - \theta_k = 0$, e então $\ell_j = 0$, o que significa que a observação em questão não é influente. Por outro lado, no caso crítico de uma

DMU cuja influência seja extrema, sua remoção faz com que pelo menos uma das unidades remanescentes apresente um valor de eficiência igual a 1, isto é, $\sum(\theta_{kj}^* - \theta_k)^2 = K - 1$, e então $\ell_j = 1$. Assim, o índice de *leverage* encontra-se dentro do intervalo $[0,1]$.

Com a informação dada pelo *leverage* é possível então identificar observações *outliers*. Para tanto, é necessário utilizar um critério específico relacionado ao desvio do índice em relação à sua média global. Sousa e Stosic (2005) sugerem um múltiplo da média global, $\tilde{\ell}_0 = c\bar{\ell}$, onde $\bar{\ell}$ representa a média global do *leverage* e c é uma constante que assume valor de 2 ou 3 de modo geral, ou, alternativamente, adota-se $\tilde{\ell}_0 = 0,02$ como critério de corte. Desta forma, DMUs com um *leverage* acima desse valor seriam caracterizadas como *outliers*.

3.4. TESTES NÃO PARAMÉTRICOS DE FRONTEIRAS DE EFICIÊNCIA

Antes de executar os modelos para cálculo das medidas de eficiência, é preciso verificar se as diferentes DMUs fazem parte de uma mesma fronteira de eficiência. Quando existem dois ou mais grupos de DMUs, pode ser que exista mais de uma fronteira de eficiência. Essa situação é ilustrada na Figura 3.

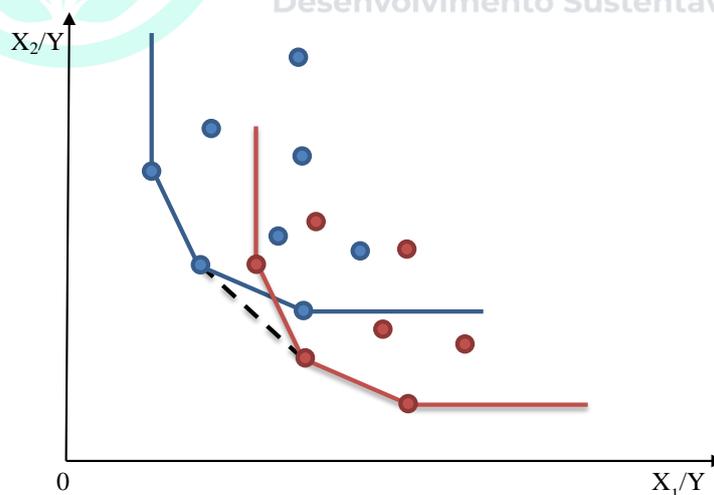


Figura 3. A possibilidade de mais de uma isoquanta eficiente

Fonte: Elaboração própria.

Na situação descrita na Figura 3, se todas as DMUs forem analisadas em conjunto, a isoquanta eficiente seria composta pela parte azul acima/esquerda, seguida pela linha tracejada e depois pela parte vermelha abaixo/direita. Porém, supondo que as DMUs azuis e vermelhas façam parte de dois grupos distintos, percebe-se que existem duas fronteiras. Isso significa que as DMUs ineficientes azuis serão projetadas para a isoquanta azul, enquanto as ineficientes vermelhas serão projetadas para a isoquanta vermelha.

Para verificar se as DMUs de dois grupos distintos fazem parte de uma mesma fronteira, o seguinte procedimento apresentado em Marinho (1996) pode ser utilizado: (1) separar os dados dos grupos previamente definidos. Em seguida, executam-se os modelos DEA em ambos os grupos separadamente; (2) em cada grupo, as DMUs ineficientes devem ser ajustadas para os seus alvos, ou seja, devem ser projetadas para a fronteira eficiente, considerando-se os seus *benchmarks*; (3) a partir dos novos bancos de dados de insumos e produtos otimizados (já projetados nas respectivas fronteiras), executar o DEA para o conjunto global formado pela união dos dois grupos ajustados; e (4) aplicar os testes estatísticos não paramétricos aos coeficientes de eficiência gerados na etapa 3 para testar a hipótese de igualdade estatística entre os grupos.

Os testes não paramétricos U de Mann-Whitney e W de Wilcoxon são utilizados para verificar se há igualdade entre as fronteiras. O teste U de Mann-Whitney avalia se, dentre dois grupos de variáveis aleatórias, uma delas é estocasticamente maior que outra, sendo assim aplicado para verificar se duas amostras independentes pertencem ou não a uma mesma população (BANKER, ZHENG e NATARAJAN, 2010). No caso de a medida de eficiência ser relativa e obtida por comparação, deve-se recorrer ao teste W de Wilcoxon, recomendado para amostras dependentes. Geralmente, ambos os testes fornecem os mesmos resultados em termos de aceitação ou rejeição da hipótese nula.

Segundo Siegel e Castellan Jr. (2006), os testes de Mann-Whitney e de Wilcoxon são dois dos mais robustos métodos estatísticos não paramétricos, e constituem-se em alternativas extremamente úteis da prova paramétrica t , quando se deseja desviar das suposições exigidas por esse teste paramétrico. As vantagens das provas estatísticas não paramétricas são, segundo Siegel e Castellan Jr. (2006): (i) as afirmações probabilísticas decorrentes da maior parte das provas estatísticas não paramétricas são probabilidades exatas, independente da forma de distribuição da população; (ii) diferente das estatísticas paramétricas, há provas estatísticas não

paramétricas adequadas para o tratamento de amostras constituídas de observações de várias populações diferentes; (iii) os métodos não paramétricos aplicam-se ao tratamento de dados simplesmente classificativos, mensurados em escala nominal, diferente de qualquer técnica paramétrica; e (iv) são essencialmente mais fáceis de aprender e de se aplicar do que as provas paramétricas.

3.5. ANÁLISE DISCRIMINANTE

Com o objetivo de verificar quais fatores possuem maior capacidade de discriminar DMUs eficientes e ineficientes, utilizar-se-á da análise discriminante. Tal análise consiste num método de estatística multivariada que possibilita a classificação de elementos de uma dada amostra de acordo com grupos previamente conhecidos, sendo também possível elaborar uma regra de classificação a ser utilizada para enquadrar eventuais novas observações nos grupos existentes (MINGOTI, 2005).

De acordo com Malhotra (2011), a análise discriminante é uma técnica de análise de dados em que a variável dependente tem natureza categórica (separa os elementos em dois ou mais grupos de categoria) e as variáveis independentes têm natureza métrica (medidas em uma escala de razão).

Malhotra (2011) apresenta como objetivos da análise discriminante: (i) estabelecer funções discriminantes, ou combinações lineares das variáveis independentes, que melhor discriminem entre as categorias da variável dependente (grupos); (ii) verificar se existem diferenças significativas entre os grupos, em termos das variáveis independentes; (iii) determinar as variáveis independentes que mais contribuem para a diferença entre os grupos; (iv) classificar os casos em um dos grupos com base no valor das variáveis independentes; e (v) avaliar a precisão da classificação.

Dependendo do número de categorias da variável dependente, a técnica de análise discriminante pode ser de dois grupos, onde é deduzida somente uma função discriminante, ou múltipla, onde pode ser estimada mais de uma função. Neste estudo a técnica utilizada será a de dois grupos.

Segundo Hair et al. (2009), a análise discriminante reduz o número de variáveis para um número menor de parâmetros, que são funções discriminantes linearmente dependentes das

variáveis originais. Os coeficientes da função discriminante indicam a contribuição da variável original para a função. O modelo é dado de acordo com a equação:

$$D = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_n \quad (38)$$

em que D é o valor da função discriminante; β_i são os coeficientes discriminantes; e X_i são os valores das variáveis independentes.

Os coeficientes β_i serão estimados de modo a diferenciar os grupos ao máximo, ou seja, os coeficientes do mesmo grupo serão os mais parecidos possíveis, e estes serão os mais diferentes possíveis dos coeficientes do outro grupo.

O processo estatístico da análise discriminante pode ser dividido em seis estágios: (i) definição dos objetivos, ou seja, definir quais os objetivos que a análise deve alcançar; (ii) pesquisa para a análise discriminante, onde compreende-se a seleção das variáveis dependentes e das variáveis independentes, a escolha do tamanho da amostra e a divisão dessa amostra; (iii) hipóteses da análise discriminante, sendo a hipótese básica da análise discriminante a existência de multivariabilidade entre as variáveis independentes e a dependente, além de uma desconhecida, mas igual estrutura de variância e covariância para os grupos formados, supondo ainda a existência de normalidade das variáveis; (iv) estimação do modelo de análise discriminante, isto é, a estimação dos parâmetros da função discriminante e a obtenção da função Z correspondente; (v) interpretação dos resultados, de forma a interpretar os parâmetros e dividir os grupos determinados pela função discriminante; e (vi) validação dos resultados, ou seja, a verificação do poder de explicação da análise, fazendo uma análise crítica dos resultados encontrados.

Um método comum na análise discriminante é o *stepwise*, que seleciona as variáveis para entrar na análise, baseando-se nas suas capacidades de discriminação. O processo inicia-se selecionando a variável que apresenta maior valor de discriminação. Esta variável é pareada com as demais variáveis, uma de cada vez, e o critério de seleção é novamente comparado. A variável que, em conjunto com a primeira selecionada, produzir o melhor valor para o critério é a segunda variável escolhida para entrar na análise. Estas duas variáveis selecionadas são combinadas com as demais remanescentes e a combinação que apresentar o maior valor para o critério de seleção determinará a terceira variável a entrar na equação. Do mesmo modo, todas

as demais variáveis são testadas através do critério de seleção, até que todas sejam ordenadas pelas suas capacidades de discriminação.

Segundo Malhotra (2011), o método de decisão estatística para o caso de dois grupos classifica uma observação no Grupo 1 se:

$$Z \geq \frac{\bar{Z}_1 + \bar{Z}_2}{2} + \ln \left[\frac{p_2 C(1/2)}{p_1 C(2/1)} \right] \quad (39)$$

e outra observação será classificada no Grupo 2 se:

$$Z < \frac{\bar{Z}_1 + \bar{Z}_2}{2} + \ln \left[\frac{p_2 C(1/2)}{p_1 C(2/1)} \right] \quad (40)$$

em que Z é o valor discriminante para uma dada observação; \bar{Z}_j é valor discriminante médio para o grupo j ; p_j é a probabilidade prévia do grupo j ; e $C(1/2)$ é o custo de classificação incorreta dentro do grupo i de uma observação que pertence ao grupo j .

Para a determinação da significância da função de discriminação, de acordo com Malhotra (2011), pode-se testar estatisticamente a hipótese nula, de que as médias de todas as funções discriminantes em todos os grupos sejam iguais. Para Hair et al. (2009), um dos métodos mais utilizados para identificar o poder discriminatório de uma variável é o teste Lambda de Wilks (L^*). Este teste considera como critério de seleção de variáveis o valor da Estatística F Multivariada, para o teste da diferença entre os centroides dos grupos. A variável que maximiza o valor da estatística F também minimiza o L^* , que é uma medida de discriminação entre os grupos.

Posteriormente, é realizada a interpretação dos resultados, em que o valor do coeficiente β_i para uma determinada variável independente depende das outras variáveis independentes que fazem parte da função discriminante. Por fim, procede-se com a avaliação da validade da análise discriminante, em que os coeficientes β_i estimados são multiplicados pelos valores das variáveis independentes na amostra retida, a fim de gerar valores discriminantes para os casos nessa amostra (MALHOTRA, 2011).

3.6. PROCEDIMENTOS E BASE DE DADOS

Os procedimentos empíricos desta parte do trabalho consistem em um conjunto de análises dividido em dois blocos, um relacionado às análises de eficiência em termos geográficos e outro em termos setoriais. Vale destacar que os procedimentos foram alterados em relação ao projeto original. Tais alterações foram necessárias em razão, principalmente, da inexistência de dados.

No projeto original, as análises seriam realizadas segundo óticas: global e FNE. Por sua vez, o modelo global teria os recortes geográfico e setorial, enquanto o modelo FNE consideraria os recortes de beneficiários e de CNAEs. Contudo, em todos os modelos propostos, a variável faturamento das empresas estaria definida como produto. Devido sua ausência nos bancos de dados disponibilizados, por si só, inviabilizou os recortes propostos no modelo global, uma vez que o faturamento das empresas seria considerado como único produto.

Outro problema detectado foi a inexistência de dados em diversos períodos, principalmente para os anos iniciais. Com isso, o procedimento de avaliação intertemporal utilizando-se análises de janela ficou inviabilizado.

Assim, de forma a proceder com as análises estática e temporal de eficiência, optou-se por modelos em dois blocos, um referente ao recorte regional e outro ao recorte setorial. Sob a ótica regional, foi necessário, ainda, utilizar as DMUs de forma mais agregada, isto é, agrupando-se as informações dos beneficiários segundo a região geográfica imediata à qual pertencem.

O detalhamento dos procedimentos realizados será apresentado a seguir.

3.6.1. MODELO REGIONAL

No modelo denominado regional, o objetivo é verificar onde ocorreram maiores índices de eficiência das aplicações de recursos do FNE na geração de emprego e massa salarial, bem como identificar as mudanças nas medidas de eficiência ao longo do período de 2000 a 2018.

Nesse modelo, as unidades tomadoras de decisão (*decision making unit* - DMUs) são todas as regiões geográficas imediatas do Nordeste e da área de abrangência da Sudene em Minas Gerais e no Espírito Santo.

Novamente, cabe salientar que houve alterações em relação aos modelos propostos originalmente, com o objetivo de adequar os procedimentos à disponibilidade de informações. Inicialmente, pensou-se em utilizar os municípios das regiões atendidas como sendo as DMUs. Contudo, para vários desses municípios não havia disponibilidade de informações, principalmente nos anos iniciais.

Para contornar essa dificuldade, foi necessária uma primeira agregação, passando-se a utilizar as regiões geográficas imediatas como sendo as DMUs. Em outras palavras, os dados de insumos e produtos referem-se aos somatórios das informações de todos os CNPJs registrados em cada região imediata.

Mesmo utilizando o banco de dados agregados por região imediata, ainda persistiram ausência de informações para diversas DMUs ao longo do período avaliado. Essa ausência impossibilita a análise temporal com base em mudanças anuais. Assim, foi necessária uma segunda agregação, desta vez por intervalos de tempo. Optou-se por agrupar as informações em três períodos distintos: de 2000 a 2008, de 2009 a 2013 e de 2014 a 2018. Uma vez que os tamanhos dos períodos são distintos, o primeiro contendo informações de nove anos e os dois últimos com informações de cinco anos, foram utilizados os valores médios das variáveis em cada período.

Após as agregações regional e temporal, a base de dados final ficou constituída por informações referentes às regiões geográficas imediatas (DMUs) em três fronteiras, isto é, uma fronteira para cada intervalo de tempo considerado. A partir dessa base de dados, foi possível avaliar as medidas de eficiência isoladamente em cada fronteira, bem como medir as mudanças ocorridas entre os intervalos. Para tanto, foram utilizadas as seguintes variáveis:

- Insumo: valor financiado pelo FNE;
- Produtos: número de empregados e massa salarial.

Os dados referentes aos financiamentos foram fornecidos pela Sudene, enquanto os valores de emprego e massa salarial foram retirados da Relação Anual de Informações Sociais (RAIS).

Antes de executar os modelos de eficiência, foi necessário verificar a existência de *outliers* nos dados, bem como a possibilidade de existirem diferentes fronteiras de eficiência entre as regiões analisadas. O teste de detecção de *outliers* foi realizado para evitar que informações discrepantes prejudiquem a correta obtenção das medidas de eficiência. Justifica-se seu uso na medida em que as análises não paramétricas de eficiência são sensíveis à presença de *outliers*, podendo sua presença distorcer toda a fronteira que servirá de referência para as unidades avaliadas.

O teste de *outliers* foi conduzido para as três fronteiras isoladamente, pressupondo-se retornos variáveis e orientação produto. Identificada uma DMU discrepante, a região foi removida das análises, evitando-se, assim, que o *outlier* interferisse nos resultados. Isso significa que uma região imediata identificada como *outlier* em uma ou mais fronteiras foi removida, permitindo a comparação de um mesmo conjunto de DMUs em todas as análises.

Após a eliminação das DMUs consideradas como *outliers*, o próximo passo consistiu na obtenção das medidas de eficiência para cada intervalo de tempo, utilizando-se a análise envoltória de dados (DEA) sob a pressuposição de retornos variáveis e orientação produto. É importante destacar que os valores das medidas de eficiência obtidos para os três intervalos de tempo não podem ser comparados entre si, uma vez que as DMUs podem pertencer a fronteiras distintas, como será abordado mais adiante.

Para cada período, foram realizadas análises, agregando-se os resultados nos seguintes blocos: por Unidade da Federação; por localização na região do semiárido; por tipologia segundo a Política Nacional de Desenvolvimento Regional (PNDR); e por estratos de medidas de pura eficiência técnica.

Para esse último bloco de análises, as DMUs foram divididas em três grupos de mesma frequência, de acordo com o valor alcançado no modelo de eficiência com retornos variáveis, ou seja, 1/3 de DMUs com menor eficiência, 1/3 com eficiência intermediária e 1/3 de maior eficiência. Em seguida, foram feitas análises comparando-se os valores médios de diversos indicadores para esses três grupos de regiões imediatas.

Além das análises estáticas das fronteiras nos três intervalos de tempo, foram realizadas técnicas que permitem abordagens temporais na eficiência. Para isso, optou-se pelo Índice de Malmquist, o qual permite a obtenção da medida de mudança na eficiência técnica ao longo do tempo, conhecido como efeito emparelhamento, conforme apresentado anteriormente. O

procedimento foi conduzido separadamente, ou seja, considerando-se as mudanças entre os períodos 1 (2000/2008) e 2 (2009/2013), e as mudanças entre os períodos 2 (2009/2013) e 3 (2014/2018).

Para cada mudança, foram realizadas análises, agregando-se os resultados nos seguintes blocos: por Unidade da Federação; por localização na região do semiárido; por tipologia segundo a Política Nacional de Desenvolvimento Regional (PNDR); e por estratos de mudança na pura eficiência técnica.

Para esse último bloco, as DMUs foram divididas em dois grupos: um grupo das que tiveram ganho de eficiência no período e outro com as DMUs que apresentaram perda de eficiência. Vale ressaltar que, como nos modelos anteriores, foi utilizada a medida de pura eficiência, considerando-se retornos variáveis à escala. Em seguida, foram feitas análises comparando-se os valores médios de diversos indicadores para os dois grupos de regiões.

Após apresentar o comportamento dos índices de mudança na eficiência, foi realizada a análise discriminante, objetivando identificar as variáveis que possam efetivamente estar discriminando as DMUs que obtiveram ganhos de eficiência daquelas que tiveram perda relativa. Tal procedimento foi conduzido para os dois intervalos isoladamente, uma vez que variáveis que discriminam um momento do tempo podem não ser as mesmas em todo período avaliado. As seguintes variáveis foram utilizadas nas análises discriminantes:

- População residente (2018)
- Índice de Desenvolvimento Humano Municipal (2010)
- Percentual da população com ensino superior completo (2010)
- Taxa de mortalidade infantil até um ano (2010)
- Densidade populacional (2018)
- Percentual da população com água encanada (2010)
- Produto Interno Bruto (2018)
- Participação da agropecuária no valor adicionado total (2018)
- Participação da indústria no valor adicionado total (2018)
- Participação dos serviços no valor adicionado total (2018)
- Participação do setor público no valor adicionado total (2018)
- Percentual do valor contratado no FNE destinado ao custeio (média 2000/2018)
- Percentual do valor contratado no FNE destinado ao capital de giro (média 2000/2018)

- Percentual do valor contratado no FNE destinado ao investimento (média 2000/2018)
- Percentual do valor contratado no FNE para outros destinos (média 2000/2018)
- Valor contratado no FNE (média 2000/2018)
- Valor da massa salarial (média 2000/2018)
- Número total de trabalhadores (média 2000/2018)
- Medida de pura eficiência técnica obtida no período inicial da mudança.

A incorporação desta última variável na análise permite verificar se a eficiência inicial pode discriminar o comportamento da mudança, fornecendo evidências sobre uma possível convergência, como será visto a seguir.

Após as análises estáticas e intertemporais das medidas de eficiência, foram realizados procedimentos que visam identificar se houve convergência na eficiência da aplicação dos recursos do FNE na geração de emprego e renda. Para isso, foi utilizado o modelo de β -Convergência absoluta, partindo-se da hipótese de que DMUs que apresentaram menores índices iniciais de eficiência tendem a ter mudanças maiores nessas medidas ao longo do tempo.

Para o modelo de β -Convergência, utilizou-se a taxa de mudança na eficiência técnica (efeito emparelhamento) sendo explicada pelo nível de eficiência inicial, calculado pelo modelo DEA, ambos conduzidos com retornos variáveis e orientação produto. Foram feitos testes de convergência, tanto para a mudança do período 1 para o período 2, quanto do período 2 para o período 3.

Contudo, mesmo identificando a existência de convergência, não se pode afirmar que tal convergência está ocorrendo para situações melhores ou piores. Isso significa que as medidas de eficiência podem estar se aproximando, tanto para patamares superiores quanto para inferiores.

Para identificar essa trajetória, foram realizados testes de verificação de igualdade nas fronteiras. O procedimento foi conduzido comparando-se pares de fronteiras, isto é, fronteira do período 1 (2000/2008) com a do período 2 (2009/2013), período 1 com período 3 (2014/2018), e, por fim, período 2 com período 3.

Caso a hipótese de existência de diferentes fronteiras seja confirmada, pode-se então afirmar que cada intervalo de tempo avaliado possui sua própria fronteira de eficiência. Por

fim, para verificar se determinada fronteira está em patamar superior ou inferior à outra, realizam-se testes de igualdade de médias entre elas.

Em síntese, os procedimentos utilizados permitem não somente identificar se houve convergência nas medidas de eficiência da alocação de recursos do FNE, mas também apontar o sentido dessa convergência. Ademais, uma vez que foram consideradas duas mudanças, pode-se verificar se a segunda mudança reforçou o processo registrado num primeiro momento ou se foram em sentido contrário.

3.6.2. MODELO SETORIAL

No modelo denominado setorial, o objetivo é verificar em quais setores da economia ocorreram maiores índices de eficiência das aplicações de recursos do FNE na geração de emprego e massa salarial, bem como identificar as mudanças nessas medidas ao longo do período de 2000 a 2018.

Nesse modelo, as unidades tomadoras de decisão (*decision making unit* - DMUs) são as diferentes atividades econômicas, definidas pela Classificação Nacional de Atividades Econômicas (CNAE).

Os procedimentos utilizados no modelo setorial são similares aos apresentados para o modelo geográfico. A diferença está na unidade de análise considerada, ou seja, as DMUs do modelo setorial referem-se às somas dos valores de insumo e produtos referentes a determinado CNAE.

No intuito de padronizar os procedimentos, a construção da base de dados para o modelo setorial também foi feita agregando-se as informações em três períodos: de 2000 a 2008, de 2009 a 2013 e de 2014 a 2018. Uma vez que os tamanhos dos períodos são distintos, foram utilizados os valores médios das variáveis em cada período.

A base de dados final ficou constituída por informações referentes aos CNAEs (DMUs) em três fronteiras, isto é, uma fronteira para cada intervalo de tempo considerado. A partir dessa base de dados, foi possível avaliar as medidas de eficiência isoladamente em cada fronteira, bem como medir as mudanças ocorridas entre os intervalos. Para tanto, foram utilizadas as seguintes variáveis:

- Insumo: valor financiado pelo FNE;

- Produtos: número de empregados e massa salarial.

Os dados referentes aos financiamentos foram fornecidos pela Sudene, enquanto os valores de emprego e massa salarial foram retirados da Relação Anual de Informações Sociais (RAIS).

Com o banco de dados contendo as informações setoriais por CNAE, utilizou-se a mesma sequência de análises abordadas para o modelo regional, resumidas a seguir:

1. Realização de testes de detecção de *outliers* isoladamente para as fronteiras, ou seja, para cada intervalo de tempo considerado.

2. Eliminação de CNAEs identificados como *outlier* em uma ou mais fronteiras.

3. Obtenção das medidas de eficiência para as DMUs nas diferentes fronteiras, utilizando-se a análise envoltória de dados (DEA) sob a pressuposição de retornos variáveis e orientação produto.

4. Realização de análises individualizadas por fronteira, agregando-se os resultados nos seguintes blocos: por divisão e seção, segundo classificação dos CNAEs; e por grupos, segundo estratos de medidas de pura eficiência técnica, onde as DMUs foram divididas de acordo com o valor alcançado no modelo de eficiência com retornos variáveis, ou seja, 1/3 de DMUs com menor eficiência, 1/3 com eficiência intermediária e 1/3 de maior eficiência.

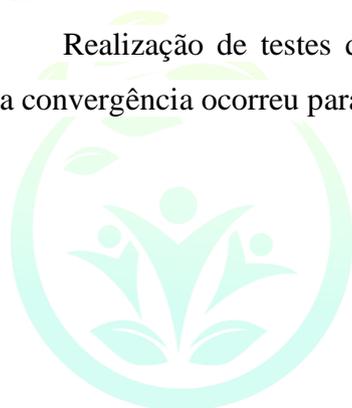
5. Obtenção das mudanças na pura eficiência técnica (efeito emparelhamento), mediante decomposição do Índice de Malmquist. Tal procedimento foi conduzido separadamente, ou seja, considerando-se as mudanças entre os períodos 1 (2000/2008) e 2 (2009/2013), e as mudanças entre os períodos 2 (2009/2013) e 3 (2014/2018).

6. Realização de análises das mudanças, cujos resultados foram apresentados nos seguintes blocos: por divisão e seção, segundo classificação dos CNAEs; e por estratos de mudança na pura eficiência técnica, onde as DMUs foram divididas em dois grupos, separando-se aquelas que tiveram ganho de eficiência no período daquelas que apresentaram perda de eficiência.

7. Utilização da análise discriminante para avaliar se existem variáveis que efetivamente discriminam os CNAEs que apresentaram ganhos ou perdas de eficiência. As seguintes variáveis foram utilizadas nas análises discriminantes:

- Valor contratado no FNE (média 2000/2018)
- Valor da massa salarial (média 2000/2018)

- Número total de trabalhadores (média 2000/2018)
 - Percentual do valor contratado no FNE destinado ao custeio (média 2000/2018)
 - Percentual do valor contratado no FNE destinado ao capital de giro (média 2000/2018)
 - Percentual do valor contratado no FNE destinado ao investimento (média 2000/2018)
 - Percentual do valor contratado no FNE para outros destinos (média 2000/2018)
 - Participação do CNAE no somatório do valor contratado por todos CNAEs (*proxy* para quociente setorial)
 - Medida de pura eficiência técnica obtida no período inicial da mudança.
8. Condução dos testes para verificar a existência de β -convergência nas medidas de eficiência.
9. Realização de testes de verificação de igualdade nas fronteiras, objetivando avaliar se a convergência ocorreu para patamares superiores ou inferiores.



4. PRODUTO 7 – AVALIAÇÃO DA EFETIVIDADE DO FNE

Essa seção descreve as metodologias utilizadas para avaliar a efetividade do acesso ao crédito concedido às empresas da região Nordeste, e demais estados abrangidos, por meio do FNE, em relação à indicadores socioeconômicos municipais, no período de 2006 a 2018⁷.

Inicialmente foi realizada uma análise descritiva com as variáveis utilizadas nos modelos estimados. Tal análise visou expor a situação média das variáveis relacionadas à atividade econômica e ao nível de desenvolvimento dos municípios da amostra, considerando sua variação no tempo, por unidade da federação, por grupo de tratado e de controle e por classificação quanto ao semiárido.

Para analisar a efetividade da política sobre indicadores de crescimento e desenvolvimento econômico dos municípios, propôs-se a utilização do procedimento de Diferenças em Diferenças (DD) que apresenta controle para heterogeneidade não observável.

O desenho tradicional do modelo DD inclui dois períodos de tempo onde nenhuma observação é tratada no primeiro período e parte delas é tratada no segundo período. No entanto, também poderá se trabalhar considerando múltiplos períodos de tempo (2006 a 2018) e variação no tempo de tratamento, ou seja, os municípios são expostos ao tratamento em momentos diferentes.

O termo *staggered design* ou *staggered rollout design* é utilizado para descrever esse tipo de situação, onde o tratamento é recebido em períodos diferentes e uma vez exposta ao tratamento, o município passa ao estado tratado de forma permanente. Atualmente, muitos trabalhos como Borusyak e Jaravel (2018), de Chaisemartin e D’Haultfoeuille (2020), Sun e Abraham (2021), Goodman-Bacon (2021) e Athey e Imbens (2022), entre outros, têm levantado preocupações em relação à dinâmica e à heterogeneidade dos efeitos de tratamento nesse tipo de análise, que caso presentes, levam a estimativas viesadas do efeito de tratamento. Para evitar tais problemas na análise da efetividade do FNE, serão utilizadas as contribuições de Wooldridge (2021) e Callaway e Sant’Anna (2021), visando análise de média condicional, compreendendo todo o período analisado.

⁷ A definição do período analisado se deu pela disponibilidade de dados que contemplam a análise voltada para o nível de desenvolvimento econômico dos municípios, isto é, as variáveis voltadas para o impacto, relacionadas à análise de desenvolvimento, só estariam disponíveis a partir de 2006. Sendo assim, para padronizar a discussão dos resultados e possibilitar uma melhor análise/comparação entre a atividade econômica e o nível de desenvolvimento, os modelos apresentados nessa seção irão trabalhar com o período de 2006 a 2018.

Contudo, além da análise de média condicional, é interessante observar os efeitos da política de crédito do FNE ao longo da distribuição dos indicadores dos municípios. Neste sentido, a metodologia empregada por Carvalho e Firpo (2014), para analisar o impacto de determinadas políticas públicas sobre o desempenho escolar, considera a heterogeneidade dos municípios e controla bem os possíveis problemas causados pela existência de variáveis não observadas que teriam o poder de afetar os resultados da efetividade da política desenvolvida pelo FNE. Sendo assim, será empregado o método de regressão quantílica aplicado em um conjunto de dados de painel que considere efeitos fixos. Tal como desenvolvido por Machado e Santos Silva (2019), o procedimento conterà informações dos municípios, compreendendo períodos entre os anos de 2006 a 2018, desenvolvendo análise de diferenças em diferenças e avaliando a efetividade da política nos diversos quantis de nível de atividade econômica. Isso se torna relevante quando se considera o fato de que os efeitos da política de concessão de crédito do FNE poderiam ser heterogêneos nos diferentes quantis de distribuição dos diferentes níveis de indicadores econômicos e de desenvolvimento dos municípios, por exemplo. Os resultados poderiam indicar se há diferenças na efetividade da política ao direcionar recursos para municípios no quantil inferior dos indicadores utilizados comparada com os do quantil superior.

4.1. ANÁLISE DESCRITIVA

Tendo o objetivo de descrever a situação dos municípios da amostra, no período de 2006 a 2018, foram elaboradas estatísticas descritivas focando nas seguintes variáveis municipais: Volume médio de recursos repassados pelo FNE; Produto Interno Bruto (PIB) per capita; Valores Adicionados per capita da Agricultura, da Indústria e de Serviços; Saldo de Empregos; Taxa de distorção Idade-série; Óbitos infantil, materno e por causas externas por cem mil habitantes.

A descrição trabalhou com a média das variáveis citadas acima considerando a evolução dos anos, as unidades da federação (UF), os grupos de tratado e de controle e a classificação quanto ao semiárido. Em complemento, foram apresentadas as dispersões em torno destas médias (desvio-padrão, máximo e mínimo).

Além disso, foram construídos Índices de Gini (IG)⁸ para as variáveis de Volume médio de recursos repassados pelo FNE, Produto Interno Bruto, Taxa de distorção idade-série e Óbitos infantis por mil habitantes.

O IG varia entre zero e 1 e é capaz de verificar o grau de concentração na distribuição de uma determinada variável. No caso dos volumes de recursos despendidos pelo FNE, o IG seria um indicativo de como esses recursos estariam sendo distribuídos entre os municípios da amostra. Quanto mais próximo de zero, o IG indicaria uma distribuição menos desigual do volume médio de recursos entre os municípios. Um IG igual a zero indicaria que os recursos do FNE seriam distribuídos igualmente entre os municípios da amostra, enquanto que um IG igual a 1 mostraria que um único município receberia todo o recurso disponibilizado. Inicialmente, o IG foi calculado para cada ano e considerando as unidades da federação e o total da amostra, possibilitando verificar a concentração destes recursos e sua evolução no tempo e a comparação entre os municípios das respectivas UF. Calculou-se também o IG possibilitando a comparação entre os grupos de tratado⁹ e controle e os grupos de municípios que estariam no semiárido e fora dele.

A concentração da atividade econômica também foi medida utilizando o IG. Para isso, utilizou-se o Produto Interno Bruto per capita. Foi possível avaliar a evolução da concentração da atividade econômica e comparar entre UFs, grupos de tratado e controle e dentro e fora do semiárido.

O nível de desenvolvimento econômico também foi contemplado com uma análise de concentração entre os municípios analisados. Foram utilizadas as variáveis de Óbitos Infantis por cem mil habitantes e a Taxa de Distorção Idade-Série. As análises possibilitaram o entendimento de como evoluiu a concentração da saúde e da educação nos municípios da amostra, possibilitando a comparação no tempo, entre UFs, entre tratados e controle e entre municípios dentro e fora da região semiárida.

⁸ O cálculo do Índice de Gini (IG) é algo bem difundido na literatura. De acordo com Hoffmann (2006), considerando uma distribuição discreta e ordenando as observações em ordem crescente, o IG para a variável X pode ser definido como:

$$IG = 1 - \frac{1}{n} \sum_{i=1}^n (\phi_{i-1} + \phi_i),$$
 com $\phi_i = \frac{\sum_{j=1}^i X_j}{\sum_{j=1}^n X_j}$ e, se $X_i < X_{i+1}$, ϕ_i representa a fração de X total apropriada pelos municípios com X inferior a X_i . n é o número de observações da amostra. Para maiores detalhes, ver Hoffmann (2006).

⁹ A definição do rupo de tratado e de controle será apresentada na sequência, na seção 4.2.

Esta análise descritiva, além de apresentar um quadro com a situação dos municípios pertencentes à amostra, também fornecerá subsídios para a discussão acerca dos resultados encontrados para os modelos que serão estimados para analisar a efetividade da política de concessão de créditos do FNE. Tais modelos são apresentados na sequência.

4.2. ESTIMADOR DE DIFERENÇAS EM DIFERENÇAS

Para avaliar se o FNE pode ser associado a um desempenho superior do município, quando comparado àqueles que não são beneficiários, pode-se utilizar, como são conhecidas na literatura, as técnicas de avaliação de impacto. Importante destacar que no presente estudo, tais métodos serão utilizados para analisar a efetividade de tal programa em nível de municípios.

A avaliação de efetividade visa a mensuração do efeito de tratamento ou de intervenção de algum programa público sobre algum produto ou indicador socioeconômico de interesse. Geralmente, esse tratamento é atribuído a algum grupo de interesse (tratado) e o grupo que não recebe tratamento é chamado de grupo de controle. É de interesse saber a efetividade desse tratamento sobre o grupo tratado em relação ao grupo de controle. No presente estudo, o programa de tratamento é o FNE, cujo objetivo é o relaxamento da restrição de crédito dos empreendimentos que tende a desencadear o aumento da produtividade, da geração de emprego e renda, e, por conseguinte, o crescimento e desenvolvimento econômico nos municípios beneficiados. Sendo assim, o grupo de tratamento deveria ser composto pelos municípios que receberam recursos provenientes do FNE e o grupo de controle os que não receberam. Entretanto, ao analisar a amostra, ficou constatado que os municípios, durante todo o período analisado, foram contemplados, em alguma medida, com os recursos do FNE. Sendo assim, o grupo de tratado foi definido pelos municípios que receberam um percentual de volume de recursos do FNE em proporção do Produto Interno Bruto municipal maior do que a média somada a um desvio-padrão em torno dessa média. A média foi construída considerando cada ano da amostra. O grupo de controle serão as demais unidades. Diante da inexistência de municípios que não foram contemplados com os recursos do FNE, buscou-se avaliar se aqueles que foram mais intensamente agraciados, em proporção do tamanho de sua economia municipal no respectivo ano, teriam apresentado resultados diferenciados dos que receberam

menos. Sendo assim, daqui em diante, sempre que se referir aos grupos de tratado e controle, será utilizada a definição aqui apresentada.

Quando se aborda a análise de impacto dos fundos constitucionais de financiamento, em nível de municípios, a literatura busca identificar a efetividade do programa sobre o PIB e a convergência do PIB (RESENDE (2014)). O objetivo da análise de efetividade, proposta no presente estudo, é mensurar a diferença existente nos produtos (*outcomes*) entre os grupos tratado e de controle para verificar o efeito da intervenção. Os modos mais comuns de se aferir o efeito de tratamento são o efeito médio de tratamento (*Average Treatment Effect - ATE*) e o efeito médio de tratamento sobre os tratados (*Average Treatment Effect on the Treated – ATT*).

O efeito médio de tratamento (ATE) pode ser definido como a diferença média existente nos produtos entre os grupos tratado e de controle:

$$\delta_{ATE} = E(y_1 - y_0) = E(y_1|X, D = 1) - E(y_0|X, D = 0) \quad (41)$$

em que, $E(\cdot)$ representa o operador de valor esperado ou a média condicional, y_1 e y_0 representam as variáveis de interesse, no caso dos municípios (Produto Interno Bruto-PIB, PIB da Agropecuária, PIB da Indústria, PIB do setor de serviços, PIB do setor de comércio, estoque de empregos nos municípios, valor adicionado em cada um dos três setores (agropecuárias, indústria e serviços) e saúde e educação), para os grupos tratado e de controle, respectivamente, D identifica o estado tratado (1) e não tratado (0) e X é o vetor de características dos municípios, conforme definido adiante.

Segundo Khandker, Koolwal e Samad (2010), o problema do ATE é que os grupos tratado e de controle podem não ter a mesma prioridade de intervenção, então, a diferença média observada entre os mesmos pode não ser totalmente devida ao programa. O problema ocorre devido à dificuldade de se definir o contrafactual (o que teria ocorrido com o grupo tratado caso não tivesse recebido o tratamento). Ao contrário dos experimentos controlados, nos estudos observacionais a decisão de participar do tratamento não é aleatória devido aos propósitos do programa e a auto seleção ao tratamento. Ao estimar o ATE por regressão linear, o fato de características observáveis e não observáveis estarem relacionadas à participação no programa ($D = 1$) introduzem correlação entre o termo de erro e a variável de tratamento (D). Isso viola um dos pressupostos do Modelo Clássico de Regressão Linear (MCRL) causando

viés e inconsistência nas estimativas de MQO do efeito médio de tratamento, conhecido nessa abordagem como viés de seleção ou viés de auto seleção.

Para evitar o problema, utiliza-se em estudos observacionais o segundo modo de mensurar o efeito de tratamento (ATT), definido como a diferença média observada nos produtos entre o grupo de tratados e o mesmo grupo caso não tivesse recebido tratamento:

$$\delta_{ATT} = E(y_1 - y_0|D = 1) = E(y_1|X, D = 1) - E(y_0|X, D = 1) \quad (42)$$

em que o último termo à direita representa o contrafactual. Como não é possível observar o contrafactual ele deve ser estimado, o método de estimação mais comum utilizado pelos estudos acima citados foi o de escore de propensão com pareamento (*Propensity Score Matching* – PSM).

Nesse método, o contrafactual é construído com base no grupo que não recebeu recursos/tratamento pelo pareamento com empreendimentos/municípios que receberam o recurso/tratamento. O procedimento consiste na estimação de um modelo de probabilidade (logit/probit) voltado para as chances de participar do programa (receber recursos) em função de características observáveis. A partir do modelo, obtém-se o escore ou a probabilidade de fazer parte do programa e através de algum método (kernel, vizinho mais próximo, estratificação, *mahalanobis*) ocorre o pareamento formando o contrafactual. A equação (43) define o ATT:

$$\delta_{ATT} = E(y_1 - y_0|p(X), D = 1) = E(y_1|p(X), D = 1) - E(y_0|p(X), D = 0) \quad (43)$$

em que, $p(X) = \text{prob}(D = 1|X) = E(D|X)$ é a probabilidade ou escore de propensão do município receber recursos do FNE dadas suas características.

Apesar de mais adequado que o ATE, para estudos observacionais, o ATT mensurado com a ajuda do PSM ignora a existência de heterogeneidade não observável, levando ao viés de seleção, caso presente. Dada a divergência de resultados da literatura, na presença de restrições de crédito, essa hipótese parece ser suficientemente forte para se sugerir a abordagem a partir de método mais adequado. O método de Diferenças em Diferenças (DD) para dados longitudinais seria uma alternativa mais robusta nesse caso.

O método DD assume que a heterogeneidade, que afeta a participação dos municípios no programa está presente e é invariante no tempo. Nesse caso, o ATT é estimado tomando-se as diferenças nos produtos entre os grupos tratado e de controle antes e depois da intervenção do programa. O desenho tradicional do modelo DD inclui dois períodos de tempo ($t = 1, 2$), em que nenhum município é tratado no primeiro período ($t = 1$) e parte deles é tratado no segundo período ($t = 2$). Ao fazer comparações antes e depois da intervenção e entre os municípios que receberam e não receberam tratamento, o DD produz uma estimativa mais consistente do contrafactual. O ATT estimado pelo método DD é definido na equação (44):

$$\delta_{ATT} = E[(y_{1,t=2} - y_{1,t=1}) - (y_{0,t=2} - y_{0,t=1})|X, D = 1] \quad (44)$$

em que, o primeiro termo à direita refere-se à diferença nos produtos antes ($y_{1,t=1}$) e depois do tratamento ($y_{1,t=2}$) para o grupo tratado. Este termo pode estar viesado devido a tendências no tempo como inflação, crescimento econômico, dentre outras. O segundo termo à direita usa diferenças nos produtos do grupo de controle ($y_{0,t=2} - y_{0,t=1}$) para remover a possibilidade desse viés. De forma geral, Wooldridge (2010), Callaway e Sant'Anna (2021), Roth et al. (2022), sugerem que para identificar o parâmetro de interesse δ_{ATT} são necessárias duas hipóteses básicas: a suposição de tendências paralelas (*parallel trends*) e a suposição de ausência de efeitos antecipatórios ao tratamento (*no anticipatory effect*).

A maior robustez associada ao estimador DD provém da utilização da estrutura de dados em painel. No entanto, de acordo com Borusyak e Jaravel (2018), de Chaisemartin e D'Haultfoeuille (2020), Sun e Abraham (2021), Callaway e Sant'Anna (2021), Goodman-Bacon (2021) e Athey e Imbens (2022), a utilização do modelo de efeitos fixos não é adequada para a estimação do ATT quando existem múltiplos períodos de tempo e variação no tempo de tratamento. Esse é o caso da presente análise de efetividade do FNE (2006 a 2018), em que municípios são expostos ao tratamento em momentos diferentes. O termo *staggered design* ou *staggered rollout design* é utilizado para descrever esse tipo de situação, em que o tratamento é recebido em períodos diferentes e uma vez exposto ao tratamento, o município passa ao estado tratado de forma permanente.

Conforme Wooldridge (2021), com vários períodos de tempo ($t = 1, 2 \dots T$) e diferentes momentos de tratamento ($G = \min\{t: D_t = 1\}$), existem diversos níveis de

exposição ao tratamento. Esse fato causa dinâmica e heterogeneidade nos efeitos de tratamento, tornando o modelo de efeitos fixos inadequado, uma vez que inclui apenas um parâmetro para medir o ATT. Callaway e Sant’Anna (2021) definem os níveis de tratamento pelo período de entrada no estado tratado, portanto, $(G = g)$ define o grupo de tratados exposto ao tratamento a partir do período g , em outras palavras, indica o primeiro momento em que o grupo G foi exposto ao tratamento. Caso os municípios nunca sejam tratados, a notação adotada para esse grupo é $(G = \infty)$. Nesse caso, surgem dois tipos de grupos de controle: os que ainda não foram tratados e os que nunca serão tratados. Dessa forma, o ATT pode ser definido no ano t para o grupo exposto ao tratamento em g :

$$\delta_{ATT}(g, t) = E[(y_{1,t} - y_{1,t=g-1}|X, G = g) - (y_{0,t} - y_{0,t=g-1}|X, G \in \mathcal{G})] \quad (45)$$

em que, o primeiro termo à direita refere-se à diferença nos produtos do grupo tratado no período (g) , antes de ser tratado, $(y_{1,t=g-1})$ e no período de interesse (t) em que se quer estimar o ATT após o tratamento $(y_{1,t})$. O segundo termo a direita usa diferenças nos produtos do grupo de controle $(y_{0,t} - y_{0,t=g-1})$ para os mesmos períodos. O grupo de controle $(G \in \mathcal{G})$ pode ser formado por municípios que nunca serão tratados $(G = \infty)$ ou por municípios que ainda não foram tratados $(G = g' > t)$, como em Callaway e Sant’Anna (2021), ou ainda por ambos, como em Wooldridge (2021).

De Chaisemartin e D’Haultfoeuille (2020) e Callaway e Sant’Anna (2021) mostram que é comum na literatura a utilização de alguma variação do modelo de efeitos fixos, conhecido nessa literatura como *Twoway fixed effects* (TWFE), para estimar o parâmetro δ_{ATT} do modelo DD, com vários períodos de tempo e variação no tempo de tratamento. Dos artigos empíricos publicados, entre 2010 e 2012, na *American Economic Review* (AER), 20% utilizavam esse tipo de regressão (DE CHAISEMARTIN E D’HAULTFOEUILLE, 2020).

O problema estaria relacionado à dinâmica e à heterogeneidade dos efeitos de tratamento nesse tipo de análise, que caso presentes, levariam a estimações viesadas dos efeitos promovidos pela intervenção. Goodman-Bacon (2021) e de Chaisemartin e D’Haultfoeuille (2020) demonstram que a utilização do TWFE para estimar o ATT em modelos DD, com múltiplos períodos e diferentes momentos de exposição ao tratamento, resulta em um parâmetro que é a soma ponderada de todos os recortes de DD possíveis com dois períodos. A

interpretação dos parâmetros estimados se torna complexa e os pesos podem assumir valores negativos, o que no limite, pode resultar na estimação de um impacto negativo ou não significativo, mesmo o verdadeiro ATT sendo positivo e estatisticamente significativo.

Segundo Wooldridge (2021), a conclusão de que o estimador TWFE gera resultados viesados na presença de efeitos de tratamentos dinâmicos e heterogêneos é o resultado da especificação utilizada e não do estimador em si. Para solucionar o problema propõe uma especificação flexível do estimador de efeitos fixos que permite que o ATT varie de acordo com o grupo de tratamento (g), com o período (t) e também com as variáveis de controle (X). A equação (46) apresenta tal especificação:

$$y_{it} = \alpha_i + \sum_{s=q}^T \phi_s f_{st} + \sum_{s=q}^T (f_{st} X_i) \pi_s + \sum_{g=q}^T \sum_{s=g}^T \delta_{gs} (D_{it} d_{ig} f_{st}) + \sum_{g=q}^T \sum_{s=g}^T (D_{it} d_{ig} f_{st} X_{ig}) \rho_{gs} + \varepsilon_{it} \quad (46)$$

em que α são os efeitos fixos para municípios, ϕ são os efeitos fixos para o tempo, f_s são variáveis *dummy* que indicam os anos a partir do primeiro período de tratamento ($s = q > 1$), supõe-se que haja pelo menos um período em que nenhum município é tratado. O vetor de parâmetros π capta a tendência das variáveis contextuais X_i , invariantes no tempo, sobre o produto de interesse ao longo dos anos do programa. O principal vetor de parâmetros, δ_{gs} , capta os efeitos de tratamento da intervenção sobre as municípios, o ATT varia por grupo de tratamento (g) e ao longo do tempo de tratamento (s), em que D assume valor 1, a partir do momento em que a município recebe o percentual de recursos do FNE em relação ao PIB municipal, com um desvios-padrão acima da média (que correspondente aos respectivos anos da amostra)¹⁰ e D assume valor 0, quando os municípios estão abaixo da soma da média com seu respectivo desvio-padrão, (d) identifica o grupo de tratamento e (g) ao qual o município

¹⁰ Diante da inexistência de municípios que não foram contemplados com os recursos do FNE, buscou-se avaliar se aqueles que foram mais agraciados, em proporção do tamanho de sua economia municipal, teriam apresentado resultado diferenciados dos que receberam menos. Uma vez que o município foi tratado, ele permanece nesta condição até o fim, até porque esses efeitos não seriam imediatamente observados, o que seria indicado no Event Study.

faz parte. Já o vetor de parâmetros ρ mostra o efeito das variáveis de controle sobre o efeito de tratamento (ATT), sendo $\dot{X}_{ig} = X_i - \bar{X}_g$, ou seja, o desvio com relação à média do grupo de tratamento (g). Por fim, ε é o termo de erro da regressão.

Segundo Bertrand et al. (2004) e Imbens e Wooldridge (2009), as estimações feitas pelo procedimento DD, com efeitos fixos, são tipicamente afetadas por problemas de autocorrelação devido à natureza das variáveis dependentes que, juntamente com o problema de heterocedasticidade, tornam as estimativas dos erros-padrão inválidas. Nesse sentido, Cameron e Trivedi (2005 e 2009) sugerem que para painéis curtos, especialmente, seja feita a correção robusta por *cluster* para que a verdadeira matriz de variância e covariância dos coeficientes seja estimada, corrigindo os problemas.

A solução apresentada por Wooldridge (2021) altera apenas a especificação do modelo TWFE que estima o ATT por meio do desenho de Diferenças em Diferenças. No entanto, vários estimadores alternativos que permitem maior dinâmica e heterogeneidade do ATT foram propostos recentemente, dentre eles estão os de Chaisemartin e D'Haultfoeuille (2020), Sun e Abraham (2021), Callaway e Sant'Anna (2021) e Athey e Imbens (2022). O estimador proposto por Callaway e Sant'Anna (2021) possui maior aplicação para a avaliação dos impactos do FNE em relação aos demais, pois permite a inclusão de variáveis de controle e possui condições de identificação mais flexíveis.

Callaway e Sant'Anna (2021) propõe um estimador (CS)¹¹ que combina as técnicas de ponderação pelo inverso da probabilidade (*Inverse Probability Weighting – IPW*) com o ajuste de regressão. Ao utilizar ambas as técnicas o estimador possui a propriedade de ser duplamente robusto, ou seja, se uma das funções estiver corretamente especificada, as estimações do ATT serão não viesadas. O estimador de Callaway e Sant'Anna (2021) pode ser definido como:

$$ATT(g, t) = E \left[\left(\frac{d_{ig}}{E[d_{ig}]} - \frac{\frac{\hat{p}(X)C}{1 - \hat{p}(X)}}{E \left[\frac{\hat{p}(X)C}{1 - \hat{p}(X)} \right]} \right) (y_{1,t} - y_{1,t=g-1} - E[y_{0,t} - y_{0,t=g-1} | X, C = 1]) \right] \quad (47)$$

¹¹ Em referência aos autores Callaway e Sant'Anna (2021).

em que, d_{ig} é uma variável *dummy* que assume valor igual a um se o município fizer parte do grupo tratado no período g , $\hat{p}(X)$ é a probabilidade ou escore de propensão do grupo g receber tratamento. O termo $(y_{1,t} - y_{1,t=g-1})$ é a diferença dos produtos para o grupo tratado em g e o termo $(y_{0,t} - y_{0,t=g-1})$ é a diferença dos produtos para o grupo de controle, nesse caso, composto por município ainda não tratados ou que não serão tratados em nenhum momento. C é definida como uma variável binária que assume valor igual a 1 para as unidades que não participam do tratamento em nenhum período de tempo.

O procedimento de estimação consiste em três etapas. A primeira etapa consiste em calcular as médias dos produtos antes e depois do tratamento para os grupos tratado e de controle. A segunda etapa é estimar o escore de propensão para calcular os pesos para a ponderação, o que adiciona a exigência da suposição de suporte comum. Já a terceira etapa calcula o ATT, a partir das diferenças ponderadas entre os grupos tratado e de controle.

O estimador duplamente robusto de Callaway e Sant'Anna (2021) concentra-se na estimação individual do ATT. No entanto, é possível fazer várias agregações para cada grupo de tratamento g e ao longo do tempo, assim como na especificação TWFE apresentada na equação (6). Para calcular erros padrão-robustos são utilizados procedimentos *bootstrap* que garantem, assintoticamente, uma inferência válida em relação à autocorrelação e *clustering*.

Atualmente, a estimação dos efeitos de tratamento ATT pelo método DD são motivo de significativa discussão na literatura. Vários estimadores foram propostos para superar as limitações levantadas em relação à regressão de efeitos fixos, principalmente relacionadas à possível dinâmica e heterogeneidade do ATT, em situações de múltiplos períodos de tempo e variação no tempo de tratamento. Segundo Woodridge (2021), as limitações levantadas são originadas da utilização de especificação inadequada, afirmando ainda que a regressão de efeitos fixos possui propriedades bem conhecidas em painéis desbalanceados e é eficiente em relação ao estimador CS, pois usa toda a informação disponível, grupos de controle nunca tratados e ainda não tratados, para a estimação do ATT. Opta-se, nesse caso, por utilizar as contribuições de Wooldridge (2021) e Callaway e Sant'Anna (2021) para estimar os efeitos de tratamento do FNE.

Um problema inerente do modelo DD refere-se à pressuposição de tendência comum entre os grupos de controle e tratado. Esta pressuposição assegura que não é necessária a

existência de uma região de suporte comum de propensão a participar do tratamento para estimação do impacto econômico. No entanto, ela exige que as tendências dos produtos dos grupos sejam iguais caso não houvesse tratamento. A pressuposição garante que o resultado encontrado se deve ao tratamento (GERTLER et al., 2011). Não é possível verificar formalmente a pressuposição, contudo, a literatura sugere que sejam feitos testes de robustez (CUNNINGHAM, 2021; HUNTINGTON-KLEIN, 2022). Um dos testes é a estimação de um *event study*, que inclui modelos em que o tratamento ocorreria em outros pontos e não no ponto onde realmente ocorre o tratamento, ou seja, incluem-se avanços e defasagens (*leads and lags*) nos modelos para verificar se existe causalidade entre o momento do tratamento e a resposta em termos de impacto.

O modelo analítico usado é definido pela equação (46), que define os determinantes do produto y no município i no ano t , podendo ser avaliada com base no seguinte modelo econométrico:

$$\ln(y_{it}) = \alpha_i + \sum_{s=q}^T \phi_s fs_t + \sum_{s=q}^T (fs_t X_i) \pi_s + \sum_{g=q}^T \sum_{s=g}^T \delta_{gs} (D_{it} d_{ig} fs_t) + \sum_{g=q}^T \sum_{s=g}^T (D_{it} d_{ig} fs_t X_{ig}) \rho_{gs} + \varepsilon_{it} \quad (48)$$

A variável y_{it} representa os produtos no município i , no ano t , aos quais tem-se interesse em verificar a efetividade do programa de concessão de crédito vinculado ao FNE, quais sejam: (Produto Interno Bruto -PIB, Valores adicionados da Agropecuária, Indústria e Setor de Serviços, Saldo de Empregos per capita, variáveis relacionadas à Saúde-óbitos maternos, infantis e por causas externas, todos por mil habitantes- e a variável de Distorção Idade Séria, relacionada à Educação nos municípios). Utiliza-se a transformação logarítmica na variável dependente devido a seus bons atributos econométricos, como a compressão das unidades de dispersão com relação à média, maior propensão à normalidade e interpretação de semi-elasticidade. O período t compreende os anos entre 2006 e 2018.

Os parâmetros ϕ_s e α_i representam os efeitos fixos invariantes no tempo para os municípios e os efeitos fixos invariantes nos municípios para os anos, respectivamente, que

controlam a heterogeneidade não observada no tempo e entre municípios. O vetor δ_{gs} representa o efeito que a participação no programa D exerce sobre a variável dependente y . A participação no programa D é dividida de acordo com o programa e a finalidade do crédito (capital de giro, custeio ou investimento), sendo variáveis dicotômicas. O efeito de tratamento pode ser calculado por grupo de tratamento e ao longo do tempo, permitindo verificar sua dinâmica.

Os vetores π e ρ representam o efeito que cada uma das variáveis observadas X exerce sobre o produto y e sobre o impacto do tratamento δ ao longo do tempo. Para compor o conjunto de características observáveis dos municípios (X_i), foram selecionadas variáveis de controle utilizadas em diversos estudos empíricos nessa vertente de pesquisa [Durlauf e Quah (1999) e Durlauf, Johnson e Temple (2004), Tsangarides (2005)], e que podem afetar as variáveis de resultado, bem como o recebimento do benefício.

Finalmente, ε é o termo de erro da regressão (equação (8)). Para calcular o ATT dos modelos DD serão utilizados os estimadores de mínimos quadrados ponderados com efeitos fixos e correção robusta por *cluster* para heterocedasticidade e autocorrelação e, o estimador CS de Callaway e Sant'Anna (2021) com correção robusta por *bootstrap* para autocorrelação e *clustering*. Para estimar o escore de propensão são utilizadas as variáveis descritas no vetor X .

4.3. REGRESSÃO QUANTÍLICA

Como se sabe, os estimadores de MQO são estimadores de médias condicionais e, como tal, possibilitam interpretar o comportamento médio de um evento estudado, que no caso é o comportamento dos índices utilizados para avaliar o impacto da política nos resultados socioeconômicos dos municípios. Sendo assim, as estimações por MQO não possibilitam investigar qual seria o impacto dos empréstimos do FNE ao longo da distribuição da atividade econômica municipal.

Diante disso, nesta seção é proposta uma metodologia de avaliação da efetividade baseada na técnica de Diferenças em Diferenças (DD), já comentada na seção anterior. Contudo, a análise irá focar em efeitos diferenciados da política ao longo da distribuição dos indicadores dos municípios, considerando um grupo de tratados e os não tratados em cada

quantil da distribuição desses indicadores. Tal análise ao longo da distribuição será trabalhada tendo como base regressões quantílicas aplicadas em dados em painel.

Segundo Meyer et. al. (1995), os estimadores de regressão quantílica reduziriam drasticamente a importância de *outliers* e permitiriam relaxar as pressuposições sobre as formas funcionais utilizadas na análise. Canay (2011) também ressalta que os modelos de regressão quantílica possibilitam o controle do efeito de variáveis não observadas e de covariáveis heterogêneas, enquanto que Santos (2012) deixa claro que outra vantagem seria o fato de permitir o relaxamento da hipótese do modelo clássico de regressão linear populacional de que os erros se distribuem normalmente.

Considerando tais vantagens, Santos (2012), define o quantil de ordem τ de uma população ou amostra como sendo o valor m , tal que $100\tau\%$ dos valores são inferiores a ele, com $0 < \tau < 1$. Seja Y com f.d.p. acumulada F^{12} , o quantil de ordem τ também pode ser definido como sendo o valor m que minimiza $E(Y - m)$, ou seja, aquele que minimiza os valores que estão acima (fora) do valor que define o quantil.

Esta ideia, de acordo com Koenker (2005), conduz ao conceito de função de perda, $\rho_\tau(\mu)$, que teria como principal objetivo minimizar a perda esperada em cada quantil possibilitando, dessa forma, trabalhar com o conjunto completo de informações dos indicadores dos municípios considerando toda sua distribuição, mediante a ponderação pelos quantis.

Sendo assim, dada uma amostra de n observações dos indicadores de municípios (Y), o quantil amostral de ordem τ resolveria o problema de minimização apresentado na equação (9), exposto por Koenker (2005), sendo abordado através de métodos de programação linear (simplex, por exemplo), uma vez que os erros, neste caso, não zeram na média:

$$\min_{q \in \mathbb{R}} \sum_{i=1}^n \rho_\tau(Y_i - q) \quad (49)$$

¹² Admitindo uma população, pode-se especificar a função densidade de probabilidade (f.d.p.) acumulada para a variável aleatória Y (indicadores dos municípios) como: $F(Y) = P(Y < y)$. Definindo a inversa da f.d.p. acumulada no ponto τ como: $F^{-1}(\tau) = \inf\{x: F(y) \geq \tau\}$

A intenção é especificar o quantil condicional de Y , dado X , sendo que X são as variáveis independentes do modelo e $\beta(\tau)$ é o vetor de parâmetros definido para cada quantil, como função linear nos parâmetros da forma:

$$Q_{\tau}(Y/X) = X\beta(\tau) \quad (50)$$

Bastaria definir um vetor $\hat{\beta}(\tau)$ que seja a solução do problema de minimização:

$$\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n \rho_{\tau}(Y_i - X_i\beta) \quad (61)$$

Minimizando a soma do erro ponderado pela perda gerada por cada quantil τ , tem-se:

$$Y_i = \beta_0(\tau) + \beta_1(\tau)X_{1i} + \dots + \beta_p(\tau)X_{pi} + u_i \quad (62)$$

em que, u_i são erros aleatórios identicamente distribuídos.

$$Q_{\tau}(Y/x) = \beta_0(\tau) + \beta_1(\tau)X_{1i} + \dots + \beta_p(\tau)X_{pi} \quad (63)$$

Deve-se destacar que β deve ser indexado a τ , pois um dos interesses, neste caso, é exatamente estudar se o vetor β assumiria diferentes valores para τ 's diferentes. Por considerar diferentes pontos de distribuição da variável condicional Y , esses modelos ajudam a obter uma visão mais completa da relação entre as variáveis estudadas.

De acordo com Canay (2011), a utilização de estimadores de regressão quantílica com um painel, permite, além das vantagens citadas na seção anterior, inserir efeitos fixos para controlar os efeitos de covariáveis não observadas. No presente trabalho, será utilizado o Método dos Momentos – Regressão Quantílica (MM-RQ) apresentado por Machado e Silva (2019). Este método para estimação dos parâmetros ao longo da distribuição da variável dependente permite que os efeitos atuem como deslocadores de localização, mas também possibilitam mensurar os efeitos individuais ao longo de toda a distribuição.

Machado e Silva (2019) indicam que, diante de determinadas pressuposições, a relação entre o fenômeno estudado e seus argumentos poderia ser definida como:

$$Y_{it} = \alpha_i + X'_{it}\beta + (\delta_i + Z'_{it}\gamma)U_{it} \quad (64)$$

Na equação (64), Y_{it} representa os indicadores de interesse dos municípios. As variáveis explicativas estariam na matriz de dados X'_{it} . Z'_{it} é um vetor de tamanho k de maneira que seria construído através de transformações em X'_{it} e teria probabilidade igual a 1¹³. O termo de erro U_{it} seria independente e identicamente distribuído (i.i.d.), normalizado para atender as condições de momento impostas por Machado e Silva (2019). Diante disso, a equação (64) pode ser escrita como:

$$Q_Y(\tau|X_{it}) = (\alpha_i + \delta_i q(\tau)) + X'_{it}\beta + Z'_{it}\gamma q(\tau) \quad (65)$$

Na equação (15) o efeito fixo no quantil τ é definido como $\alpha_i(\tau) \equiv (\alpha_i + \delta_i q(\tau))$, também chamado de efeito distributivo em τ . Gouveia (2020) indica o impacto de efeitos fixos, que seriam distintos ao longo da distribuição condicional de Y . O tempo é representado pelo subscrito $t = 0, 1$ e os municípios são representadas pelo subscrito $i = 1, 2, \dots, n$.

Nas análises de efetividade para os municípios, respectivamente, serão separados em grupos de tratados e os não tratados. Entretanto, como mencionado anteriormente, a efetividade pode ter apresentado efeitos distintos ao longo da distribuição dos indicadores considerados. Por exemplo, os empréstimos podem ter tido maior impacto positivo sobre o PIB em municípios tratados, quando comparados aos não tratados, quando é trabalhado em municípios que estão concentrados na parte inicial da distribuição do PIB, ou seja, nas de menor PIB. Isto poderia ser um indicador de que o impacto seria maior quando a política se concentrasse nesses municípios.

Autores como Meyer et. al. (1995) e Athey e Imbens (2006) apresentam estimadores de diferenças em diferenças (DD) quando se trabalha com regressões quantílicas. Carvalho e Firpo (2014) enfatizam que, quando se trabalha com DD em regressões quantílicas, as pressuposições

¹³ Machado e Silva (2019) apresentam as condições para a transformação no vetor Z .

seriam mais fortes. Seria necessária realizar a imposição de que, considerando as variáveis de controle, as variáveis não observáveis deveriam ter a mesma distribuição para os grupos de controle e de tratamento em qualquer período no tempo, independente do período de adoção ou não da política analisada.

Considerando a inserção da política no painel quantílico, a equação (15) seria definida como:

$$Q_Y(\tau|X_{it}) = (\alpha_i + \delta_i q(\tau)) + \theta_\tau D_{it} + X'_{it}\beta + Z'_{it}\gamma q(\tau) \quad (66)$$

D_{it} é uma variável *dummy* definida igual a 1 caso o município seja considerado não tratado no período $t=0$ e passe a ser tratado no período $t=1$. D_{it} assumira valor zero caso o município esteja no grupo de não tratados em ambos os períodos. O parâmetro θ_τ seria responsável por avaliar se houve impacto diferenciado nos municípios tratados e de controle. Considerando a mediana ($\theta_{\tau=0,5}$), tem-se:

$$\theta_{\tau=0,5} = E[(Y_{1,t=1} - Y_{1,t=0}) - (Y_{0,t=1} - Y_{0,t=0}) | X'_{it}, Z'_{it}, D_{i,t=1} = 1] \quad (67)$$

Observa-se que o primeiro termo da equação (67) apresenta o diferencial nos índices dos municípios que receberam o tratamento no segundo momento (tratados) e o segundo termo é a diferença de desempenho do grupo de controle, ou seja, a diferença entre os períodos analisados dos índices dos municípios que não receberam empréstimos do FNE na intensidade desejada. $\theta_{\tau=0,5} > 0$ seria o indicativo de que a política de crédito do FNE, a partir de uma intensidade pré-determinada, concedido aos municípios localizados na mediana da distribuição dos indicadores utilizados gerou impactos positivos.

Para realizar a análise, serão avaliados os efeitos da política de concessão de crédito sobre indicadores socioeconômicos dos municípios da região atendida, ou seja, o impacto tende a ir além do que a política implementada inicialmente objetivava. Como citado anteriormente, no período 0, nenhum dos municípios devem receber o tratamento. No período 1, a amostra será dividida entre municípios tratados, e municípios de controle.

Sendo assim, as estimações buscarão avaliar como a política de crédito afetou o PIB *per capita*, o Saldo de Empregos, os Valores Adicionados da Agropecuária, Indústria e

Serviços, além das variáveis relacionadas à qualidade de vida, saúde e educação dos municípios, óbitos maternos, infantis e por causas externas e Distorção Idade Séria, respectivamente, (Y_{it}). O modelo é apresentado na equação (68).

$$\begin{aligned}
 Y_{it} = & \alpha_{it} + \theta_{\tau}^M D_{it}^M + \beta_{1\tau} FNE_pc_{it} + \beta_{2\tau} PRIMARIO_{it} + \beta_{3\tau} SECUNDARIO_{it} \\
 & + \beta_{4\tau} TECIARIO_{it} + \beta_{5\tau} CARENIA_{it} + \beta_{6\tau} TX_JUROS_{it} \\
 & + \beta_{7\tau} RRENDIMENTO_{it} + \beta_{8\tau} SEM_{it} + \beta_{9i} fT_{it} + u_{it}
 \end{aligned} \quad (68)$$

Na equação (68), Y_{it} representa o PIB *per capita* municipal a preços constantes, o valor adicionado *per capita* do setor agropecuário, o valor adicionado *per capita* do setor industrial, o valor adicionado *per capita* do setor de serviços e o Saldo de empregos nos municípios e como proxy de saúde, óbitos infantis, maternos e por causas externas, e como proxy de educação, foi utilizada a variável de distorção idade-série do Ensino Médio, totalizando nove modelos. Cada modelo irá considerar a distribuição de Y_{it} nos quantis $\tau = 0,10; 0,25; 0,50; 0,75; 0,90$. A variável de tratamento é representada por D_{it}^M , sendo que para os tratados $D_{it}^M = 0$ se o município não recebeu proporção maior que a média no respectivo ano (Proporção de recursos do FNE Total em relação ao PIB), mais 1 desvio padrão, em $t = 0$ e $D_{it}^M = 1$ a partir do momento que o município recebeu proporção maior que a média, mais 1 desvio padrão, em $t = 1$, passando assim a ser tratado em todos os demais anos até o final do painel. No caso do grupo de controle, $D_{it}^M = 0$ considerando que em nenhum dos períodos, o município foi contemplado com o crédito destinado ao setor. Como exemplo, considerando o setor agrícola, um $\theta_{\tau}^M > 0$ indica que os financiados estariam gerando um impacto positivo quando comparado com os municípios que não receberam os recursos.

4.4. FONTE E TRATAMENTO DOS DADOS

Para a realização da análise de efetividade em nível de município, foram utilizados dados extraídos de diferentes fontes. As informações referentes ao FNE (valor contratado, taxa de juros e prazo de carência) e à RAIS (rendimento médio) foram disponibilizadas pela equipe técnica da Sudene em nível individual. Para a construção da base de dados utilizada na presente investigação, os dados supracitados foram agregados em nível de município, no intuito de se

se obter as variáveis efetivamente utilizadas: (i) valor total contratado do FNE per capita, (ii) taxa de juros média, (iii) prazo de carência médio e (iv) média do rendimento médio.

Com exceção do estoque de empregos, o qual foi retirado da base de dados disponibilizada pela equipe técnica da Sudene, a partir da extração de dados da RAIS, as variáveis de resultado relacionadas ao crescimento econômico municipal foram obtidas junto ao IBGE (2022). De modo específico, o PIB e os VABs setoriais foram extraídos do Sistema IBGE de Recuperação Automática (SIDRA), sendo posteriormente divididos pela população municipal para a obtenção das variáveis efetivamente utilizadas: (i) PIB per capita, (ii) VAB per capita da agropecuária, (iii) VAB per capita da Indústria e (iv) VAB per capita do setor de serviços.

No tocante às variáveis de resultado relacionadas ao desenvolvimento econômico municipal, as mesmas foram extraídas das bases de dados do Ministério da Saúde e do Ministério da Educação. Em específico, as seguintes variáveis foram obtidas do site do Departamento de Informática do Sistema Único de Saúde (DATASUS, 2022): (i) total de óbitos infantis, (ii) total de óbitos maternos e (iii) o total de óbitos por causas externas. A variável de distorção idade-série no ensino médio, por seu turno, foi retirada do site do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep, 2022).

Por fim, tem-se a variável dicotômica utilizada para indicar se o município pertencia, a cada ano analisado, à delimitação legal do semiárido. Para construção dessa variável foram consideradas as legislações específicas sobre o tema, quais sejam: a Portaria Nº 1.182/1999 da Sudene (BRASIL, 1999), para classificação entre 2000 e 2004; a Portaria Nº 89/2005 do Ministério da Integração Nacional, (BRASIL 2005) para classificação entre 2005 e 2016; e a Resolução Nº 115/2017 do Conselho Deliberativo da Sudene (BRASIL, 2017), para classificação entre 2017 e 2020.

5. PRODUTO 8 – AVALIAÇÃO DO RETORNO ECONÔMICO E SOCIAL DO FNE

Tendo em vista o objetivo de se avaliar o retorno econômico e social da aplicação dos recursos do Fundo Constitucional de Financiamento do Nordeste (FNE) por meio das abordagens por insumo-produto (IP) e por equilíbrio geral computável (EGC), essa seção apresenta as bases dos modelos empregados na avaliação, além da fonte de dados utilizadas e procedimentos adotados.

5.1. ANÁLISE DE INSUMO-PRODUTO

A estrutura básica do modelo refere-se aos fluxos monetários de venda de produtos de cada setor produtor para cada setor consumidor e da demanda final pertinente. Além disso, retratam-se as despesas de importação, pagamento de impostos e remuneração de fatores de produção de cada atividade econômica. Logo, o elemento central nesse modelo é uma tabela de transações intersetoriais, como exemplificado na Figura 4. As linhas de setores mostram como eles distribuem (vendem) suas produções em toda economia e as colunas demonstram os requerimentos (compras) de insumos para composição do valor do produto final.

Figura 4 – Matriz Básica de Insumo-Produto

		Setores Compradores				Demanda Final				Valor Bruto da Produção X_i
		Setor A	Setor B	Setor C	Subtotal	Investimento (I)	Gasto do Governo (G)	Exportações (E)	Consumo das Famílias (C)	
setores Vendedores	Setor A									
	Setor B									
	Setor C									
	Subtotal 1									
Importações										
Tributos Indiretos Líquidos										
Valor Adicionado										
Valor Bruto da Produção X_j										

Fonte: elaborado pelos autores com base em Miller e Blair (2009)

De acordo com a matriz apresentada, as vendas de um setor são usadas como insumo para a produção de outro setor, ou podem ser utilizadas por qualquer outro componente da demanda final. Por outro lado, nas colunas percebe-se que o processo produtivo de um setor requer uma determinada quantidade de insumos, que podem ser produzidos pela própria economia daquela região ou importado, além de pagamento dos impostos e remuneração de fatores de produção (salários, lucros, aluguéis e juros) representado pelo valor adicionado.

Desse modo, a construção matemática de um sistema de insumo-produto para uma região consiste em um conjunto de n equações lineares com n incógnitas. A demanda de um setor qualquer denominado i por insumos oriundos de outros setores, é relacionada diretamente com a quantidade de bens produzidos por esse mesmo setor i para atender a demanda final: famílias (**C**), governo (**G**), investimento (**I**) e exportações (**E**).

Assume-se, então, que a economia é dividida em n setores:

$$x_i = z_{i1} + \dots + z_{ij} + \dots + z_{in} + Y_i = \sum_{j=1}^n z_{ij} + Y_i \quad (69)$$

em que, x_i representa a produção total do setor i , com i variando de 1 até n ; z_{ij} é o fluxo monetário do setor i para o setor j (índice alternativo para outro setor); Y_i é a demanda final total pelo produto do setor i .

Em seguida, pode-se estimar os coeficientes técnicos (a_{ij}), que representam a relação fixa dos fluxos interindustriais do setor i para o setor j , da seguinte forma:

$$a_{ij} = \frac{z_{ij}}{x_j} \quad (70)$$

Assume-se então, por restrição imposta pela base de dados, a hipótese de retornos constantes de escala, isto é, em qualquer nível de produção, a tecnologia (proporção de compra dos “insumos”) utilizada por todos os setores será a mesma para um certo conjunto de dados coletados em um período. Mas (70) pode tornar-se:

$$z_{nj} = a_{nj}x_j \quad (71)$$

Substituindo (71) em (69), tem-se que:

$$x_n = a_{n1}x_1 + \dots + a_{ni}x_i + \dots + a_{nn}x_n + Y_n \quad (72)$$

É possível representar o sistema acima compactamente de maneira matricial:

$$x = Ax + Y \quad (73)$$

que pode ser representado também da seguinte forma:

$$(I - A)x = Y \quad (74)$$

ou, alternativamente:

$$x = (I - A)^{-1}Y \quad (75)$$

Por fim, por (75), tem-se a matriz inversa de Leontief “ $(I - A)^{-1}$ ”, representando a tabela de requerimentos totais, ou dos efeitos diretos e indiretos das modificações exógenas da demanda final sobre a produção dos n setores. Logo, cada elemento da matriz contém as condições diretas e indiretas da produção total do setor i para a produção de uma unidade de demanda final do setor j .

Ademais, um desenvolvimento natural do modelo base de insumo-produto é a abordagem inter-regional. Nesse caso, além das representações típicas intersetoriais, o sistema econômico da matriz contém os fluxos inter-regionais de bens e serviços, demanda final e renda entre duas ou mais regiões com cada uma tendo dois ou mais setores ao mesmo tempo. A Figura 5 abaixo apresenta esse tipo de matriz de maneira estilizada.

Figura 5 – Matriz Básica de Insumo-Produto Inter-regional

Setor de vendas	Setor de compras					
	Região L			Região M		
	1	2	3	1	2	
Região L	1	z_{11}^{LL}	z_{12}^{LL}	z_{13}^{LL}	z_{11}^{LM}	z_{12}^{LM}
	2	z_{21}^{LL}	z_{22}^{LL}	z_{23}^{LL}	z_{21}^{LM}	z_{22}^{LM}
	3	z_{31}^{LL}	z_{32}^{LL}	z_{33}^{LL}	z_{31}^{LM}	z_{32}^{LM}
Região M	1	z_{11}^{ML}	z_{12}^{ML}	z_{13}^{ML}	z_{11}^{MM}	z_{12}^{MM}
	2	z_{21}^{ML}	z_{22}^{ML}	z_{23}^{ML}	z_{21}^{MM}	z_{22}^{MM}

Fonte: elaborado pelos autores com base em Miller e Blair (2009)

Na tabela apresentada na Figura 5, temos que z^{LL} representa o fluxo monetário do setor i para o setor j dentro da região L ; z^{LM} representa o fluxo monetário do setor i da região L para o setor j da região M ; z^{MM} representa o fluxo monetário do setor i para o setor j dentro da região M ; z^{ML} representa o fluxo monetário do setor i da região M para o setor j da região L .

Para encontrar a produção total do setor 1 na região L , faz-se:

$$X_1^L = z_{11}^{LL} + z_{12}^{LL} + z_{13}^{LL} + z_{11}^{LM} + z_{12}^{LM} + Y_1^r \quad (76)$$

em Y_1^r são as vendas do setor 1 para a demanda final.

Como no modelo base, a abordagem inter-regional assume a hipótese de que os coeficientes técnicos são fixos, ou seja, a economia opera com retornos constantes de escala. Os coeficientes intrarregionais são:

$$a_{ij}^{LL} = \frac{z_{ij}^{LL}}{X_i^L} \quad a_{ij}^{MM} = \frac{z_{ij}^{MM}}{X_j^M} \quad (77)$$

em que, a_{ij}^{LL} estima quanto o setor j na região L compra do setor i e da região L ; a_{ij}^{MM} estima quanto o setor j na região M compra do setor i e da região M .

Quanto aos coeficientes inter-regionais tem-se que:

$$a_{ij}^{LM} = \frac{z_{ij}^{LM}}{X_i^M} \qquad a_{ij}^{ML} = \frac{z_{ij}^{ML}}{X_i^L} \qquad (78)$$

em que a_{ij}^{LM} estima quanto o setor j na região M compra do setor i e da região L ; a_{ij}^{ML} estima quanto o setor j na região L compra do setor i e da região M .

Em seguida, substitui-se em (76) os coeficientes encontrados em (77) e (78),

$$X_1^L = a_{11}^{LL}X_1^L + a_{12}^{LL}X_2^L + a_{13}^{LL}X_3^L + a_{11}^{LM}X_1^M + a_{12}^{LM}X_2^M + Y_1^L \qquad (79)$$

Na forma matricial, o impacto na produção regional de uma mudança na demanda final da região r será encontrado da seguinte forma:

$$X^r = (I - A^{rr})^{-1}Y^r \qquad (80)$$

A modelagem econômica por Insumo-Produto permite também gerar indicadores e estimação de efeitos na economia a partir de choques exógenos no modelo. Isso é feito por meio dos chamados multiplicadores econômicos, particularmente os de produção, renda e emprego. Nesse sentido, a tabela de requerimentos totais (isto é, a matriz inversa de Leontief derivada na seção anterior) mostra o efeito direto, indireto, e as mudanças induzidas na produção, na renda e no emprego da economia requeridas para atender a um aumento da demanda final, isto é, estimar efeitos na economia a partir de choques exógenos no modelo.

Sendo assim, o multiplicador de produção é definido da seguinte forma:

$$MP_j = \sum_{i=1}^n b_{ij} \qquad (81)$$

com $j = 1, \dots, n$ em que MP_j é o multiplicador de produção do j -ésimo setor e b_{ij} é um elemento qualquer da matriz inversa de Leontief (denominada B). De novo, o valor calculado representa o valor total da produção de toda a economia que é acionado para atender a variação de uma unidade na demanda final do setor j .

Quando esse efeito é limitado à demanda de insumos intermediários, eles são chamados de multiplicadores do tipo I. Contudo, se a demanda das famílias é endogenizada no sistema, tem-se um efeito adicional na economia denominado “efeito induzido”. Tais multiplicadores recebem o nome de “tipo II”.

Para o multiplicador tipo II, calcula-se a matriz inversa de Leontief $\bar{B} = (I - \bar{A})^{-1}$, que é obtida a partir de uma matriz \bar{A} de coeficientes técnicos, onde o consumo das famílias é endogeneizado. Com esta \bar{B} obtém-se os seguintes multiplicadores de produção setoriais do tipo II:

$$\overline{MP}_j = \sum_{i=1}^n \bar{B}_{ij} \quad (82)$$

com $j = 1, \dots, n$ em que \overline{MP}_j é o multiplicador de produção do j -ésimo setor do tipo II. De novo, o valor calculado representa o valor total da produção de toda a economia que é acionado para atender a variação de uma unidade na demanda final do setor j , porém, agora considerando os efeitos diretos, indiretos e induzidos pela renda das famílias.

Especificamente sobre o multiplicador de renda (ou salários), o multiplicador do tipo I informa o aumento direto e indireto na renda gerado na economia para um aumento da demanda final. Em paralelo, o multiplicador de renda do tipo II fornece o aumento na renda gerados na economia dado uma elevação na demanda final, incorporando efeitos diretos, indiretos e induzidos (pelos salários das famílias). Os cálculos se dão da seguinte forma:

$$C_i^R = R_i / VBP_i \quad (83)$$

em que C_i^R é o coeficiente de renda do setor i ; R_i é a renda (estimada pelo Salários) no setor i ; e VBP_i é o valor bruto da produção do setor i . De posse desses coeficientes, obtém-se os multiplicadores de renda diretos e indiretos do tipo I na economia com a seguinte multiplicação:

$$MR_{1 \times n} = C_{1 \times n}^R (I - A)_{n \times n}^{-1} \quad (84)$$

Para encontrar os multiplicadores de renda do tipo II basta usar $(I - A)^{-1}$ na fórmula acima, que tem os rendimentos das famílias endogenizados. Logo, os valores obtidos para os coeficientes de geração de renda levarão em conta, além do efeito direto e indireto, o efeito induzido. Este efeito adicional corresponde à geração de renda ocasionado pelo aumento do consumo das famílias.

De resto, destaca-se que optou-se neste relatório por apresentar somente a média dos multiplicadores de produção e renda do Tipo I e II. Além da economia considerável de espaço, tornando a visualização e interpretação dos resultados muito mais amigável, normalmente entende-se que, em certa medida, os multiplicadores do Tipo I possivelmente subestimam os impactos econômicos dado que a reação endógena das famílias é ignorada. Em contrapartida, os multiplicadores do Tipo II provavelmente superestimam os impactos econômicos por causa das hipóteses fortes implícitas acerca dos salários e padrões de consumo no processo de endogeneização no consumo das famílias no modelo. Isto é, não é tão razoável supor que as famílias permaneçam consumindo no mesmo patamar e com padrão similar, após aumentos sucessivos dos seus níveis de renda advindos dos choques exógenos no modelo. Sendo assim, essas duas maneiras de calcular os multiplicadores (sem e com famílias endógenas) podem ser consideradas os limites superiores e inferiores dos intervalos onde estarão os verdadeiros efeitos diretos e indiretos do aumento na demanda final.

Na modelagem por insumo-produto o aumento da produção de um determinado setor tem dois tipos de impactos econômicos em outros setores da economia. O primeiro é aquele que corresponde ao aumento da demanda de insumo em função da necessidade de aumentar a produção. Esse é a ligação para trás (*backward linkage effects* ou poder de dispersão) donde se extraem os índices de ligação para trás. Por outro lado, o consequente aumento da produção do setor significa que quantidades adicionais da sua produção estará disponível para uso como insumos de outros setores da economia para seus processos produtivos. Nesse caso, temos a interconexão para frente (*forward linkage effects* ou sensibilidade de dispersão) e as estimativas dos índices de ligação para frente

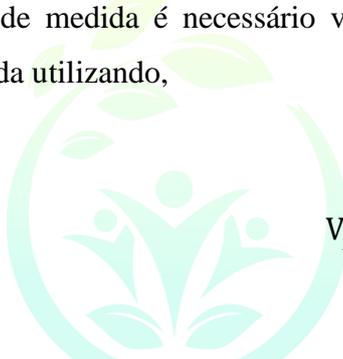
Nesse sentido, os índices de ligação são usados para estabelecer os setores que teriam o maior poder de encadeamento dentro da economia. Assim, setores com elevado grau de encadeamento junto à cadeia produtiva, com efeitos para trás e para frente acima da média, são considerados setores-chave. Logo, um setor-chave é aquele que demanda insumos dos demais

setores em uma quantidade superior que a média e cuja produção é largamente utilizada pelos demais setores.

O método utilizado para a identificação desses setores foi desenvolvido por Rasmussen e é baseado na matriz inversa de Leontief. Em um sistema contendo n setores, no qual a inversa de Leontief é uma matriz \mathbf{b} ($n \times n$), tem-se,

$$R_j = \sum_{i=1}^n b_{ij} \quad (j = 1, 2, \dots, n) \quad (85)$$

em que \mathbf{R}_j simboliza a soma dos elementos da coluna j da inversa de Leontief, ou seja, o efeito no produto de toda a economia provocada por um aumento de uma unidade na demanda final do setor j . Assim, cada elemento da coluna j propicia o impacto direto e indireto de um acréscimo de uma unidade da demanda final pela produção do setor j em cada um dos n setores. Para fins de medida é necessário verificar a magnitude relativa de \mathbf{R}_j , em que pode ser normalizada utilizando,



$$V_j = \frac{\frac{1}{n} R_j}{\frac{1}{n^2} \sum_{i,j} b_{ij}} \quad (86)$$

no qual o numerador é a média dos elementos da coluna j e o denominador é a média de todos os elementos da inversa de Leontief. Logo V_j é o *índice de ligação para trás* na medida que os seus valores demonstram a resposta que os outros setores têm após um aumento de seu produto.

De outra forma, a medida das ligações interindustriais que é definida usando as linhas da matriz inversa de Leontief é o que se segue,

$$R_i = \sum_{j=1}^n b_{ij} \quad (j = 1, 2, \dots, n) \quad (87)$$

\mathbf{R}_i mede o produto que seria gerado no setor i caso a demanda final de cada setor fosse aumentada por uma unidade. Como em (86) \mathbf{R}_i pode ser normalizado utilizando,

$$U_i = \frac{\frac{1}{n} R_i}{\frac{1}{n^2} \sum_{i,j} b_{ij}} \quad (88)$$

no qual o numerador é a média dos elementos da linha i e o denominador é a média de todos os elementos da inversa de Leontief. Logo U_i é o *índice de ligação para frente* na medida que os seus valores demonstram a oferta que esse setor produz para os demais setores após um aumento de seu produto.

Dessa forma, os setores chaves são aqueles nos quais V_j e U_i são maiores que a unidade, embora alguns autores considerem que aqueles setores com um ou outro maior que 1 também podem ser considerados setores-chaves ou, pelo menos muito relevantes para aquela região em estudo.

5.2. MODELO DE EQUILÍBRIO GERAL COMPUTÁVEL

Os modelos de Equilíbrio Geral Computável (EGC) assumem a estrutura de uma economia concorrencial *walrasiana*. A pressuposição básica é que nessa economia há três agentes principais: firmas, famílias e governos que produzem, consomem bens, serviços e fatores e pagam impostos, no mercado nacional e internacional. Neste projeto usou-se um modelo de EGC capaz de analisar o impacto de políticas públicas: o Projeto de Análise de Equilíbrio Geral da Economia Brasileira (PAEG) que pode representar de forma desagregada as regiões brasileiras.

O PAEG é um modelo econômico multirregional e multissetorial, construído com vistas a analisar a economia brasileira de forma regional, sendo em cada uma das regiões é representada por uma estrutura de demanda intermediária e final, composta por setores selecionados e despesas públicas e privadas com bens e serviços (Teixeira et al., 2013). Ele está plenamente integrado ao modelo e base de dados do *Global Trade Analysis Project* (GTAP), que inclui os fluxos de transações econômicas globais pela compatibilização de matrizes de insumo-produto (MIPs) nacionais.

O PAEG baseia-se nas pressuposições microeconômicas clássicas para o comportamento dos agentes econômicos: o consumidor representativo e o governo buscam otimizar seu bem-estar sujeito a uma restrição orçamentária, e os setores produtivos combinam

insumos intermediários e fatores primários de produção, com vistas a minimizar os custos, dada a tecnologia.

Por hipótese, as preferências são contínuas e convexas, das quais resultam funções de demanda contínuas e homogêneas de grau zero em relação aos preços, ou seja, somente os preços relativos podem ser determinados. Pelo lado das firmas, a tecnologia é representada por uma função de produção com rendimentos constantes de escala, significando que o lucro econômico das firmas é nulo no equilíbrio, agindo em mercados perfeitamente competitivos.

Dessa forma, podem-se enumerar três condições essenciais de consistências da base de dados: equilíbrio dos mercados (oferta igual à demanda para todos os bens e fatores); balanço da renda, isto é, renda líquida igual à despesa líquida para cada agente econômico; e, por último, a renda é exaurida pelas unidades produtivas, dado um conjunto de identidades que se aplicam a cada um dos setores produtivos: lucro econômico igual a zero.

Nesse sentido, do ponto de vista microeconômico, opera-se estritamente numa economia pela sua ótica produtiva ou real, desconsiderando, portanto, o fenômeno inflacionário. Em certo sentido, opera-se numa perspectiva de longo prazo por causa da mobilidade entre setores e regiões de todos os recursos da economia e pelo fato de os agentes tomadores de decisões terem, por hipótese, todas as condições de ajustar suas escolhas ótimas à medida que os preços relativos se alterem.

Quanto às regras de fechamento macroeconômicas, elas são, na verdade, desdobramentos naturais das anteriores, conferindo coerência teórica ao modelo. Como dito anteriormente, a demanda de investimento e os fluxos de capitais internacionais são fixados exogenamente nos valores do ano base. Dessa forma, mudanças na taxa real de câmbio devem ocorrer para acomodar alterações nos fluxos de exportações e importações após os choques, mantendo o saldo do balanço de pagamentos constante.

A dotação total de fatores de produção não se altera, mas são móveis entre setores e entre as regiões brasileiras. Na verdade, há três tipos de mobilidade de fatores de produção no modelo utilizado: no primeiro caso, sem migração de fatores, os rendimentos de capital e trabalho se equalizam entre setores somente dentro da sua própria região de origem. No segundo tipo, denominada mobilidade imperfeita, ocorre a equalização da renda nominal dos fatores, sem considerar as variações nos preços de bens e serviços, entre setores e regiões do

Brasil. E por último, mobilidade perfeita de fatores com equalização dos rendimentos reais entre setores e regiões brasileiras mais uma vez.

Deste modo, os pressupostos macroeconômicos reforçam a ideia de uma economia modelada pelo lado real e com equilíbrio de longo prazo. Evidentemente, como o investimento é fixo, portanto, sem acumulação de capital, e não há um fluxo endógeno de renda líquida de fatores externos pela conta financeira no balanço de pagamentos, não se faz inferências de trajetória de crescimento do Produto Interno Bruto (PIB) no modelo.

Em termos computacionais, cabe comentar que o modelo PAEG tem flexibilidade para permitir diversas possibilidades de desagregações setoriais e regionais. A partir do modelo padrão PAEG/GTAPinGAMS descrito acima, fez-se um conjunto de alterações no código base em MPSGE no GAMS a fim de representar a nova desagregação regional e setorial desenvolvida. Em seguida, checkou-se a existência do equilíbrio de preços e quantidades em todos os mercados representados, preparando-o para a implementação dos choques previstos, medidos em termos de alterações da produção.

A proxy para verificar os ganhos de bem-estar nas regiões do modelo em função das simulações previstas neste relatório foi a medida de variação equivalente, que tem sido comumente utilizada para mensuração de ganhos de bem-estar em modelos aplicados de equilíbrio geral. As mudanças nos níveis de bem-estar auferidas pelo conceito de variação equivalente, expressa em termos percentual e monetário, reflete as consequências de diversas mudanças no consumo, na renda, na produção, fluxos comerciais e preços da economia. Ela pode ser calculada da seguinte forma:

$$VE = \left(\frac{U^f - U^0}{U^0} \right) \cdot C^0 \quad (89)$$

em que VE representa a variação equivalente, U^f representa o nível de utilidade final, U^0 retrata o nível de utilidade inicial e C^0 descreve a renda das famílias no equilíbrio inicial.

Assim, a medida de variação equivalente reflete a mudança no consumo necessário, a partir de um novo conjunto de preços, para que se mantenha o mesmo nível de utilidade das famílias, aos preços do equilíbrio inicial. A rigor, reflete o aumento da utilidade em termos de aumento do consumo. Ela indica aumentos de bem-estar para valores positivos e queda de bem-

estar para valores negativos. Sendo assim, a possível elevação do bem-estar em virtude de choques no modelo representará ganhos diretos para as famílias em termos de aumento de renda para consumir mais e, portanto, ganhos de utilidade.

5.3. FONTE E TRATAMENTO DOS DADOS

A desagregação e estimação das matrizes de insumo-produto para os estados jurisdicionados da Sudene para o período de 2002 a 2018 foram realizados a partir de uma ampla gama de informações coletadas da base de dados do Projeto de Análise de Equilíbrio Geral da Economia Brasileira (PAEG), do Instituto Brasileiro de Geografia e Estatística (IBGE) e do Núcleo de Economia Regional e Urbana da Universidade de São Paulo (NEREUS). Logo, é de se esperar que tenha exigido procedimentos técnicos sequenciados logicamente para consistência e consolidação dos dados.

Vale ressaltar que embora os anos de 2000, 2001 e 2019 façam parte do horizonte relevante da avaliação da política, não foi possível estimar as matrizes para os dois primeiros em função de alterações estruturais substanciais nas contas regionais do IBGE de 2001 para 2002, elevando significativamente os custos de estimação das matrizes para esse ano. Quanto ao ano de 2019, não houve envio de informações referentes as operações de crédito com recursos do FNE.

Assim sendo, o primeiro passo foi a elaboração de um tradutor de setores que pudesse equiparar o conjunto de atividades entre as bases de dados e a agregação desses nas suas respectivas regiões por meio do somatório correspondente, permitindo o passo seguinte. Definiu-se para esse relatório um total de 8 setores econômicos na matriz para os onze estados jurisdicionado da Sudene mais uma região reunindo todos os demais estados, denominado Restante do Brasil - RB, conforme Tabela 1 abaixo.

Tabela 1 – Desagregação setorial e regional das MIPs - Sudene

Setor	Descrição	Região	Descrição
S1	Agropecuária	MA	Maranhão
S2	Extrativa	PI	Piauí
S3	Transformação	CE	Ceará

S4	Utilities	RN	Rio Grande do Norte
S5	Construção	PB	Paraíba
S6	Comércio	PE	Pernambuco
S7	Transporte	AL	Alagoas
S8	Serviços	SE	Sergipe
		BA	Bahia
		MG	Minas Gerais
		ES	Espírito Santo
		RS	Resto do Brasil

Fonte: elaborado pelos autores

Em seguida, partindo-se de uma matriz de insumo-produto inter-regional de 2011 com todos os estados brasileiros desagregados estimada por Guilhoto (2019) procedeu-se a estimação e compatibilização temporal das MIPs requeridas para cada ano do período utilizando-se dos cálculos do Valor Bruto da Produção (VBP), Valor Adicionado (VA), Investimentos, Consumo Intermediário (CI) e Impostos Indiretos Líquidos (IIL) para os 8 setores e 12 regiões selecionadas, extraídas no Sistema de Contas Regionais do IBGE de 2002-2018¹⁴. A partir dessas informações, obteve-se, pelas funções de identidade da Contabilidade Nacional, os valores de Demanda Final e Produto Nacional.

Em seguida, de posse desses vetores “de margem” das transações intermediárias dos setores da MIP de cada ano foi possível atualizar a matriz de coeficientes técnicos para cada ano do período por meio uma técnica não-censitária denominada *RAS*¹⁵. A técnica basicamente estabelece um método iterativo de atualização desses coeficientes que os faça resultar nos vetores linha e coluna de consumo intermediário fixados de antemão (e conhecidos censitariamente no ano de interesse fornecido pelo IBGE, nesse relatório, para os anos de 2002 a 2018). Recebe esse nome por causa do “visual” da expressão matricial de estimação abaixo:

¹⁴<https://www.ibge.gov.br/estatisticas/economicas/contas-nacionais/9054-contas-regionais-do-brasil.html?=&t=resultados>

¹⁵ Não-censitário exatamente porque são usados métodos matemáticos de estimação desses coeficientes, ou seja, não foram coletados pelos órgãos competentes a partir de informações oficiais da própria base geradora dos fluxos. Esses fluxos intersetoriais e inter-regionais são, por sua própria natureza, de difícil e custosa apuração ficando, por isso, com defasagens de tempo consideráveis.

$$A(1) = \hat{r}A(0)\hat{s} \quad (90)$$

em que $A(1)$ é a matriz de coeficiente técnicos atualizada (2018, por exemplo) e \hat{r} e \hat{s} são matrizes calculadas a partir de ajustamentos dos vetores linha e coluna conhecidos do consumo intermediários de cada setor em cada região nas “bordas” (2018), onde cada elemento r_i de \hat{r} multiplica cada elemento i da linha de $A(0)$ e cada elemento s_j de \hat{s} multiplica cada elemento j da coluna de $A(0)$ ¹⁶, para todo $i, j = 1, \dots, n$.

A técnica RAS de atualização dos coeficientes técnicos é capaz de capturar certos fenômenos econômicos estruturais chamados efeito-substituição e o efeito-fabricação. O primeiro refere-se ao surgimento de substitutos no processo produtivo. Por exemplo, se houve uma redução do uso de borracha no ano mais recente, este insumo terá proporcionalmente sua utilização reduzida em todos os setores que fazem uso dele. Já o efeito-fabricação considera a alteração na proporção entre valor adicionado e consumo intermediário nas compras totais do setor, ou seja, um setor que demonstre variação na relação citada terá todos os seus coeficientes de consumo intermediário ajustados proporcionalmente (Miller e Blair, 2009).

Enfim, checou-se o balanceamento de cada uma das MIPs estimada, isto é, o somatório de todas as linhas da matriz quadrada deve ser necessariamente igual ao somatório das colunas. Registra-se também a consistência do produto interno bruto (PIB) calculado pela ótica da renda, do produto e da demanda agregada com as informações oficiais do IBGE para cada ano do período avaliado.

De resto, a compatibilização dos setores das Matrizes de Insumo Produto com os setores que receberam repasses do FNE foi feita com base nos códigos do IBGE, no Sistema de Contas Regionais e as divisões da CNAE 2.0. Cabe ressaltar que os dados dos estados de Minas Gerais e Espírito Santo considera os valores do estado com um todo e não somente os municípios da área de atuação da SUDENE. Tal decisão foi pautada pela disponibilidade dos dados do Sistema de Contas Regionais, disponibilizadas pelo IBGE.

5.4. ESTRATÉGIA DE AVALIAÇÃO

¹⁶ Nesse relatório, o tempo “0” de $A(0)$ é 2011. Para os anos anteriores a lógica de estimação se mantém.

Além dos cálculos dos multiplicadores de produção e renda, índices de ligação e identificação de setores-chaves, a aplicação de choques nos modelos de insumo-produto e equilíbrio geral computável construídos para análise das economias regionais da Sudene com vistas a avaliar o retorno econômico e social do FNE seguiu uma estratégia direta e intuitiva.

Para representar a concessão de recursos do FNE nos modelos, assumiu-se que os setores receptores de recursos do FNE em cada estado, em cada ano, ampliam a produção em um montante equivalente e, ao fazê-lo, os recursos se distribuam na economia pela necessidade de elevação da Demanda Final, no caso do modelo IP, e de consumo de insumos intermediários e fatores produtivos, no modelo EGC, que por sua vez terão impactos sobre os demais setores da economia e sobre a oferta disponível para consumo intermediário e final e formação de investimentos.

Dessa forma, no modelo EGC, os choques aplicados expressam os ganhos de produção de cada setor nas regiões devido a investimentos via FNE, da seguinte forma¹⁷:

- 1) Calcula-se a proporção dos recursos do FNE em relação ao valor da produção setorial em cada estado;
- 2) Aplica-se a proporção calculada na produção setorial nos estados.

A hipótese assumida considera, portanto, que a injeção de novos recursos na economia não será capaz de captar possíveis aspectos relacionados ao custo de oportunidade de se empregar os recursos do FNE em outras atividades, ou dos recursos serem utilizados de outra forma pelo poder público. Dessa forma, os resultados de impactos obtidos dizem respeito aos efeitos esperados da aplicação de recursos que, se não fossem destinados ao FNE, não seriam utilizados para qualquer outro fim.

Um último aspecto importante sobre a estratégia de implementação é a hipótese de que os montantes dos recursos do FNE já circulam na economia no ano de construção da matriz de insumo-produto base do modelo e não afetam os possíveis efeitos de longo prazo que serão captados a partir do choque de injeção de recursos nas economias. Dois pontos podem ser listados para justificar tal hipótese:

- 1) Os montantes aplicados pelo FNE são proporções pequenas em relação ao produto

¹⁷ Outra estratégia possível de simulação do choque seria a de remoção dos recursos do FNE aplicados em determinado ano, de forma a captar como a economia seria impactada caso esses recursos não tivessem sido fornecidos. A não opção por esse tipo de choque é que ele seria apropriado apenas para captar os efeitos de curto prazo.

agregado da economia do Estado, e a matriz de insumo-produto, sendo uma base de dados estática para o modelo, não reflete os efeitos potenciais de encadeamentos desses recursos;

- 2) Os impactos proporcionados por esses recursos vão além de impactos de curto prazo, ou seja, se propagam por vários períodos, na medida em que são liberados e utilizados aos poucos para a compra de insumos intermediários e contratação de fatores produtivos, o que permite supor que os valores na matriz de insumo-produto captam apenas parte desses efeitos.

Dessa forma, os gastos do FNE implícitos no ano base da matriz de insumo-produto apenas significam um ponto de partida ligeiramente superior nos valores iniciais do modelo para efeitos de simulação e, portanto, não devem afetar os resultados das simulações baseadas na injeção desses recursos no modelo.



IPPDS
Instituto de Políticas Públicas e
Desenvolvimento Sustentável

6. CONSIDERAÇÕES FINAIS

Este documento tratou das metodologias de avaliação empregadas para a execução dos Produtos 5, 6, 7 e 8 da Meta 2 – Avaliação de eficácia, eficiência, efetividade e retorno econômico do FNE. Considerando essa base indicativa, a equipe técnica buscou detalhar de modo pormenorizado as ferramentas metodológicas utilizadas, as quais abarcaram métodos matemáticos, estatísticos e econométricos.

Para os Produtos 5 e 7, que tratam, respectivamente, da avaliação da eficácia e da efetividade do FNE, foram utilizadas ferramentas metodológicas similares. Tais análises foram pautadas na aplicação de ferramentas econométricas de avaliação de impacto, como o estimador de Diferenças em Diferenças e o método de Regressão Quantílica. Para o Produto 5 em específico, foram empregadas também as Funções Dose-Resposta. Enquanto a análise conduzida no Produto 5 foi realizada em nível de empresas, a feita no Produto 7 considerou dados em nível de município.

O Produto 6, que considera a avaliação da eficiência do FNE, teve por base metodológica a utilização de um modelo não-paramétrico, chamada a Análise Envoltória de Dados (*Data Envelopment Analysis* – DEA). Devido à ausência de informações quanto ao faturamento das empresas investigadas, a análise de eficiência foi pautada em dois modelos distintos, porém complementares. Primeiro, foi aplicado um modelo regional, onde a análise foi conduzida em nível de municípios. Depois, foi conduzido um modelo setorial, onde a análise foi conduzida em nível de setores econômicos.

Por fim, tem-se o Produto 8, o qual versa sobre a avaliação do retorno social econômico do FNE. Nesse caso, foram empregadas as abordagens de Insumo-Produto (IP) e de Equilíbrio Geral Computável (EGC). Além dos cálculos dos multiplicadores de produção e renda, índices de ligação e identificação de setores-chaves, foram aplicados “choques” nos modelos IP e EGC construídos para a análise das economias regionais da área de atuação da Sudene, sendo possível assim a avaliação do retorno econômico e social do FNE.

ANEXO 1. ROTINAS DE PROGRAMAÇÃO EMPREGADAS NO PRODUTO 5

O Produto 5 da Meta 2, intitulado *Avaliação da eficácia da aplicação dos recursos do FNE*, lançou mão de ferramentas econométricas de avaliação de impacto para a obtenção dos resultados desejados. Em específico, foi utilizado o software estatístico *Stata*. Abaixo são apresentadas as linhas de programação que formam a rotina empregada para a análise empírica da eficácia do FNE.

```
* IMPACTOS ECONÔMICOS E SOCIAIS DO FNE - EMPRESAS

* Abrindo a base de dados a ser utilizada (nível: empresas)
use "fne_2000_2018_aval_merge_rais_nid.dta", clear

* Mantendo a base de dados entre 2000 e 2018 (disponibilizada BNB)
keep if ano >=2000 & ano<=2018

* Declarando dados em painel
xtset new_id_firm ano

* Definindo as variáveis que serão utilizadas

** Variáveis de controle

*** Grandes setores da economia (IBGE)
gen industria = subs_ibge >=1 & subs_ibge <=14
gen civil = subs_ibge ==15
gen comercio = subs_ibge >=16 & subs_ibge <=17
gen servicos = subs_ibge >=18 & subs_ibge <=24
gen agropecuaria = subs_ibge ==25

*** Tamanho do estabelecimento
label define tam 0 "0 empregado" 1 "De 1 a 4" 2 "De 5 a 9" 3 "De 10 a 19" ///
4 "De 20 a 49" 5 "De 50 a 99" 6 "De 100 a 249" 7 "De 250 a 499" 8 "De 500 a 999" 9 "Mais
de 1000"
label values tamestab tam

*** Definindo variáveis binárias (dummies) para tamanho do estabelecimento (classificação
do Sebrae/IBGE - http://www.sebrae-sc.com.br/leis/default.asp?vcdtexto=4154)

*Indústria:
*Micro: com até 19 empregados
*Pequena: de 20 a 99 empregados
*Média: 100 a 499 empregados
```

*Grande: mais de 500 empregados

*Comércio e Serviços

*Micro : até 9 empregados

*Pequena: de 10 a 49 empregados

*Média: de 50 a 99 empregados

*Grande: mais de 100 empregados

gen micro_pequena=0

replace micro_pequena=1 if tamestab<=4 | tamestab==5 & industria==1

gen media=0

replace media=1 if tamestab==5 & industria==0 | tamestab==6 & industria==1 |
tamestab==7 & industria==1

gen grande=0

replace grande=1 if tamestab>=6 & industria==0 | tamestab==8 & industria==1 |
tamestab==9 & industria==1

*** Definindo o estado em que a empresa está localizada

encode nm_uf, gen (uf)

gen AL = 0

replace AL = 1 if uf==1

sort uf

xi, prefix() i.uf

rename (_Iuf_2 _Iuf_3 _Iuf_4 _Iuf_5 _Iuf_6 _Iuf_7 _Iuf_8 _Iuf_9 _Iuf_10 _Iuf_11) ///
(BA CE ES MA MG PB PE PI RN SE)

global x_list AL BA CE ES MA MG PB PE PI RN SE micro_pequena media grande
industria civil comercio servicios agropecuaria med_grau_instr

*** IPCA

gen ipca=.

replace ipca=32.5527 if ano==2000

replace ipca=34.7795 if ano==2001

replace ipca=37.7184 if ano==2002

replace ipca=43.2686 if ano==2003

replace ipca=46.1231 if ano==2004

replace ipca=49.2916 if ano==2005

replace ipca=51.3537 if ano==2006

replace ipca=53.2236 if ano==2007

replace ipca=56.2460 if ano==2008

replace ipca=58.9953 if ano==2009

```

replace ipca=61.9679 if ano==2010
replace ipca=66.0804 if ano==2011
replace ipca=69.6510 if ano==2012
replace ipca=73.9724 if ano==2013
replace ipca=78.6542 if ano==2014
replace ipca=85.7565 if ano==2015
replace ipca=93.2509 if ano==2016
replace ipca=96.4647 if ano==2017
replace ipca=100.000 if ano==2018

** Variáveis de resultado (outcomes)

*** Estoque de emprego

egen ee = rowtotal(ens_analf- ens_superior_i) // ee = estoque de emprego
gen ln_ee = ln(ee)
*gen ln_tt = ln(tt_vinculo) // total de vínculos em dezembro
label var ee "Estoque de emprego"
label var ln_ee "ln do estoque de emprego"
*label var ln_tt "ln do total de vínculos"

*** Massa salarial
gen ln_ms = ln((100/ipca)*rem_med_r) // normalmente os salários são corrigidos pela
inflação
label var ln_ms "ln da massa salarial"

*** Salário médio
gen ln_sm = ln((100/ipca)*med_rem_med_r)
label var ln_sm "ln do salário médio"

** Variáveis de tratamento (staggered)

*** Recebimento de recursos do FNE (x=1 a partir do primeiro ano que recebeu recursos)

gen w_fne = 1 if v1_contrato != .
recode w_fne (1=.) if v1_contrato == 0
by new_id_firm (ano), sort: replace w_fne = w_fne[_n-1] if w_fne>=.
recode w_fne (.=0)

*** Recebimento de recursos do FNE por finalidade (x=1 a partir do primeiro ano que
recebeu recursos)
encode finalidade_new, gen(finalidade_new2)

**** Capital de Giro e Custeio
gen w_fne_gc = 1 if finalidade_new2==1 | finalidade_new2==2 //Capital de Giro e Custeio
by new_id_firm (ano), sort: replace w_fne_gc = w_fne_gc[_n-1] if w_fne_gc>=.
recode w_fne_gc (.=0)

```

**** Investimento

```
gen w_fne_in = 1 if finalidade_new2==3           //Investimento
by new_id_firm (ano), sort: replace w_fne_in = w_fne_in[_n-1] if w_fne_in>=.
recode w_fne_in (.=0)
```

drop finalidade_new2

*** Recebimento de recursos do FNE por programa (x=1 a partir do primeiro ano que recebeu recursos)

```
encode programa_new, gen(programa_new2)
```

**** FNE AGRIN

```
gen w_agrin = 1 if programa_new2==1
by new_id_firm (ano), sort: replace w_agrin = w_agrin[_n-1] if w_agrin>=.
recode w_agrin (.=0)
```

**** FNE AQUIPESCA

```
gen w_aquipesca = 1 if programa_new2==2
by new_id_firm (ano), sort: replace w_aquipesca = w_aquipesca[_n-1] if w_aquipesca>=.
recode w_aquipesca (.=0)
```

**** FNE COMÉRCIO E SERVIÇOS

```
gen w_com_ser = 1 if programa_new2==3
by new_id_firm (ano), sort: replace w_com_ser = w_com_ser[_n-1] if w_com_ser>=.
recode w_com_ser (.=0)
```

**** FNE INDUSTRIAL

```
gen w_ind = 1 if programa_new2==4
by new_id_firm (ano), sort: replace w_ind = w_ind[_n-1] if w_ind>=.
recode w_ind (.=0)
```

**** FNE IRRIGAÇÃO

```
gen w_irri = 1 if programa_new2==5
by new_id_firm (ano), sort: replace w_irri = w_irri[_n-1] if w_irri>=.
recode w_irri (.=0)
```

**** FNE MPE

```
gen w_mpe = 1 if programa_new2==6
by new_id_firm (ano), sort: replace w_mpe = w_mpe[_n-1] if w_mpe>=.
recode w_mpe (.=0)
```

**** FNE OUTROS

```
gen w_outros = 1 if programa_new2==7
by new_id_firm (ano), sort: replace w_outros = w_outros[_n-1] if w_outros>=.
recode w_outros (.=0)
```

**** FNE PROATUR

```
gen w_proatur = 1 if programa_new2==8
by new_id_firm (ano), sort: replace w_proatur = w_proatur[_n-1] if w_proatur>=.
recode w_proatur (.=0)
```

**** FNE RURAL

```
gen w_rural = 1 if programa_new2==9
by new_id_firm (ano), sort: replace w_rural = w_rural[_n-1] if w_rural>=.
recode w_rural (.=0)
```

**** FNE SECA

```
gen w_seca = 1 if programa_new2==10
by new_id_firm (ano), sort: replace w_seca = w_seca[_n-1] if w_seca>=.
recode w_seca (.=0)
```

**** FNE VERDE

```
gen w_verde = 1 if programa_new2==11
by new_id_firm (ano), sort: replace w_verde = w_verde[_n-1] if w_verde>=.
recode w_verde (.=0)
```

**** PRODETEC

```
gen w_prodetec = 1 if programa_new2==12
by new_id_firm (ano), sort: replace w_prodetec = w_prodetec[_n-1] if w_prodetec>=.
recode w_prodetec (.=0)
```

**** PRODIN

```
gen w_prodin = 1 if programa_new2==13
by new_id_firm (ano), sort: replace w_prodin = w_prodin[_n-1] if w_prodin>=.
recode w_prodin (.=0)
```

**** PROGER

```
gen w_proger = 1 if programa_new2==14
by new_id_firm (ano), sort: replace w_proger = w_proger[_n-1] if w_proger>=.
recode w_proger (.=0)
```

**** PROINFRA

```
gen w_proinfra = 1 if programa_new2==15
by new_id_firm (ano), sort: replace w_proinfra = w_proinfra[_n-1] if w_proinfra>=.
recode w_proinfra (.=0)
```

**** PROMICRO

```
gen w_promicro = 1 if programa_new2==16
by new_id_firm (ano), sort: replace w_promicro = w_promicro[_n-1] if w_promicro>=.
recode w_promicro (.=0)
```

```
drop programa_new2
```

* Gerando variáveis para estimação dos estimadores TWFE Wooldridge (2021)

** Gerando dummies de ano (para anos em que houve tratamento)

```
gen f2000 = ano == 2000
gen f2001 = ano == 2001
gen f2002 = ano == 2002
gen f2003 = ano == 2003
gen f2004 = ano == 2004
gen f2005 = ano == 2005
gen f2006 = ano == 2006
gen f2007 = ano == 2007
gen f2008 = ano == 2008
gen f2009 = ano == 2009
gen f2010 = ano == 2010
gen f2011 = ano == 2011
gen f2012 = ano == 2012
gen f2013 = ano == 2013
gen f2014 = ano == 2014
gen f2015 = ano == 2015
gen f2016 = ano == 2016
gen f2017 = ano == 2017
gen f2018 = ano == 2018
```

** Gerando o tempo de exposição ao tratamento em anos

```
foreach var of varlist w_fne-w_promicro {
by new_id_firm (ano), sort: egen yt_`var' = min(`var'*ano) if `var' != 0 // yt: ano de
tratamento
by new_id_firm (ano), sort: egen ft_`var' = max(yt_`var') // ft: primeiro ano de tratamento
recode ft_`var' (. = 0)
drop yt_`var'
}
```

** Gerando dummies para grupos de tratamento, de acordo com o tempo de exposição ao tratamento

```
foreach var of varlist ft_w_fne-ft_w_promicro {
gen d0`var' = `var' == 2000 // d0=1 (19 anos de exposição ao tratamento, entrada em
2000)
gen d1`var' = `var' == 2001 // d1=1 (18 anos de exposição ao tratamento, entrada em
2001)
gen d2`var' = `var' == 2002 // d2=1 (17 anos de exposição ao tratamento, entrada em
2002)
gen d3`var' = `var' == 2003 // d3=1 (16 anos de exposição ao tratamento, entrada em
2003)
```

```

gen d4`var' = `var' == 2004 // d4=1 (15 anos de exposição ao tratamento, entrada em
2004)
gen d5`var' = `var' == 2005 // d5=1 (14 anos de exposição ao tratamento, entrada em
2005)
gen d6`var' = `var' == 2006 // d6=1 (13 anos de exposição ao tratamento, entrada em
2006)
gen d7`var' = `var' == 2007 // d7=1 (12 anos de exposição ao tratamento, entrada em
2007)
gen d8`var' = `var' == 2008 // d8=1 (11 anos de exposição ao tratamento, entrada em
2008)
gen d9`var' = `var' == 2009 // d9=1 (10 anos de exposição ao tratamento, entrada em
2009)
gen d10`var' = `var' == 2010 // d10=1 (9 anos de exposição ao tratamento, entrada em
2010)
gen d11`var' = `var' == 2011 // d11=1 (8 anos de exposição ao tratamento, entrada em
2011)
gen d12`var' = `var' == 2012 // d12=1 (7 anos de exposição ao tratamento, entrada em
2012)
gen d13`var' = `var' == 2013 // d13=1 (6 anos de exposição ao tratamento, entrada em
2013)
gen d14`var' = `var' == 2014 // d14=1 (5 anos de exposição ao tratamento, entrada em
2014)
gen d15`var' = `var' == 2015 // d15=1 (4 anos de exposição ao tratamento, entrada em
2015)
gen d16`var' = `var' == 2016 // d16=1 (3 anos de exposição ao tratamento, entrada em
2016)
gen d17`var' = `var' == 2017 // d17=1 (2 anos de exposição ao tratamento, entrada em
2017)
gen d18`var' = `var' == 2018 // d18=1 (1 ano de exposição ao tratamento, entrada em
2018)
}

```

* Gerando variáveis para a estimação do Event Study

** Gerando variável TimeToTreat para estimar o event study

```

foreach var of varlist ft_w_fne-ft_w_promicro {
    gen ttt_`var' = ano - `var'
    replace ttt_`var' =. if `var' ==0 // ttt: tempo para o tratamento
}

```

* Estimação dos modelos

** Definindo diferentes recortes da amostra

*** Setor da economia

```

gen primario =1 if agropecuaria==1
recode primario (.=0)
gen secundario =1 if industria==1 | civil==1
recode secundario (.=0)
gen terciario =1 if comercio==1 | servicos==1
recode terciario (.=0)
gen setores = primario+secundario+terciario

*** Tamanho da empresa

*keep if micro_pequena==1
gen media_grande = media + grande
gen porte_all=micro_pequena+media_grande

* Estimando o efeito do tratamento (Average Treatment on the Treated - ATT) - FNE
GERAL

** ATT Constante (médio):

*** TWFE Wooldridge (2021)

foreach var of varlist ln_ee ln_ms ln_sm {
xtreg `var' w_fne ///
    f2000-f2018 ///
    $x_list ///
    , fe vce(cluster new_id_firm) // com controles
}

*Adicionar variáveis de controle não altera a estimativa do ATT, elas
*são adicionadas para relaxar o pressuposto de tendências paralelas

**** Por setor e tamanho de empresa

foreach var of varlist ln_ee ln_ms ln_sm {
    foreach setor of varlist primario secundario terciario setores {
        foreach porte of varlist micro_pequena media grande porte_all {
xtreg `var' w_fne ///
    f2000-f2018 ///
    $x_list ///
    if `setor'==1 & `porte'==1, ///
    fe vce(cluster new_id_firm) // com controles
        }
    }
}

*** Event Study

```

```

foreach var of varlist ln_ee ln_ms ln_sm {
eventdd `var' /// // com controles
    f2000-f2018 ///
    $x_list ///
    , inrange leads(3) lags(5) timevar(ttt_ft_w_fne) method(fe) cluster(new_id_firm)
}

**** Por setor e tamanho de empresa

foreach var of varlist ln_ee ln_ms ln_sm {
    foreach setor of varlist primario secundario terciario setores {
        foreach porte of varlist micro_pequena media grande porte_all {
eventdd `var' /// // com controles
            f2000-f2018 ///
            $x_list ///
            if `setor'==1 & `porte'==1, ///
            inrange leads(3) lags(5) timevar(ttt_ft_w_fne) method(fe) cluster(new_id_firm)
                }
        }
    }
}

*** DRDID Callaway e Sant'Anna (2021)

foreach var of varlist ln_ee ln_ms ln_sm {
csdid `var' $x_list /// // com controles
    , ivar(new_id_firm) time(ano) gvar(ft_w_fne) method(drimp) wboot agg(simple) notyet
}

**** Por setor e tamanho de empresa

foreach var of varlist ln_ee ln_ms ln_sm {
    foreach setor of varlist primario secundario terciario setores {
        foreach porte of varlist micro_pequena media grande porte_all {
csdid `var' $x_list /// // com controles
            if `setor'==1 & `porte'==1, ///
            ivar(new_id_firm) time(ano) gvar(ft_w_fne) method(drimp) wboot agg(simple)
notyet
                }
        }
    }
}

** ATT variando ao longo do tempo:

*** TWFE Wooldridge (2021)

*foreach var of varlist ln_ee-ln_sm {

```

```

*xtreg `var' c.w_fne#c.(f2000-f2018) ///
*   f2000-f2018 ///
*   $x_list ///
*   , fe vce(cluster new_id_firm)           // com controles
*}

** ATT variando por grupo de exposição ao tratamento:

*** TWFE Wooldridge (2021)

*foreach var of varlist ln_ee-ln_sm {
*xtreg `var' c.w_fne#c.(d0ft_w_fne-d18ft_w_fne) ///
*   f2000-f2018 ///
*   $x_list ///
*   , fe vce(cluster new_id_firm)           // com controles
*}

** ATT variando por grupo e ao longo o tempo:

*foreach var of varlist ln_ee-ln_sm {
*xtreg `var' c.(f2000-f2018)#c.(d0ft_w_fne-d18ft_w_fne) ///
*   f2000-f2018 ///
*   $x_list ///
*   , fe vce(cluster new_id_firm)           // com controles
*}

* Estimando o efeito do tratamento (Average Treatment on the Treated - ATT) - FNE
FINALIDADE

** ATT Constante (médio):

*** TWFE Wooldridge (2021)

foreach var of varlist ln_ee ln_ms ln_sm {
xtreg `var' w_fne_gc w_fne_in ///
   f2000-f2018 ///
   $x_list ///
   , fe vce(cluster new_id_firm)           // com controles
}

*Adicionar variáveis de controle não altera a estimativa do ATT, elas
*são adicionadas para relaxar o pressuposto de tendências paralelas

foreach var of varlist ln_ee ln_ms ln_sm {
   foreach setor of varlist primario secundario terciario setores {
       foreach porte of varlist micro_pequena media grande porte_all {
xtreg `var' w_fne_gc w_fne_in ///

```

```

f2000-f2018 ///
  $x_list ///
  if `setor'==1 & `porte'==1, ///
  fe vce(cluster new_id_firm) // com controles
    }
  }
}

*** Event Study

*capital de giro e custeio (problema na estimação)
*foreach y of varlist ln_ee ln_ms ln_sm {
*eventdd `y' /// // com controles
*   f2000-f2018 ///
*   $x_list /// // capital de giro e custeio
*   , inrange leads(3) lags(5) timevar(ttt_ft_w_fne_gc) method(fe) cluster(new_id_firm)
*}

**** Por setor e tamanho de empresa

foreach var of varlist ln_ee ln_ms ln_sm {
  foreach setor of varlist primario secundario terciario setores {
    foreach porte of varlist micro_pequena media grande porte_all {
eventdd `var' /// // com controles
  f2000-f2018 ///
  $x_list ///
  if `setor'==1 & `porte'==1, ///
  inrange leads(3) lags(5) timevar(ttt_ft_w_fne_gc) method(fe) cluster(new_id_firm)
    }
  }
}

***** Investimento

foreach y of varlist ln_ee ln_ms ln_sm {
eventdd `y' /// // com controles
  f2000-f2018 ///
  $x_list /// // investimento
  , inrange leads(3) lags(5) timevar(ttt_ft_w_fne_in) method(fe) cluster(new_id_firm)
}

**** Por setor e tamanho de empresa

foreach var of varlist ln_ee ln_ms ln_sm {
  foreach setor of varlist primario secundario terciario setores {
    foreach porte of varlist micro_pequena media grande porte_all {
eventdd `var' /// // com controles

```

```

f2000-f2018 ///
$x_list ///
if `setor'==1 & `porte'==1, ///
inrange leads(3) lags(5) timevar(ttt_ft_w_fne_in) method(fe) cluster(new_id_firm)
    }
}
}

*** DRDID Callaway e Sant'Anna (2021)

**** Capital de giro e custeio

foreach y of varlist ln_ee ln_ms ln_sm {
csdid `y' $x_list /// // com controles
, ivar(new_id_firm) time(ano) gvar(ft_w_fne_gc) method(drimp) wboot agg(simple)
notyet
}

**** Por setor e tamanho de empresa

foreach var of varlist ln_ee ln_ms ln_sm {
    foreach setor of varlist primario secundario terciario setores {
        foreach porte of varlist micro_pequena media grande porte_all {
csdid `var' $x_list /// // com controles
if `setor'==1 & `porte'==1, ///
ivar(new_id_firm) time(ano) gvar(ft_w_fne_gc) method(drimp) wboot agg(simple)
notyet
        }
    }
}

**** Investimento

foreach y of varlist ln_ee ln_ms ln_sm {
csdid `y' $x_list /// // com controles
, ivar(new_id_firm) time(ano) gvar(ft_w_fne_in) method(drimp) wboot agg(simple)
notyet
}

**** Por setor e tamanho de empresa

foreach var of varlist ln_ee ln_ms ln_sm {
    foreach setor of varlist primario secundario terciario setores {
        foreach porte of varlist micro_pequena media grande porte_all {
csdid `var' $x_list /// // com controles
if `setor'==1 & `porte'==1, ///

```

```

        ivar(new_id_firm) time(ano) gvar(ft_w_fne_in) method(drimp) wboot agg(simple)
notyet
    }
}

```

** ATT variando ao longo do tempo:

*** TWFE Wooldridge (2021)

```

*foreach var of varlist ln_ee-ln_sm {
*xtnreg `var' ///
*   c.(w_fne_gc w_fne_in)#c.(f2000-f2018) ///
*   f2000-f2018 ///
*   $x_list ///
*   , fe vce(cluster new_id_firm)           // com controles
*}

```

** ATT variando por grupo de exposição ao tratamento:

*** TWFE Wooldridge (2021)

```

*foreach var of varlist ln_ee-ln_sm {
*xtnreg `var' ///
*   c.w_fne_gc#c.(d0ft_w_fne_gc-d18ft_w_fne_gc) ///
*   c.w_fne_in#c.(d0ft_w_fne_in-d18ft_w_fne_in) ///
*   f2000-f2018 ///
*   $x_list ///
*   , fe vce(cluster new_id_firm)           // com controles
*}

```

** ATT variando por grupo e ao longo o tempo:

```

*foreach var of varlist ln_ee-ln_sm {
*xtnreg `var' ///
*   c.(f2000-f2018)#c.(d0ft_w_fne_gc-d18ft_w_fne_gc) ///
*   c.(f2000-f2018)#c.(d0ft_w_fne_in-d18ft_w_fne_in) ///
*   f2000-f2018 ///
*   $x_list ///
*   , fe vce(cluster new_id_firm)           // com controles
*}

```

* Estimando o efeito do tratamento (Average Treatment on the Treated - ATT) - FNE PROGRAMAS

** ATT Constante (médio):

*** TWFE Wooldridge (2021)

```
foreach var of varlist ln_ee-ln_sm {
xtreg `var' w_agrin-w_promicro ///
    f2000-f2018 ///
    $x_list ///
    , fe vce(cluster new_id_firm)           // com controles
}
```

*Adicionar variáveis de controle não altera a estimativa do ATT, elas

*são adicionadas para relaxar o pressuposto de tendências paralelas

**** Por setor e tamanho de empresa

```
foreach var of varlist ln_ee ln_ms ln_sm {
    foreach setor of varlist primario secundario terciario setores {
        foreach porte of varlist micro_pequena media grande porte_all {
xtreg `var' w_agrin-w_promicro ///
    f2000-f2018 ///
    $x_list ///
    if `setor'==1 & `porte'==1, ///
    fe vce(cluster new_id_firm) // com controles
        }
    }
}
```

*** Event Study

```
foreach y of varlist ln_ee-ln_sm {
    foreach x of varlist ttt_ft_w_agrin-ttt_ft_w_promicro {
eventdd `y' /// // com controles
    f2000-f2018 ///
    $x_list ///
    , inrange leads(3) lags(5) timevar(`x') method(fe) cluster(new_id_firm)
    }
}
```

**** Por setor e tamanho de empresa

```
foreach var of varlist ln_ee ln_ms ln_sm {
    foreach setor of varlist primario secundario terciario setores {
        foreach porte of varlist micro_pequena media grande porte_all {
            foreach x of varlist ttt_ft_w_agrin-ttt_ft_w_promicro {
eventdd `var' /// // com controles
    f2000-f2018 ///
    $x_list ///
    if `setor'==1 & `porte'==1, ///
    inrange leads(3) lags(5) timevar(`x') method(fe) cluster(new_id_firm)
            }
        }
    }
}
```

```

    }
  }
}

*** DRDID Callaway e Sant'Anna (2021)

foreach y of varlist ln_ee-ln_sm {
  foreach x of varlist ft_w_agrin-ft_w_promicro {
csdid `y' $x_list /// // com controles
  , ivar(new_id_firm) time(ano) gvar(`x') method(drimp) wboot agg(simple) notyet
  }
}

**** Por setor e tamanho de empresa

foreach var of varlist ln_ee ln_ms ln_sm {
  foreach setor of varlist primario secundario terciario setores {
    foreach porte of varlist micro_pequena media grande porte_all {
      foreach x of varlist ft_w_agrin-ft_w_promicro {
csdid `var' $x_list /// // com controles
      if `setor'==1 & `porte'==1, ///
        ivar(new_id_firm) time(ano) gvar(`x') method(drimp) wboot agg(simple) notyet
      }
    }
  }
}

** ATT variando ao longo do tempo:

*** TWFE Wooldridge (2021)

*foreach var of varlist ln_ee-ln_sm {
*xxtreg `var' ///
* c.(w_agrin-w_promicro)#c.(f2000-f2018) ///
* f2000-f2018 ///
* $x_list ///
* , fe vce(cluster new_id_firm) // com controles
*}

** ATT variando por grupo de exposição ao tratamento:

*** TWFE Wooldridge (2021)

*foreach var of varlist ln_ee-ln_sm {
*xxtreg `var' ///
* c.w_agrin#c.(d0ft_w_agrin-d18ft_w_agrin) ///

```

```

*      c.w_aquipesca#c.(d0ft_w_aquipesca-d18ft_w_aquipesca) ///
*      c.w_com_ser#c.(d0ft_w_com_ser-d18ft_w_com_ser) ///
*      c.w_ind#c.(d0ft_w_ind-d18ft_w_ind) ///
*      c.w_irri#c.(d0ft_w_irri-d18ft_w_irri) ///
*      c.w_mpe#c.(d0ft_w_mpe-d18ft_w_mpe) ///
*      c.w_outros#c.(d0ft_w_outros-d18ft_w_outros) ///
*      c.w_proatur#c.(d0ft_w_proatur-d18ft_w_proatur) ///
*      c.w_rural#c.(d0ft_w_rural-d18ft_w_rural) ///
*      c.w_seca#c.(d0ft_w_seca-d18ft_w_seca) ///
*      c.w_verde#c.(d0ft_w_verde-d18ft_w_verde) ///
*      c.w_prodetec#c.(d0ft_w_prodetec-d18ft_w_prodetec) ///
*      c.w_prodin#c.(d0ft_w_prodin-d18ft_w_prodin) ///
*      c.w_proger#c.(d0ft_w_proger-d18ft_w_proger) ///
*      c.w_proinfra#c.(d0ft_w_proinfra-d18ft_w_proinfra) ///
*      c.w_promicro#c.(d0ft_w_promicro-d18ft_w_promicro) ///
*      f2000-f2018 ///
*      $x_list ///
*      , fe vce(cluster new_id_firm)          // com controles
* }

```

** ATT varia por grupo e ao longo o tempo:

```

*foreach var of varlist ln_ee-ln_sm {
*xtreg `var' ///
*      c.(f2000-f2018)#c.(d0ft_w_agrin-d18ft_w_agrin) ///
*      c.(f2000-f2018)#c.(d0ft_w_aquipesca-d18ft_w_aquipesca) ///
*      c.(f2000-f2018)#c.(d0ft_w_com_ser-d18ft_w_com_ser) ///
*      c.(f2000-f2018)#c.(d0ft_w_ind-d18ft_w_ind) ///
*      c.(f2000-f2018)#c.(d0ft_w_irri-d18ft_w_irri) ///
*      c.(f2000-f2018)#c.(d0ft_w_mpe-d18ft_w_mpe) ///
*      c.(f2000-f2018)#c.(d0ft_w_outros-d18ft_w_outros) ///
*      c.(f2000-f2018)#c.(d0ft_w_proatur-d18ft_w_proatur) ///
*      c.(f2000-f2018)#c.(d0ft_w_rural-d18ft_w_rural) ///
*      c.(f2000-f2018)#c.(d0ft_w_seca-d18ft_w_seca) ///
*      c.(f2000-f2018)#c.(d0ft_w_verde-d18ft_w_verde) ///
*      c.(f2000-f2018)#c.(d0ft_w_prodetec-d18ft_w_prodetec) ///
*      c.(f2000-f2018)#c.(d0ft_w_prodin-d18ft_w_prodin) ///
*      c.(f2000-f2018)#c.(d0ft_w_proger-d18ft_w_proger) ///
*      c.(f2000-f2018)#c.(d0ft_w_proinfra-d18ft_w_proinfra) ///
*      c.(f2000-f2018)#c.(d0ft_w_promicro-d18ft_w_promicro) ///
*      f2000-f2018 ///
*      $x_list ///
*      , fe vce(cluster new_id_firm)          // com controles
* }

```

* Iniciando a análise quantílica

```
summarize ln_ee, detail
summarize pib_sm, detail
summarize pib_ms, detail
```

```
gen teste_ln_ee = .
replace teste_ln_ee = 1 if ln_ee>0
replace teste_ln_ee = 0 if teste_ln_ee==.
```

- * o gráfico abaixo mostra os quantis das variáveis dependentes, estoque de emprego (ln_ee), massa salarial (ln_ms) e salário médio (ln_sm), e indica o formato da distribuição da respectiva variável. Pontos em cima da reta de 45° apontariam para uma distribuição uniforme dos dados.

```
quantile ln_ee, recast(line) xline(.10) xline(.25) xline(.50) xline(.75) xline(.90)
quantile ln_sm, recast(line) xline(.10) xline(.25) xline(.50) xline(.75) xline(.90)
quantile ln_ms, recast(line) xline(.10) xline(.25) xline(.50) xline(.75) xline(.90)
```

- *No entanto, é possível identificar que existe uma concentração de valores que se encontrariam acima do 25° quantil, indicando que a análise ficaria prejudicada caso se optasse apenas por MQO.

- *Como os dados utilizados nas regressões são obtidos de empresas atendidas pela Sudene, que possuem como características serem heterogêneas, elas vão se comportar como clusters e essa estrutura precisa ser levada em conta na estimativa desses modelos. Aldieri e Vinci (2017) mostraram a importância de se identificar correlações intra-cluster para se obter estimativas consistentes nessa classe de modelos. Para medir se existe esta possível correlação, aplica-se o teste de Parente Santos-Silva (PARENTE; SANTOS SILVA, 2016) que rejeitou a hipótese nula de ausência de correlação entre cluster.

```
qreg2 ln_ee w_fne f2000-f2018, cluster(new_id_firm)
qreg2 ln_sm w_fne f2000-f2018, cluster(new_id_firm)
qreg2 ln_ms w_fne f2000-f2018, cluster(new_id_firm)
```

- *O teste, feito para a mediana, indica a rejeição da hipótese nula de ausência de correlação intra-classe, apontando para a utilização da especificação quantílica.

- *No modelo a seguir estima-se o impacto dos empréstimos do FNE nas empresas atendidas desconsiderando a heterogeneidade individual dos municípios. Isso significa que os parâmetros estimados levam em conta os efeitos de w_fne, sobre o ln_ee, ln_sm, ln_ms ao longo dos quantis, independentemente das características individuais dos mesmos. Portanto, servem para isolar os efeitos das variáveis independentes sobre a variável dependente ao longo de toda sua distribuição. Este modelo permite com que os coeficientes respondam à pergunta: como as concessões de empréstimos do FNE afetam a geração de emprego, a massa salarial e a produtividade nos diferentes quantis, controlando pelas características individuais das empresas?

Vocês irão perceber que a estimação vai ser precedida do seguinte aviso**
***"WARNING: .15634283% of the fitted values of the scale function are not positive"
***Pesquisando na internet, no forum do Stata, o Prórpio Santos Silva,
***questionado sobre isso, indica que: "I do not know why the messaged appear
*** but with such small percentage I would not worry about."

***Definindo um segundo grupo de variáveis. Neste segundo grupo, para a análise
*** quantífica, não houvera as dummies de tamanho de empresa.

global x_list2 AL BA CE ES MA MG PB PE PI RN SE industria civil comercio servicos
agropecuaria med_grau_instr

** Estimando o peso geral para ln_ee (estoque de empregos)

xtqreg ln_ee w_fne f2000-f2018 \$x_list, i(new_id_firm) q(0.10)

xtqreg ln_ee w_fne f2000-f2018 \$x_list, i(new_id_firm) q(0.25)

xtqreg ln_ee w_fne f2000-f2018 \$x_list, i(new_id_firm) q(0.50)

xtqreg ln_ee w_fne f2000-f2018 \$x_list, i(new_id_firm) q(0.75)

xtqreg ln_ee w_fne f2000-f2018 \$x_list, i(new_id_firm) q(0.90)

** Estimando o peso geral para ln_ms (massa salarial)

xtqreg ln_ms w_fne f2000-f2018 \$x_list, i(new_id_firm) q(0.10)

xtqreg ln_ms w_fne f2000-f2018 \$x_list, i(new_id_firm) q(0.25)

xtqreg ln_ms w_fne f2000-f2018 \$x_list, i(new_id_firm) q(0.50)

xtqreg ln_ms w_fne f2000-f2018 \$x_list, i(new_id_firm) q(0.75)

xtqreg ln_ms w_fne f2000-f2018 \$x_list, i(new_id_firm) q(0.90)

** Estimando o peso geral para ln_sm (salário médio)

xtqreg ln_sm w_fne f2000-f2018 \$x_list, i(new_id_firm) q(0.10)

xtqreg ln_sm w_fne f2000-f2018 \$x_list, i(new_id_firm) q(0.25)

xtqreg ln_sm w_fne f2000-f2018 \$x_list, i(new_id_firm) q(0.50)

xtqreg ln_sm w_fne f2000-f2018 \$x_list, i(new_id_firm) q(0.75)

```
xtqreg ln_sm w_fne f2000-f2018 $x_list, i(new_id_firm) q(0.90)
```

** Na sequência, foram estimados os parâmetros para os quantis e construídos os gráficos no Excel.

```
xtqreg ln_ee w_fne f2000-f2018 $x_list, i(id_municipio) q(0.05(0.05)0.95)
```

```
xtreg ln_ee w_fne f2000-f2018 $x_list, fe vce(bootstrap)
```

```
xtqreg ln_ms w_fne f2000-f2018 $x_list, i(id_municipio) q(0.05(0.05)0.95)
```

```
xtreg ln_ms w_fne f2000-f2018 $x_list, fe vce(bootstrap)
```

```
xtqreg ln_sm w_fne f2000-f2018 $x_list, i(id_municipio) q(0.05(0.05)0.95)
```

```
xtreg ln_sm w_fne f2000-f2018 $x_list, fe vce(bootstrap)
```

*** Dummies de ano (para os anos em que houve tratamento)

```
gen f2000 = ano == 2000
```

```
gen f2001 = ano == 2001
```

```
gen f2002 = ano == 2002
```

```
gen f2003 = ano == 2003
```

```
gen f2004 = ano == 2004
```

```
gen f2005 = ano == 2005
```

```
gen f2006 = ano == 2006
```

```
gen f2007 = ano == 2007
```

```
gen f2008 = ano == 2008
```

```
gen f2009 = ano == 2009
```

```
gen f2010 = ano == 2010
```

```
gen f2011 = ano == 2011
```

```
gen f2012 = ano == 2012
```

```
gen f2013 = ano == 2013
```

```
gen f2014 = ano == 2014
```

```
gen f2015 = ano == 2015
```

```
gen f2016 = ano == 2016
```

```
gen f2017 = ano == 2017
```

```
gen f2018 = ano == 2018
```

** Variáveis de resultado (outcomes)

*** Estoque de emprego

```
egen ee = rowtotal(ens_analf-ens_superior_i) // ee = estoque de emprego
```

```
gen ln_ee = ln(ee)
```

```

*gen ln_tt = ln(tt_vinculo) // total de vínculos em dezembro
label var ee "Estoque de emprego"
label var ln_ee "ln do estoque de emprego"
*label var ln_tt "ln do total de vínculos"

*** Massa salarial
gen ln_ms = ln((100/ipca)*rem_med_r)
label var ln_ms "massa salarial"

** Salário médio
gen ln_sm = ln((100/ipca)*med_rem_med_r)
label var ln_sm "salário médio"

* Estimando os impacto econômicos - Funções dose-resposta (FDR)

** Mudando a unidade dos valores contratados para mil reais

*Para o comando ctreatreg a dose do tratamento deve estar no intervalo de 0 a 100

gen fne_gps = (vl_contrato_21/1000)
recode fne_gps (.=0)

sum fne_gps, d
*drop if fne_gps == r(max) // retirando valores extremos

sum fne_gps, d
gen fne_treat=(fne_gps-0)/(r(max)-0)*100

sum fne_gps if micro_pequena==1, d
gen fne_treat_mp=(fne_gps-0)/(r(max)-0)*100 if micro_pequena==1

sum fne_gps if media==1, d
gen fne_treat_me=(fne_gps-0)/(r(max)-0)*100 if media==1

sum fne_gps if grande==1, d
gen fne_treat_gr=(fne_gps-0)/(r(max)-0)*100 if grande==1

gen gc_gps = ((vl_fne_giro_21 + vl_fne_custeio_21)/1000)
recode gc_gps (.=0)

sum gc_gps, d
*drop if gc_gps == r(max) // retirando valores extremos

sum gc_gps, d
gen gc_treat=(gc_gps-0)/(r(max)-0)*100

sum gc_gps if micro_pequena==1, d

```

```

gen gc_treat_mp=(gc_gps-0)/(r(max)-0)*100 if micro_pequena==1

sum gc_gps if media==1, d
gen gc_treat_me=(gc_gps-0)/(r(max)-0)*100 if media==1

sum gc_gps if grande==1, d
gen gc_treat_gr=(gc_gps-0)/(r(max)-0)*100 if grande==1

gen in_gps = ((vl_fne_invest_21)/1000)
recode in_gps (.=0)

sum in_gps, d
gen in_treat=(in_gps-0)/(r(max)-0)*100

sum in_gps if micro_pequena==1, d
gen in_treat_mp=(in_gps-0)/(r(max)-0)*100 if micro_pequena==1

sum in_gps if media==1, d
gen in_treat_me=(in_gps-0)/(r(max)-0)*100 if media==1

sum in_gps if grande==1, d
gen in_treat_gr=(in_gps-0)/(r(max)-0)*100 if grande==1

** Definindo as variáveis de tratamento (staggered)

*** Recebimento de recursos do FNE (x=1 a partir do primeiro ano que recebeu recursos)

gen w_fne = 1 if vl_contrato != .
recode w_fne (1=.) if vl_contrato == 0

by new_id_firm (ano), sort: replace fne_treat = fne_treat[_n-1] if w_fne>=. // Complemento
de comando anterior
recode fne_treat (.=0)
by new_id_firm (ano), sort: replace fne_treat_mp = fne_treat_mp[_n-1] if w_fne>=.
recode fne_treat_mp (.=0)
by new_id_firm (ano), sort: replace fne_treat_me = fne_treat_me[_n-1] if w_fne>=.
recode fne_treat_me (.=0)
by new_id_firm (ano), sort: replace fne_treat_gr = fne_treat_gr[_n-1] if w_fne>=.
recode fne_treat_gr (.=0)

by new_id_firm (ano), sort: replace w_fne = w_fne[_n-1] if w_fne>=.
recode w_fne (.=0)

*** Recebimento de recursos do FNE por finalidade (x=1 a partir do primeiro ano que
recebeu recursos)
encode finalidade_new, gen(finalidade_new2)

```

```

**** Capital de Giro e Custeio
gen w_fne_gc = 1 if finalidade_new2==1 | finalidade_new2==2 //Capital de Giro e Custeio

by new_id_firm (ano), sort: replace gc_treat = gc_treat[_n-1] if w_fne_gc>=. //
Complemento de comando anterior
recode gc_treat (.=0)
by new_id_firm (ano), sort: replace gc_treat_mp = gc_treat_mp[_n-1] if w_fne_gc>=.
recode gc_treat_mp (.=0)
by new_id_firm (ano), sort: replace gc_treat_me = gc_treat_me[_n-1] if w_fne_gc>=.
recode gc_treat_me (.=0)
by new_id_firm (ano), sort: replace gc_treat_gr = gc_treat_gr[_n-1] if w_fne_gc>=.
recode gc_treat_gr (.=0)

by new_id_firm (ano), sort: replace w_fne_gc = w_fne_gc[_n-1] if w_fne_gc>=.
recode w_fne_gc (.=0)

**** Investimento
gen w_fne_in = 1 if finalidade_new2==3 //Investimento

by new_id_firm (ano), sort: replace in_treat = in_treat[_n-1] if w_fne_in>=. // Complemento
de comando anterior
recode in_treat (.=0)
by new_id_firm (ano), sort: replace in_treat_mp = in_treat_mp[_n-1] if w_fne_in>=.
recode in_treat_mp (.=0)
by new_id_firm (ano), sort: replace in_treat_me = in_treat_me[_n-1] if w_fne_in>=.
recode in_treat_me (.=0)
by new_id_firm (ano), sort: replace in_treat_gr = in_treat_gr[_n-1] if w_fne_in>=.
recode in_treat_gr (.=0)

by new_id_firm (ano), sort: replace w_fne_in = w_fne_in[_n-1] if w_fne_in>=.
recode w_fne_in (.=0)

drop finalidade_new2

*As estimações serão divididas por porte das empresas e finalidade do crédito

** Criando um arquivo log para armazenar os resultados
log using dose_resposta.txt, text replace

global x_list f2000-f2018 AL BA CE ES MA MG PB PE PI RN SE micro_pequena media
grande industria civil comercio servicos agropecuaria med_grau_instr

** Função Dose Resposta (FDR) - FNE GERAL

*** Todas as empresas
ctreatreg ln_ee w_fne $x_list , ///

```

```

graphdrf ci(5) delta(10) model(ct-ols) ct(fne_treat) m(2) s(10)
graph save "DRFIC" "DRFIC_fne_ee_all.gph"

ctreatreg ln_ms w_fne $x_list , ///
graphdrf ci(5) delta(10) model(ct-ols) ct(fne_treat) m(2) s(10)
graph save "DRFIC" "DRFIC_fne_ms_all.gph"

ctreatreg ln_sm w_fne $x_list , ///
graphdrf ci(5) delta(10) model(ct-ols) ct(fne_treat) m(2) s(10)
graph save "DRFIC" "DRFIC_fne_sm_all.gph"

*** Micro e pequenas empresas
ctreatreg ln_ee w_fne $x_list if micro_pequena==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(fne_treat_mp) m(2) s(10)
graph save "DRFIC" "DRFIC_fne_ee_mp.gph"

ctreatreg ln_ms w_fne $x_list if micro_pequena==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(fne_treat_mp) m(2) s(10)
graph save "DRFIC" "DRFIC_fne_ms_mp.gph"

ctreatreg ln_sm w_fne $x_list if micro_pequena==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(fne_treat_mp) m(2) s(10)
graph save "DRFIC" "DRFIC_fne_sm_mp.gph"

*** Empresas de médio porte
ctreatreg ln_ee w_fne $x_list if media==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(fne_treat_me) m(2) s(10)
graph save "DRFIC" "DRFIC_fne_ee_me.gph"

ctreatreg ln_ms w_fne $x_list if media==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(fne_treat_me) m(2) s(10)
graph save "DRFIC" "DRFIC_fne_ms_me.gph"

ctreatreg ln_sm w_fne $x_list if media==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(fne_treat_me) m(2) s(10)
graph save "DRFIC" "DRFIC_fne_sm_me.gph"

*** Empresas de grande porte
ctreatreg ln_ee w_fne $x_list if grande==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(fne_treat_gr) m(2) s(10)
graph save "DRFIC" "DRFIC_fne_ee_gr.gph"

ctreatreg ln_ms w_fne $x_list if grande==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(fne_treat_gr) m(2) s(10)
graph save "DRFIC" "DRFIC_fne_ms_gr.gph"

ctreatreg ln_sm w_fne $x_list if grande==1, ///

```

```
graphdrf ci(5) delta(10) model(ct-ols) ct(fne_treat_gr) m(2) s(10)
graph save "DRFIC" "DRFIC_fne_sm_gr.gph"
```

**** Função Dose Resposta (FDR) - FNE CAPITAL DE GIRO E CUSTEIO**

***** Todas as empresas**

```
ctreatreg ln_ee w_fne_gc $x_list , ///
graphdrf ci(5) delta(10) model(ct-ols) ct(gc_treat) m(2) s(10)
graph save "DRFIC" "DRFIC_gc_ee_all.gph"
```

```
ctreatreg ln_ms w_fne_gc $x_list , ///
graphdrf ci(5) delta(10) model(ct-ols) ct(gc_treat) m(2) s(10)
graph save "DRFIC" "DRFIC_gc_ms_all.gph"
```

```
ctreatreg ln_sm w_fne_gc $x_list , ///
graphdrf ci(5) delta(10) model(ct-ols) ct(gc_treat) m(2) s(10)
graph save "DRFIC" "DRFIC_gc_sm_all.gph"
```

***** Micro e pequenas empresas**

```
ctreatreg ln_ee w_fne_gc $x_list if micro_pequena==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(gc_treat_mp) m(2) s(10)
graph save "DRFIC" "DRFIC_gc_ee_mp.gph"
```

```
ctreatreg ln_ms w_fne_gc $x_list if micro_pequena==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(gc_treat_mp) m(2) s(10)
graph save "DRFIC" "DRFIC_gc_ms_mp.gph"
```

```
ctreatreg ln_sm w_fne_gc $x_list if micro_pequena==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(gc_treat_mp) m(2) s(10)
graph save "DRFIC" "DRFIC_gc_sm_mp.gph"
```

***** Empresas de médio porte**

```
ctreatreg ln_ee w_fne_gc $x_list if media==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(gc_treat_mp) m(2) s(10)
graph save "DRFIC" "DRFIC_gc_ee_me.gph"
```

```
ctreatreg ln_ms w_fne_gc $x_list if media==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(gc_treat_mp) m(2) s(10)
graph save "DRFIC" "DRFIC_gc_ms_me.gph"
```

```
ctreatreg ln_sm w_fne_gc $x_list if media==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(gc_treat_mp) m(2) s(10)
graph save "DRFIC" "DRFIC_gc_sm_me.gph"
```

***** Empresas de grande porte**

```
ctreatreg ln_ee w_fne_gc $x_list if grande==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(gc_treat_gr) m(2) s(10)
```

```
graph save "DRFIC" "DRFIC_gc_ee_gr.gph"

ctreatreg ln_ms w_fne_gc $x_list if grande==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(gc_treat_gr) m(2) s(10)
graph save "DRFIC" "DRFIC_gc_ms_gr.gph"

ctreatreg ln_sm w_fne_gc $x_list if grande==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(gc_treat_gr) m(2) s(10)
graph save "DRFIC" "DRFIC_gc_sm_gr.gph"

** Função Dose Resposta (FDR) - FNE INVESTIMENTO

*** Todas as empresas
ctreatreg ln_ee w_fne_in $x_list , ///
graphdrf ci(5) delta(10) model(ct-ols) ct(in_treat) m(2) s(10)
graph save "DRFIC" "DRFIC_in_ee_all.gph"

ctreatreg ln_ms w_fne_in $x_list , ///
graphdrf ci(5) delta(10) model(ct-ols) ct(in_treat) m(2) s(10)
graph save "DRFIC" "DRFIC_in_ms_all.gph"

ctreatreg ln_sm w_fne_in $x_list , ///
graphdrf ci(5) delta(10) model(ct-ols) ct(in_treat) m(2) s(10)
graph save "DRFIC" "DRFIC_in_sm_all.gph"

*** Micro e pequenas empresas
ctreatreg ln_ee w_fne_in $x_list if micro_pequena==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(in_treat_mp) m(2) s(10)
graph save "DRFIC" "DRFIC_in_ee_mp.gph"

ctreatreg ln_ms w_fne_in $x_list if micro_pequena==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(in_treat_mp) m(2) s(10)
graph save "DRFIC" "DRFIC_in_ms_mp.gph"

ctreatreg ln_sm w_fne_in $x_list if micro_pequena==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(in_treat_mp) m(2) s(10)
graph save "DRFIC" "DRFIC_in_sm_mp.gph"

*** Empresas de médio porte
ctreatreg ln_ee w_fne_in $x_list if media==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(in_treat_me) m(2) s(10)
graph save "DRFIC" "DRFIC_in_ee_me.gph"

ctreatreg ln_ms w_fne_in $x_list if media==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(in_treat_me) m(2) s(10)
graph save "DRFIC" "DRFIC_in_ms_me.gph"
```

```
ctreatreg ln_sm w_fne_in $x_list if media==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(in_treat_me) m(2) s(10)
graph save "DRFIC" "DRFIC_in_sm_me.gph"

*** Empresas de grande porte
ctreatreg ln_ee w_fne_in $x_list if grande==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(in_treat_gr) m(2) s(10)
graph save "DRFIC" "DRFIC_in_ee_gr.gph"

ctreatreg ln_ms w_fne_in $x_list if grande==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(in_treat_gr) m(2) s(10)
graph save "DRFIC" "DRFIC_in_ms_gr.gph"

ctreatreg ln_sm w_fne_in $x_list if grande==1, ///
graphdrf ci(5) delta(10) model(ct-ols) ct(in_treat_gr) m(2) s(10)
graph save "DRFIC" "DRFIC_in_sm_gr.gph"

log close
```



ANEXO 2. ROTINAS DE PROGRAMAÇÃO EMPREGADAS NO PRODUTO 6

O Produto 6 da Meta 2, intitulado *Avaliação da eficiência da aplicação dos recursos do FNE*, lançou mão de ferramentas de programação matemática para a solução de problemas de otimização e a consequente obtenção dos resultados desejados. Em específico, foi utilizado os softwares estatísticos *DEAP*, para a Análise Envoltória de Dados, *R*, para o teste de outliers, e *SPSS*, para o teste de igualdade de fronteiras, a análise discriminante e o teste de convergência. Abaixo são apresentadas as linhas de programação que formam a rotina empregada para a análise empírica da eficiência do FNE.

MALM.prn	DATA FILE NAME		
MALM.out	OUTPUT FILE NAME		
169	NUMBER OF FIRMS		
3	NUMBER OF TIME PERIODS		
2	NUMBER OF OUTPUTS		
1	NUMBER OF INPUTS		
1	0=INPUT AND 1=OUTPUT ORIENTATED		
1	0=CRS AND 1=VRS		
2	0=DEA(MULTI-STAGE),	1=COST-DEA,	2=MALMQUIST-DEA,
3	3=DEA(1-STAGE), 4=DEA(2-STAGE)		
P1_RV.prn	DATA FILE NAME		
P1_RV.out	OUTPUT FILE NAME		
169	NUMBER OF FIRMS		
1	NUMBER OF TIME PERIODS		
2	NUMBER OF OUTPUTS		
1	NUMBER OF INPUTS		
1	0=INPUT AND 1=OUTPUT ORIENTATED		
1	0=CRS AND 1=VRS		
4	0=DEA(MULTI-STAGE),	1=COST-DEA,	2=MALMQUIST-DEA,
3	3=DEA(1-STAGE), 4=DEA(2-STAGE)		
P2_RV.prn	DATA FILE NAME		
P2_RV.out	OUTPUT FILE NAME		
169	NUMBER OF FIRMS		
1	NUMBER OF TIME PERIODS		
2	NUMBER OF OUTPUTS		
1	NUMBER OF INPUTS		
1	0=INPUT AND 1=OUTPUT ORIENTATED		
1	0=CRS AND 1=VRS		
4	0=DEA(MULTI-STAGE),	1=COST-DEA,	2=MALMQUIST-DEA,
3	3=DEA(1-STAGE), 4=DEA(2-STAGE)		

P3_RV.prn DATA FILE NAME
P3_RV.out OUTPUT FILE NAME
169 NUMBER OF FIRMS
1 NUMBER OF TIME PERIODS
2 NUMBER OF OUTPUTS
1 NUMBER OF INPUTS
1 0=INPUT AND 1=OUTPUT ORIENTATED
1 0=CRS AND 1=VRS
4 0=DEA(MULTI-STAGE), 1=COST-DEA, 2=MALMQUIST-DEA,
3=DEA(1-STAGE), 4=DEA(2-STAGE)

F1_2.prn DATA FILE NAME
F1_2.out OUTPUT FILE NAME
338 NUMBER OF FIRMS
1 NUMBER OF TIME PERIODS
2 NUMBER OF OUTPUTS
1 NUMBER OF INPUTS
1 0=INPUT AND 1=OUTPUT ORIENTATED
1 0=CRS AND 1=VRS
4 0=DEA(MULTI-STAGE), 1=COST-DEA, 2=MALMQUIST-DEA,
3=DEA(1-STAGE), 4=DEA(2-STAGE)

F1_3.prn DATA FILE NAME
F1_3.out OUTPUT FILE NAME
338 NUMBER OF FIRMS
1 NUMBER OF TIME PERIODS
2 NUMBER OF OUTPUTS
1 NUMBER OF INPUTS
1 0=INPUT AND 1=OUTPUT ORIENTATED
1 0=CRS AND 1=VRS
4 0=DEA(MULTI-STAGE), 1=COST-DEA, 2=MALMQUIST-DEA,
3=DEA(1-STAGE), 4=DEA(2-STAGE)

F2_3.prn DATA FILE NAME
F2_3.out OUTPUT FILE NAME
338 NUMBER OF FIRMS
1 NUMBER OF TIME PERIODS
2 NUMBER OF OUTPUTS
1 NUMBER OF INPUTS
1 0=INPUT AND 1=OUTPUT ORIENTATED
1 0=CRS AND 1=VRS
4 0=DEA(MULTI-STAGE), 1=COST-DEA, 2=MALMQUIST-DEA,
3=DEA(1-STAGE), 4=DEA(2-STAGE)

```
##### Estimador Outliers SUDENE - Regiões imediatas (2022) #####
# Referências: #
# SIMAR, L. WILSON, P.W. Estimation and inference in two-stage, semi-parametric
models of production processes.#
# Journal of Econometrics. vol. 136, p?g. 31-64, 2007. #
# SIMAR, L. WILSON, P.W. Non-parametric tests of returns to scale. European
Journal of Operational Research. #
# vol. 139, p?g. 115-132, 2002. #
# BANKER, R. D. NATARAJAN, R. Statistical tests based on DEA efficiency
estimators.School of Management, #
# The University of Texas at Dallas, Richardson, USA. p?g. 1-23, 2008.
#
# SAMPAIO DE SOUSA, M. C.; STOSIC, B. Technical efficiency of the
Brazilian municipalities: correcting #
# nonparametric frontier measurements for outliers. Journal of Productivity Analysis,
v. 24, n. 2, p. 157-181, #
# 2005. #
# FERREIRA, C. M. C.; GOMES, A. P., Introdu??o ? An?lise Envolt?ria de Dados:
Teoria, Modelos e Aplica??es. #
# 2 ed. Vi?osa, MG: Editora UFV, 2020.
#
#####
#####

#### Remover dados existentes

rm(list=ls())

## Carregar Pacotes (? necess?rio instalar os pacotes previamente)

library(sfsmisc)
library(AER)
library(FEAR)
library(boot)

## A rotina utiliza os quatro pacotes. Favor atentar para poss?veis licen?as para
utiliza??o dos pacotes.

#### Ler a base de dados

Imediatas<-read.table("C:/Users/USUARIO/Desktop/imediatas_p3.prn", h=T)

## fne = financiamento FNE
## renda = rendimento do trabalho
## trabalho = n?mero de trabalhadores
```

```
#####
## Teste para detecção de outlier - leverage ##
#####

## Dada a configuração dos dados (apenas 1 insumo) ocorre uma inconsistência com a
rotina, por isso acrescentamos
## aqui um novo insumo fne2=fne para todas as dmus o que não altera nossa análise.

## Criar novas matrizes input e output

x=matrix(nrow=2,ncol=174)
x[1,]=Imediatas$fne
x[2,]=Imediatas$fne

y=matrix(nrow=2,ncol=174)
y[1,]=Imediatas$renda
y[2,]=Imediatas$trabalho

#####
#####

## Estimador Sampaio-Stosic ##### retornos variáveis #####

## Como são poucos dados, podemos construir o leverage sem utilizar o artifício das
## Bubbles de Sampaio-Stosic.

## Estimando os coeficientes de eficiência para as K dmus

deat <- dea(XOBS=x,YOBS=y, ORIENTATION=2)

## Temos de tirar a primeira observação j e estimar um dea para k-1 dmus, tirar a
segunda e fazer
## outro, a terceira e assim sucessivamente.

## construímos uma sequência da extensão dos nossos dados:
## Temos de obter também as matrizes. Uma para todos os insumos sem cada uma das
## observações, outra para os produtos:

## Primeiro desenvolve-se uma função para calcular a DEA original, com todas as
informações

first <- function(x, y, rts, orientation){
```

```

j <- seq(1:ncol(x))

rts <- seq(1:3)

orientation <- c(1,2)

saida1 <- NULL

dea_t <- dea(XOBS=x,YOBS=y, RTS = rts[1], ORIENTATION = orientation[2])

saida1 <- matrix(c(saida1,dea_t), nrow=length(j), ncol=1)

}

dea_t <- first(x,y, rts=1, orientation=2)

dea_t

## A segunda parte ? retirar as dmus uma a uma, formulando para tanto outra fun??o.
## A fun??o dea.out serve para identificar as observa??es outliers, retirando as dmus
uma a uma e estimando o dea. Essa
## fun??o requer os mesmos par?metros que a fun??o dea, por?m o primeiro par?metro
? o n?mero de observa??es neste caso = 208

dea.out <- function(j, x, y, rts, orientation){

j <- seq(1:ncol(x))

jn <- length(j)

rts <- seq(1:3)

orientation <- c(1,2)

saida2 <- NULL

xi <- 0
yi <- 0

for (i in 1:jn){
xi <- x[,-i]
yi <- y[,-i]
dea <- dea(XOBS=xi,YOBS=yi, RTS = rts[1], ORIENTATION = orientation[2])

saida2 <- c(saida2, dea)
}

```

```

saida2

}

dea <- dea.out(174,x,y, rts=1, orientation=2)

## Lembrar de transformar dea em uma matrix com 'jn-1' linhas e 'jn' colunas:

dea_j <- matrix(dea, nrow =173, ncol=174)

## Agora ? preciso colocar a fórmula presente em Sampaio de Sousa e Stosic (2005,
pp.162)
## Função construída Jackstrap (jackknife+bootstrap):
## Primeiro inseri-se os valores "0" nas diagonais [i,i] no lugar de cada observação a fim
de calcular o leverage:

orientation <- c(1,2)
rts <- seq(1:3)
comp1 <- 0
comp2 <- 0
comp3 <- 0
j <- seq(1:ncol(x))
jn <- length(j)
jnm <- length(j)-1
completar<- matrix(0, nrow=jn, ncol=jn)
saida3 <- NULL

## Fazemos a primeira e a última coluna da matriz de eficiência com os 0's enxertados.

comp1 <- c(completar[1,1], dea_j[1:jn-1,1])
comp3 <- c(dea_j[1:jn-1,jn], completar[jn,jn])

dea_j1 <- 0
dea_j2 <- 0
dea_j3 <- 0

## o miolo da matrix enxertada:

comp2 <- for (i in 2:jnm){

dea_j1 <- dea_j[1:(i-1),i]
dea_j2 <- completar[i,i]
dea_j3 <- dea_j[i:jnm,i]
vetor <- as.vector(c(dea_j1, dea_j2, dea_j3))

saida3 <- c(saida3,vetor)

```

```

}
comp2 <- saida3

dea_J <- matrix(c(comp1, comp2, comp3),nrow=jn, ncol=jn)

dea_J

## Sai a matriz com os índices de eficiência após a retirada de uma a uma das observações
## linhas = n observações, colunas = n observações.

## Agora a função da fórmula leverage propriamente, sem os bubbles.

jackknife.lev <- function (DEA_t,DEA_j)
{
  K <- length(j)

  dea_t <- dea(XOBS=x,YOBS = y, RTS = rts[1], ORIENTATION = orientation[2])

  theta <- c(dea_t)

  saida4 <- NULL

  lev <- for (i in 1:jn){
    theta_e <- dea_J[,i]
    square <- (theta_e-theta)^2
    soma <- sum(square[-i])
    leverage <- sqrt(soma/(K-1))
    saida4 <- c(saida4,leverage)
  }

  nleverage <- round(saida4, digits = 5)

  histograma <- hist(nleverage)
  kernel_lev <- density(nleverage)
  grafo <- lines(kernel_lev$x, kernel_lev$y, col="blue", lwd=1.5)

  ## Quais são os outliers? Usamos a regra de Bolso em Sampaio-Stosic lev >= 0.02

  outliers <- nleverage >= 0.02
  posicao <- matrix(c(nleverage,j), ncol=2, byrow=FALSE)

  ## Pedindo para listar apenas os outliers

```

```
nlev <- nleverage[outliers]
```

```
outliers
```

```
c(nlev,posicao[,2][posicao[,1]>=0.02])
```

```
}
```

```
jackknife.lev(dea_t, dea_J)
```

```
## Caso sejam identificados outliers eles devem ser extra?dos da amostra.
```

```
#####  
#####
```



ANEXO 3. ROTINAS DE PROGRAMAÇÃO EMPREGADAS NO PRODUTO 7

O Produto 7 da Meta 2, intitulado *Avaliação da efetividade da aplicação dos recursos do FNE*, lançou mão de ferramentas econométricas de avaliação de impacto para a obtenção dos resultados desejados. Em específico, foi utilizado o software estatístico *Stata*. Abaixo são apresentadas as linhas de programação que formam a rotina empregada para a análise empírica da efetividade do FNE.

```
* Abrindo a base de dados a ser utilizada (nível: municípios)
use "Base_fne_longa2_total_desenv.dta", clear
```

```
* Renomeando e rotulando variáveis
```

```
** Variáveis de identificação (ano, município e UF)
```

```
label var ano "2000 a 2018"
```

```
label var id_municipio "código ibge (7 dígitos)"
```

```
label var id_uf "unidade da federação"
```

```
** Variáveis de resultado (outcomes)
```

```
label var pib_def "pib deflacionado (IPCA dez/21)"
```

```
label var va_agropecuaria_def "va agropecuária (IPCA dez/21)"
```

```
label var va_industria_def "va indústria (IPCA dez/21)"
```

```
label var va_servicos_def "va serviços (IPCA dez/21)"
```

```
label var va_adespss_def "va setor público (IPCA dez/21)"
```

```
label var saldo "saldo de emprego"
```

```
label var idhm "índice de desen. humano municipal"
```

```
** Variáveis de controle
```

```
rename taxa_analfabetismo_25_mais x1
```

```
label var x1 "% da pop. 25+ analfabeta"
```

```
rename taxa_fundamental_25_mais x2
```

```
label var x2 "% da pop. 25+ com ens. fund. completo"
```

```
rename taxa_medio_25_mais x3
```

```
label var x3 "% da pop. 25+ com ens. médio incompleto?"
```

```
rename taxa_superior_25_mais x4
```

```
label var x4 "% da pop. 25+ com ens. sup. completo"
```

```
rename mortalidade_1 x5
```

```
label var x5 "taxa de mortalidade infantil até 1 ano"
```

```
rename pea x6
```

```
label var x6 "pop. 10+ economicamente ativa"
```

```
rename denspop x7
```

```
label var x7 "densidade populacional (hab/km2)"
```

```
rename taxa_agua_encanada x8
```

```

label var x8 "% pop. com acesso a água encanada"
rename gov x9
label var x9 "tamanho do governo (despesa corrente/pib)"
rename opcred_def x10
label var x10 "R$ total das operações de crédito (IPCA dez/21)"
rename agencias x11
label var x11 "nº de agências bancárias exceto BNB"
rename dist_capital x12
label var x12 "distância para a capital km"
rename regioao_metropolitana x13
label var x13 "=1 se região metropolitana"
rename med_rem_med_def x14
label var x14 "média do rendimento médio (IPCA dez/21)"
rename agen_bnb x15
label var x15 "nº de agências BNB"
gen                                primario                                =
va_agropecuaria_def/(va_agropecuaria_def+va_industria_def+va_servicos_def+va_adesps
s_def)
label var primario "participação % do setor primário sobre o VA"
gen                                secundario                                =
va_industria_def/(va_agropecuaria_def+va_industria_def+va_servicos_def+va_adesps
s_def)
label var secundario "participação % do setor secundário sobre o VA"
gen                                terciario                                =
va_servicos_def/(va_agropecuaria_def+va_industria_def+va_servicos_def+va_adesps
s_def)
label var terciario "participação % do setor terciário sobre o VA"
gen                                set_publico                                =
va_adesps_def/(va_agropecuaria_def+va_industria_def+va_servicos_def+va_adesps
s_def)
label var set_publico "participação % do setor público sobre o VA"

** Variáveis de tratamento

*** Semiárido
rename semiarido w_semi // Variável correta é a semiarido2
label var w_semi "=1 se semiárido"
label define semi 0 "fora" 1 "semiárido"
label values w_semi semi

*** Valor total contratado FNE
rename valor_def fne
label var fne "R$ total contratado (IPCA dez/21)"

*** Valor total contratado FNE por setor

```

```
rename (valor_defAGRICOLA valor_defAGROINDUSTRIA valor_defINDUSTRIA  
valor_defINFRAESTRUTURA valor_defPECUARIA valor_defTURISMO) (fnes1 fnes2  
fnes3 fnes4 fnes5 fnes6)
```

```
label var fnes1 "setor: agrícola"  
label var fnes2 "setor: agroindústria"  
label var fnes3 "setor: indústria"  
label var fnes4 "setor: infraestrutura"  
label var fnes5 "setor: pecuária"  
label var fnes6 "setor: turismo"
```

*** Valor total contratado FNE por programa

```
rename (valor_defPROCULTURA valor_defPRODESA valor_defPRODETEC  
valor_defPRODIN valor_defPROGER valor_defPROINFRA valor_defPROMICRO  
valor_defPROMOC valor_defPRONAF valor_defPROPEC valor_defPROTAD) (fnep01  
fnep02 fnep03 fnep04 fnep05 fnep06 fnep07 fnep08 fnep09 fnep10 fnep11)
```

```
label var fnep01 "programa: procultura"  
label var fnep02 "programa: prodesa"  
label var fnep03 "programa: prodetec"  
label var fnep04 "programa: prodin"  
label var fnep05 "programa: proger"  
label var fnep06 "programa: proinfra"  
label var fnep07 "programa: promicro"  
label var fnep08 "programa: promoc"  
label var fnep09 "programa: pronaf"  
label var fnep10 "programa: propec"  
label var fnep11 "programa: protad"
```

*** Valor total contratado FNE por finalidade

```
rename (valor_defCUSTEIO valor_defINVESTIMENTO valor_defOUTROS) (fnef1 fnef2  
fnef3)
```

```
label var fnef1 "finalidade: custeio"  
label var fnef2 "finalidade: investimento"  
label var fnef3 "finalidade: outros"
```

*** Valor total contratado FNE por objetivo

```
rename (valor_defAMPLIACAO valor_defAQUISICAO valor_defIMPLANTACAO  
valor_defMODERNIZACAO valor_defRELOCALIZACAO) (fneo1 fneo2 fneo3 fneo4  
fneo5)
```

```
label var fneo1 "objetivo: ampliação"  
label var fneo2 "objetivo: aquisição"  
label var fneo3 "objetivo: implantação"  
label var fneo4 "objetivo: modernização"  
label var fneo5 "objetivo: realocização"
```

*** Valor total contratado FNE por porte/tamanho

```

rename      (valor_defGRANDE      valor_defMEDIO      valor_defMEDIO_GRANDE
valor_defMINI_MICRO valor_defPEQUENO valor_defPEQUENO_MEDIO) (fnet1 fnet2
fnet3 fnet4 fnet5 fnet6)
label var fnet1 "porte: grande"
label var fnet2 "porte: médio"
label var fnet3 "porte: médio e grande"
label var fnet4 "porte: mini e micro"
label var fnet5 "porte: pequeno"
label var fnet6 "porte: pequeno e médio"

* Ajustando a base de dados

** Mantendo a base de dados entre 2000 e 2018 (disponibilizada BNB)
keep if ano >=2002 & ano<=2018

** Declarando dados em painel
xtset id_municipio ano

** Adicionando variáveis de controle TWFE Wooldridge (2021)

* Gerando variáveis de controle como desvios da média
*sum x1 if d4
*gen x1_dm4 = x1 - r(mean)
*sum x1 if d5
*gen x1_dm5 = x1 - r(mean)
*sum x1 if d6
*gen x1_dm6 = x1 - r(mean)
*sum x1 if w
*gen x1_dmw = x1 - r(mean)

*foreach var of varlist d0sum_w_prop-d18sum_w_prop {
*      sum x14 if `var'
*      gen x14_`var' = x14 - r(mean)
* }

** Logaritmizando as variáveis de resultado

gen pib_pc = pib_def/populacao
gen saldo_pc = saldo/populacao
gen va_agro_pc = va_agropecuaria_def/populacao
gen va_ind_pc = va_industria_def/populacao
gen va_ser_pc = va_servicos_def/populacao
gen va_ade_pc = va_adespss_def/populacao
gen obito_inf_pc = obitos_inf/ populacao
gen ln_obito_inf_pc = ln(obito_inf_pc)
gen obito_mat_pc = obitos_mat/ populacao
gen ln_obito_mat_pc = ln(obito_mat_pc)

```

```

gen obito_ext_pc = obitos_ext/ populacao
gen ln_obito_ext_pc = ln(obito_ext_pc)

* Gerando variáveis de tratamento

** Proporção de recursos do FNE (Total) em relação ao PIB (staggered)

gen prop = fne/pib_def
sum prop
gen sup = r(mean) + 2*r(sd)
keep if prop >=0 & prop <=sup
recode prop (.=0)

** Gerando variáveis para o cálculo do efeito do tratamento (Average Treatment on the
Treated - ATT)

foreach var of varlist pib_pc-va_ade_pc {
    gen ln_`var' = ln(`var')
}

** Gerando variáveis de controle em relação aos recursos recebidos

foreach var of varlist fne-fnet6 {
    gen `var'_pc = `var'/populacao/1000
    recode `var'_pc (.=0)
}

** Definindo o tratamento (x=1 a partir do momento em que o município recebe proporção
maior do que a média mais 1 desvio-padrão)

by ano, sort: egen m_proporcao=mean(prop)
by ano, sort: egen d_proporcao=sd(prop)
sum prop
gen w_prop = 1 if prop >= m_proporcao + d_proporcao
by id_municipio (ano), sort: replace w_prop = w_prop[_n-1] if w_prop >=.
recode w_prop (.=0)

xtset id_municipio ano

** Definindo a proporção de recursos do FNE (por setor) em relação ao PIB (staggered)

gen props1 = fnes1/pib_def
gen props2 = fnes2/pib_def
gen props3 = fnes3/pib_def
gen props4 = fnes4/pib_def
gen props5 = fnes5/pib_def
gen props6 = fnes6/pib_def

```

```
recode props1-props6 (.=0)
sum props1-props6
```

** Definindo o tratamento (x=1 a partir do momento em que o município recebe proporção maior do que a média mais 1 desvio-padrão)

```
foreach var of varlist props1-props6 {
sum `var'
gen w_`var' = 1 if `var' > r(mean) + r(sd)
by id_municipio (ano), sort: replace w_`var' = w_`var'[_n-1] if w_`var' >=.
recode w_`var' (.=0)
}
```

** Definindo a proporção de recursos do FNE (por finalidade) em relação ao PIB (staggered)

```
gen propf1 = fnef1/pib_def
gen propf2 = fnef2/pib_def
gen propf3 = fnef3/pib_def
recode propf1-propf3 (.=0)
sum propf1-propf3
```

** Definindo o tratamento (x=1 a partir do momento em que o município recebe proporção maior do que a média mais 1 desvio-padrão)

```
foreach var of varlist propf1-propf3 {
sum `var'
gen w_`var' = 1 if `var' > r(mean) + r(sd)
by id_municipio (ano), sort: replace w_`var' = w_`var'[_n-1] if w_`var' >=.
recode w_`var' (.=0)
}
```

** Criando as dummies para os setores que receberam recursos do FNE (Agrícola e Agropecuária; Agroindústria; Indústria; Infraestrutura; e Serviços (Turismo))

```
gen d_agp = .
replace d_agp = 1 if (props1 > 0 | props5 > 0)
replace d_agp = 0 if d_agp == .
gen d_ag_ind = .
replace d_ag_ind = 1 if props2 > 0
replace d_ag_ind = 0 if d_ag_ind == .
gen d_ind = .
replace d_ind = 1 if props3 > 0
replace d_ind = 0 if d_ind == .
gen d_infra = .
replace d_infra = 1 if props4 > 0
replace d_infra = 0 if d_infra == .
gen d_serv = .
```

```
replace d_serv = 1 if props6 > 0  
replace d_serv = 0 if d_serv == .
```

** Gerando dummies (variáveis binárias) para verificar o não recebimento de recursos relacionados a nenhum setor

```
gen total = (d_agp + d_ag_ind + d_ind + d_infra + d_serv)
```

```
gen d_nenhum = .  
replace d_nenhum = 1 if total > 0  
replace d_nenhum = 0 if total == 0
```

** Cruzando os setores de agropecuária, agroindústria, indústria, infraestrutura e serviços

```
gen d_agp_infra = .  
replace d_agp_infra = 1 if d_agp == 1 & d_infra == 1  
replace d_agp_infra = 0 if d_agp_infra == .
```

```
gen d_ag_ind_infra = .  
replace d_ag_ind_infra = 1 if d_ag_ind == 1 & d_infra == 1  
replace d_ag_ind_infra = 0 if d_ag_ind_infra == .
```

```
gen d_ind_infra = .  
replace d_ind_infra = 1 if d_ind == 1 & d_infra == 1  
replace d_ind_infra = 0 if d_ind_infra == .
```

```
gen d_serv_infra = .  
replace d_serv_infra = 1 if d_serv == 1 & d_infra == 1  
replace d_serv_infra = 0 if d_serv_infra == .
```

*** As dummies criadas não fizeram muito sentido no modelo estimado pois todas apresentaram parâmetros negativos.

*** No entanto, elas servem para justificar a ausência de emprego quando fica possível visualizar que uma pequena

*** parcela dos municípios receberam recursos da infraestrutura e da indústria.

```
tab d_ind  
tab d_infra
```

*** Além disso, é possível verificar que os recursos destinados para investimentos e infraestrutura estariam

*** concentrados nos municípios maiores e mais próximos das capitais.

```
pwcorr fnes1 fnes2 fnes3 fnes4 fnes5 fnes6 populacao x7 x12, sig
```

** Criando uma variável de crise que assume valor 0 antes de 2014 e valor 1 para o ano de 2014 em diante

*** Com isso é possível verificar que o saldo de empregos está negativa e significamente relacionado com a crise.

*** Contudo, observa-se que a crise teria apresentado correlação negativa com a produção e, conseqüentemente, com os repasses do FNE.

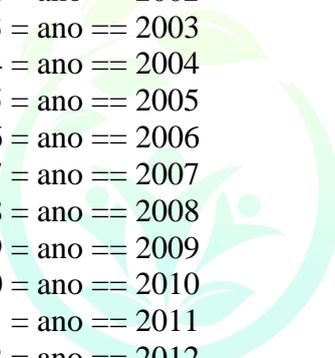
```
gen crise = .
replace crise =1 if ano >= 2014
replace crise = 0 if ano < 2014
```

```
pwcorr saldo_pc pib_pc fne_pc crise , sig
```

* Gerando variáveis para estimação dos estimadores TWFE Wooldridge (2021)

** Gerando dummies de ano (para anos em que houve tratamento)

```
gen f2000 = ano == 2000
gen f2001 = ano == 2001
gen f2002 = ano == 2002
gen f2003 = ano == 2003
gen f2004 = ano == 2004
gen f2005 = ano == 2005
gen f2006 = ano == 2006
gen f2007 = ano == 2007
gen f2008 = ano == 2008
gen f2009 = ano == 2009
gen f2010 = ano == 2010
gen f2011 = ano == 2011
gen f2012 = ano == 2012
gen f2013 = ano == 2013
gen f2014 = ano == 2014
gen f2015 = ano == 2015
gen f2016 = ano == 2016
gen f2017 = ano == 2017
gen f2018 = ano == 2018
```



IPPDS
Instituto de Políticas Públicas e
Desenvolvimento Sustentável

** Gerar o tempo de exposição ao tratamento em anos

```
*egen wsum = sum(w), by(id)
```

```
foreach var of varlist w_prop w_props1-w_props6 w_propf1-w_propf3 {
    egen sum_`var'=sum(`var'), by(id_municipio)
}
```

* Gerar dummies para grupos de tratamento, de acordo com o tempo de exposição ao tratamento

```
*gen d4 = wsum == 3
```

```

*gen d5 = wsum == 2
*gen d6 = wsum == 1

*foreach var of varlist sum_w_valor_def-sum_w_valor_defPEQUENO_MEDIO {
*   foreach num of numlist 1/19 {
*   gen d`num`var' = `var' == `num' // d1=1 (1 ano de exposição ao tratamento, entrada em
2018)
*   } // d19=1 (19 anos de exposição ao tratamento, entrada em 2000)
* }

foreach var of varlist sum_w_prop-sum_w_propf3 {
gen d0`var' = `var' == 19 // d0=1 (19 anos de exposição ao tratamento, entrada em 2000)
gen d1`var' = `var' == 18 // d1=1 (18 anos de exposição ao tratamento, entrada em 2001)
gen d2`var' = `var' == 17 // d2=1 (17 anos de exposição ao tratamento, entrada em 2002)
gen d3`var' = `var' == 16 // d3=1 (16 anos de exposição ao tratamento, entrada em 2003)
gen d4`var' = `var' == 15 // d4=1 (15 anos de exposição ao tratamento, entrada em 2004)
gen d5`var' = `var' == 14 // d5=1 (14 anos de exposição ao tratamento, entrada em 2005)
gen d6`var' = `var' == 13 // d6=1 (13 anos de exposição ao tratamento, entrada em 2006)
gen d7`var' = `var' == 12 // d7=1 (12 anos de exposição ao tratamento, entrada em 2007)
gen d8`var' = `var' == 11 // d8=1 (11 anos de exposição ao tratamento, entrada em 2008)
gen d9`var' = `var' == 10 // d9=1 (10 anos de exposição ao tratamento, entrada em 2009)
gen d10`var' = `var' == 9 // d10=1 (9 anos de exposição ao tratamento, entrada em 2010)
gen d11`var' = `var' == 8 // d11=1 (8 anos de exposição ao tratamento, entrada em 2011)
gen d12`var' = `var' == 7 // d12=1 (7 anos de exposição ao tratamento, entrada em 2012)
gen d13`var' = `var' == 6 // d13=1 (6 anos de exposição ao tratamento, entrada em 2013)
gen d14`var' = `var' == 5 // d14=1 (5 anos de exposição ao tratamento, entrada em 2014)
gen d15`var' = `var' == 4 // d15=1 (4 anos de exposição ao tratamento, entrada em 2015)
gen d16`var' = `var' == 3 // d16=1 (3 anos de exposição ao tratamento, entrada em 2016)
gen d17`var' = `var' == 2 // d17=1 (2 anos de exposição ao tratamento, entrada em 2017)
gen d18`var' = `var' == 1 // d18=1 (1 ano de exposição ao tratamento, entrada em 2018)
}

* Gerando variáveis para estimação dos estimadores duplo robusto de Callaway and
Sant'Anna (2020)

** Gerando variável que identifica o primeiro ano de exposição ao tratamento
*by id (year), sort: egen year_treat = min(w*year) if w != 0
*by id (year), sort: egen first_treat = max(year_treat)
*recode first_treat (. = 0)
*drop year_treat

foreach var of varlist w_prop w_props1-w_props6 w_propf1-w_propf3 {
by id_municipio (ano), sort: egen yt_`var' = min(`var'*ano) if `var' != 0 // yt: ano de
tratamento
by id_municipio (ano), sort: egen ft_`var' = max(yt_`var') // ft: primeiro ano de tratamento
recode ft_`var' (. = 0)
drop yt_`var'

```

```

}

* Gerando variáveis para a estimação do Event Study

** Gerando a variável TimeToTreat para estimar o event study
*gen TimeToTreat = year - first_treat
*replace TimeToTreat =. if TimeToTreat >=10

foreach var of varlist w_prop w_props1-w_props6 w_propf1-w_propf3 {
    gen ttt_`var' = ano - ft_`var'
    replace ttt_`var' =. if ft_`var' ==0 // ttt: tempo para o tratamento
}

* Estimando o modelo - ATT Constante (médio):

** TWFE Wooldridge (2021)

*Exemplo:
*xrtreg logy w f2014 f2015 f2016, fe vce(cluster id)

*** Tratamento: fne total (w_prop)
foreach y of varlist ln_pib_pc-ln_va_ade_pc {
xrtreg `y' w_prop fne_pc ///
    f2002-f2018 ///
    primario secundario terciario x14 semiarido_2 pz_carencia tx_juros ///
    , fe vce(cluster id_municipio)
}

*** Tratamento: fne por setor (w_props1-w_props6)
foreach y of varlist ln_pib_pc-ln_va_ade_pc {
xrtreg `y' w_props1-w_props6 fnes1_pc-fnes6_pc ///
    f2002-f2018 ///
    primario secundario terciario x14 semiarido_2 pz_carencia tx_juros ///
    d_agp d_ag_ind d_ind d_infra d_serv ///
    , fe vce(cluster id_municipio)
}

*** Tratamento: fne por finalidade (w_propf1-w_propf3)
foreach y of varlist ln_pib_pc-ln_va_ade_pc {
xrtreg `y' w_propf1-w_propf3 fnef1_pc-fnef3_pc ///
    f2002-f2018 ///
    primario secundario terciario x14 semiarido_2 pz_carencia tx_juros ///
    d_agp d_ag_ind d_ind d_infra d_serv ///
    , fe vce(cluster id_municipio)
}

** Event Study

```

```

*Exemplo:
*eventdd logy ///
*   f2014 f2015 f2016 ///
*       , timevar(TimeToTreat) method(fe) cluster(id)

*** Tratamento: fne total (w_prop)
foreach y of varlist ln_pib_pc-ln_va_ade_pc {
eventdd `y' ///
    primario secundario terciario x14 semiarido_2 pz_carencia tx_juros ///
    f2000-f2018 ///
    fne_pc ///
    , inrange leads(3) lags(5) timevar(ttt_w_prop) method(fe) cluster(id_municipio)
}

*** Tratamento: fne por setor (w_props1-w_props6)
foreach y of varlist ln_pib_pc-ln_va_ade_pc {
    foreach x of varlist ttt_w_props1-ttt_w_props6 { // props6 não estima leads
eventdd `y' ///
    primario secundario terciario x14 semiarido_2 pz_carencia tx_juros ///
    f2000-f2018 ///
    fnes1_pc-fnes6_pc ///
    , inrange leads(3) lags(5) timevar(`x') method(fe) cluster(id_municipio)
    }
}

*** Tratamento: fne por finalidade (w_propf1-w_propf3)
foreach y of varlist ln_pib_pc-ln_va_ade_pc {
    foreach x of varlist ttt_w_propf1-ttt_w_propf3 {
eventdd `y' ///
    primario secundario terciario x14 semiarido_2 pz_carencia tx_juros ///
    f2000-f2018 ///
    fnef1_pc-fnef3_pc ///
    , inrange leads(3) lags(5) timevar(`x') method(fe) cluster(id_municipio)
    }
}

** DRDID Callaway e Sant'Anna (2021)

*Exemplo:
*csdid logy, ivar(id) time(year) gvar(first_treat) method(drimp) wboot agg(simple)
*csdid logy, ivar(id) time(year) gvar(first_treat) method(drimp) wboot agg(simple) notyet //
Opção notyet

*** Tratamento: fne total (w_prop)
foreach y of varlist ln_pib_pc-ln_va_ade_pc {
csdid `y' /// // com controles

```

```

fne_pc primario secundario terciario x14 semiarido_2 pz_carencia tx_juros ///
, ivar(id_municipio) time(ano) gvar(ft_w_prop) method(drimp) wboot notyet
estat all
estat event, window(-3 5)
csdid_plot
}

*** Tratamento: fne por setor (w_props1-w_props6)
foreach y of varlist ln_pib_pc-ln_va_ade_pc {
    foreach x of varlist ft_w_props1-ft_w_props6 {
csdid `y' /// // com controles
    fnes1_pc-fnes6_pc primario secundario terciario x14 semiarido_2 pz_carencia tx_juros
///
    , ivar(id_municipio) time(ano) gvar(`x') method(drimp) wboot notyet
estat all
estat event, window(-3 5)
csdid_plot
    }
}

*** Tratamento: fne por finalidade (w_propf1-w_propf3)
foreach y of varlist ln_pib_pc-ln_va_ade_pc {
    foreach x of varlist ft_w_propf1-ft_w_propf3 {
csdid `y' /// // com controles
    fnef1_pc-fnef3_pc primario secundario terciario x14 semiarido_2 pz_carencia tx_juros
///
    , ivar(id_municipio) time(ano) gvar(`x') method(drimp) wboot notyet
estat all
estat event, window(-3 5)
csdid_plot
    }
}

* Estimando o modelo - ATT variando ao longo do tempo:

** TWFE Wooldridge (2021)

*Exemplo:
*xtreg logy c.w#c.f2014 c.w#c.f2015 c.w#c.f2016 ///
*      f2014 f2015 f2016, fe vce(cluster id)

*** Tratamento: fne total (w_prop)
foreach y of varlist ln_pib_pc-ln_va_ade_pc {
xtreg `y' c.w_prop#c.(f2000-f2018) fne_pc ///
    f2000-f2018 ///
    primario secundario terciario x14 semiarido_2 pz_carencia tx_juros ///
    , fe vce(cluster id_municipio)

```

```

}

*** Tratamento: fne por setor (w_props1-w_props6)
foreach y of varlist ln_pib_pc-ln_va_ade_pc {
xtreg `y' c.(w_props1-w_props6)#c.(f2000-f2018) fnes1_pc-fnes6_pc ///
    f2000-f2018 ///
        primario secundario terciario x14 semiarido_2 pz_carencia tx_juros ///
        , fe vce(cluster id_municipio)
}

*** Tratamento: fne por finalidade (w_propf1-w_propf3)
foreach y of varlist ln_pib_pc-ln_va_ade_pc {
xtreg `y' c.(w_propf1-w_propf3)#c.(f2000-f2018) fnef1_pc-fnef3_pc ///
    f2000-f2018 ///
        primario secundario terciario x14 semiarido_2 pz_carencia tx_juros ///
        , fe vce(cluster id_municipio)
}

* Estimando o modelo - ATT variando por grupo de exposição ao tratamento:

** TWFE Wooldridge (2021)

*Exemplo:
*xtreg logy c.w#c.d4 c.w#c.d5 c.w#c.d6 ///
*    f2014 f2015 f2016, fe vce(cluster id)

*** Tratamento: fne total (w_prop)
foreach y of varlist ln_pib_pc-ln_va_ade_pc {
xtreg `y' c.w_prop#c.(d0sum_w_prop-d18sum_w_prop) fne_pc ///
    f2000-f2018 ///
        primario secundario terciario x14 semiarido_2 pz_carencia tx_juros ///
        , fe vce(cluster id_municipio)
}

*** Tratamento: fne por setor (w_props1-w_props6)
foreach y of varlist ln_pib_pc-ln_va_ade_pc {
xtreg `y' fnes1_pc-fnes6_pc ///
    c.(w_props1)#c.(d0sum_w_props1-d18sum_w_props1) ///
    c.(w_props2)#c.(d0sum_w_props2-d18sum_w_props2) ///
    c.(w_props3)#c.(d0sum_w_props3-d18sum_w_props3) ///
    c.(w_props4)#c.(d0sum_w_props4-d18sum_w_props4) ///
    c.(w_props5)#c.(d0sum_w_props5-d18sum_w_props5) ///
    c.(w_props6)#c.(d0sum_w_props6-d18sum_w_props6) ///
    f2000-f2018 ///
        primario secundario terciario x14 semiarido_2 pz_carencia tx_juros ///
        , fe vce(cluster id_municipio)
}

```

```

*** Tratamento: fne por finalidade (w_propf1-w_propf3)
foreach y of varlist ln_pib_pc-ln_va_ade_pc {
xtreg `y' fnef1_pc-fnef3_pc ///
    c.(w_propf1)#c.(d0sum_w_propf1-d18sum_w_propf1) ///
    c.(w_propf2)#c.(d0sum_w_propf2-d18sum_w_propf2) ///
    c.(w_propf3)#c.(d0sum_w_propf3-d18sum_w_propf3) ///
    f2000-f2018 ///
    primario secundario terciario x14 semiarido_2 pz_carencia tx_juros ///
    , fe vce(cluster id_municipio)
}

```

* Estimando o modelo - ATT variando por grupo e ao longo o tempo:

*Exemplo:

```

*xtreg logy c.d4#c.f2014 c.d4#c.f2015 c.d4#c.f2016 ///
*    c.d5#c.f2015 c.d5#c.f2016 ///
*    c.d6#c.f2016 ///
*    f2014 f2015 f2016, fe vce(cluster id)

```

```

*** Tratamento: fne total (w_prop)
foreach y of varlist ln_pib_pc-ln_va_ade_pc {
xtreg `y' c.(f2000-f2018)#c.(d0sum_w_prop-d18sum_w_prop) fne_pc ///
    f2000-f2018 ///
    primario secundario terciario x14 semiarido_2 pz_carencia tx_juros ///
    , fe vce(cluster id_municipio)
}

```

```

*** Tratamento: fne por setor (w_props1-w_props6)
foreach y of varlist ln_pib_pc-ln_va_ade_pc {
xtreg `y' fnes1_pc-fnes6_pc ///
    c.(f2000-f2018)#c.(d0sum_w_props1-d18sum_w_props1) ///
    c.(f2000-f2018)#c.(d0sum_w_props2-d18sum_w_props2) ///
    c.(f2000-f2018)#c.(d0sum_w_props3-d18sum_w_props3) ///
    c.(f2000-f2018)#c.(d0sum_w_props4-d18sum_w_props4) ///
    c.(f2000-f2018)#c.(d0sum_w_props5-d18sum_w_props5) ///
    c.(f2000-f2018)#c.(d0sum_w_props6-d18sum_w_props6) ///
    f2000-f2018 ///
    primario secundario terciario x14 semiarido_2 pz_carencia tx_juros ///
    , fe vce(cluster id_municipio)
}

```

```

*** Tratamento: fne por finalidade (w_propf1-w_propf3)
foreach y of varlist ln_pib_pc-ln_va_ade_pc {
xtreg `y' fnef1_pc-fnef3_pc ///
    c.(f2000-f2018)#c.(d0sum_w_propf1-d18sum_w_propf1) ///
    c.(f2000-f2018)#c.(d0sum_w_propf2-d18sum_w_propf2) ///

```

```
c.(f2000-f2018)#c.(d0sum_w_propf3-d18sum_w_propf3) ///
f2000-f2018 ///
    primario secundario terciario x14 semiarido_2 pz_carencia tx_juros ///
    , fe vce(cluster id_municipio)
}
```

* Definindo uma regressão quantílica para dados em painel com efeitos fixos COM CONTROLES

***O estimador xtgqreg estima regressão quantílica com efeitos fixos utilizando
 ***o método de Machado and Santos Silva (2019).
 ***Este método considera a estimação de parâmetros de localização e
 ***escala. Os parâmetros de localização representam os parâmetros no ponto médio.
 ***Esses parâmetros mudariam a distribuição para esquerda ou para a direita.
 ***Os de escala comprimem ou estendem toda a distribuição.
 ***Chamaremos ao coeficiente escalar α_i (τ) $\equiv \alpha_i + \delta_{iq}$ (τ) o efeito fixo do
 ***quantil- τ para o indivíduo i , ou o efeito distributivo em τ .
 ***O efeito distributivo difere do efeito fixo usual por não ser, em geral,
 ***uma mudança de localização apenas. Ou seja, o efeito distributivo representa o
 ***efeito de características individuais invariantes no tempo que, como outras
 ***variáveis, podem ter impactos diferentes em regiões diferentes da distribuição
 ***condicional de Y . O fato de $\int_0^1 q(\tau) d\tau = 0$ implica que α_i
 ***pode ser interpretado como o efeito médio para o indivíduo i .
 ***Os efeitos individuais permitem que os efeitos individuais afetem toda a
 ***distribuição ao invés de serem apenas deslocadores de localização.

***Quando apenas um quantil é estimado, xtgqreg salva os seguintes resultados em

* e ():

* Escalares

* e (N) número de observações

* Macros

* e (cmd) xtgqreg

* e (depvar) nome da variável dependente

* e (propriedades) b V

* Matrizes

* e (b) vetor de coeficiente

* e (V) matriz de covariância

* e (q) quantil estimado dos erros escalonados

* e (b_location) vetor de coeficientes de localização

* e (V_location) matriz de covariância de coeficientes de localização

* Vetor de coeficientes de escala e (b_scale)

* e (V_scale) matriz de covariância de coeficientes de escala

*Referencia: Machado, J.A.F. and Santos Silva, J.M.C. (2019), Quantiles via Moments. Journal of Econometrics, 213(1), pp. 145-173.

xtqreg lpibm lkf lkh [if] [in] [, options] /// é um exemplo

*xtqreg é a sintaxe para o painel quantílico com efeitos fixos

*lpibm é a variável dependente logaritmo do pib percapta

*lkf e lkh são as variáveis independentes, logaritmo do capital físico e do capital humano, respectivamente.

*Dentro das opções temos as possibilidades abaixo.

* quantile(#[#[# ...]]) estima # quantil; o padrão é a mediana 0,5

* id especifica a variável que define o painel

* ls exhibe as estimativas dos parâmetros de localização e escala

* savefe(newvar) salva os efeitos fixos para cada quantil estimado

* predict(newvar) salva os valores ajustados para cada quantil estimado

*É importante ressaltar que as variáveis independentes podem conter variáveis

*dummies e que o estimador não permite a inserção de pesos

*No modelo a seguir estima-se o impacto dos empréstimos do FNE nos municípios atendidos

*descontando a heterogeneidade individual dos municípios. Isso significa que os

* parâmetros estimados levam em conta os efeitos de w_prop, w_props e w_propf sobre o ln_pib_pc ao

*longo dos quantis, independentemente das características individuais dos mesmos.

*Portanto, servem para isolar os efeitos das variáveis independentes sobre a

*variável dependente ao longo de toda sua distribuição. Este modelo permite com

*que os coeficientes respondam à pergunta: como as concessões de empréstimos do

*FNE afetam a atividade econômica nos diferentes quantis, controlando pelas

*características individuais dos municípios?

Você irá perceber que a estimação vai ser precedida do seguinte aviso**

***"WARNING: .15634283% of the fitted values of the scale function are not positive"

***Pesquisando na internet, no fórum do Stata, o Prórpio Santos Silva,

***questionado sobre isso, indica que: "I do not know why the message appear

*** but with such small percentage I would not worry about."

* Estimando a regressão quantílica para as variáveis de resultado

** Variável de resultado: PIB per capita

```
xtqreg ln_pib_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14 semiarido_2
pz_carencia tx_juros , i(id_municipio) q(0.10)
```

xtqreg ln_pib_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14 semiarido_2
pz_carencia tx_juros , i(id_municipio) q(0.25)

xtqreg ln_pib_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14 semiarido_2
pz_carencia tx_juros , i(id_municipio) q(0.50)

xtqreg ln_pib_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14 semiarido_2
pz_carencia tx_juros , i(id_municipio) q(0.75)

xtqreg ln_pib_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14 semiarido_2
pz_carencia tx_juros , i(id_municipio) q(0.90)

** Variável de resultado: Saldo de empregos

xtqreg ln_saldo_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.10)

xtqreg ln_saldo_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.25)

xtqreg ln_saldo_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.50)

xtqreg ln_saldo_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.75)

xtqreg ln_saldo_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.90)

** Variável de resultado: Valor adicionado - Agropecuária

xtqreg ln_va_agro_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.10)

xtqreg ln_va_agro_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.25)

xtqreg ln_va_agro_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.50)

xtqreg ln_va_agro_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.75)

xtqreg ln_va_agro_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.90)

** Variável de resultado: Valor adicionado - Indústria

xtqreg ln_va_ind_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.10)

xtqreg ln_va_ind_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.25)

xtqreg ln_va_ind_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.50)

xtqreg ln_va_ind_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.75)

xtqreg ln_va_ind_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.90)

** Variável de resultado: Valor adicionado - Serviços

xtqreg ln_va_ser_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.10)

xtqreg ln_va_ser_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.25)

xtqreg ln_va_ser_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.50)

xtqreg ln_va_ser_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.75)

xtqreg ln_va_ser_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.90)

** Variável de resultado: Valor adicionado - Administração Pública

xtqreg ln_va_ade_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.10)

xtqreg ln_va_ade_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.25)

xtqreg ln_va_ade_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.50)

xtqreg ln_va_ade_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.75)

```
xtqreg ln_va_ade_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.90)
```

** Obtendo os parâmetros de cada quantil para construção dos gráficos no Excel

```
xtqreg ln_pib_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14 semiarido_2
pz_carencia tx_juros , i(id_municipio) q(0.05(0.05)0.95)
```

```
xtreg ln_pib_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14 semiarido_2
pz_carencia tx_juros , fe vce(bootstrap)
```

```
xtqreg ln_saldo_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.05(0.05)0.95)
```

```
xtreg ln_saldo_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , fe vce(bootstrap)
```

```
xtqreg ln_va_agro_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.05(0.05)0.95)
```

```
xtreg ln_va_agro_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , fe vce(bootstrap)
```

```
xtqreg ln_va_ind_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.05(0.05)0.95)
```

```
xtreg ln_va_ind_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , fe vce(bootstrap)
```

```
xtqreg ln_va_ser_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.05(0.05)0.95)
```

```
xtreg ln_va_ser_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , fe vce(bootstrap)
```

```
xtqreg ln_va_ade_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , i(id_municipio) q(0.05(0.05)0.95)
```

```
xtreg ln_va_ade_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2 pz_carencia tx_juros , fe vce(bootstrap)
```

* Estimando a regressão quantílica para indicadores de nível de desenvolvimento

** Óbitos infantis per capita

xtqreg obito_inf_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2, i(id_municipio) q(0.10)

xtqreg obito_inf_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2, i(id_municipio) q(0.25)

xtqreg obito_inf_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2, i(id_municipio) q(0.50)

xtqreg obito_inf_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2, i(id_municipio) q(0.75)

xtqreg obito_inf_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2, i(id_municipio) q(0.90)

** Óbitos maternos per capita

xtqreg obito_mat_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2, i(id_municipio) q(0.10)

xtqreg obito_mat_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2, i(id_municipio) q(0.25)

xtqreg obito_mat_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2, i(id_municipio) q(0.50)

xtqreg obito_mat_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2, i(id_municipio) q(0.75)

xtqreg obito_mat_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2, i(id_municipio) q(0.90)

** Óbitos por causas externas per capita

xtqreg obito_ext_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2, i(id_municipio) q(0.10)

xtqreg obito_ext_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2, i(id_municipio) q(0.25)

xtqreg obito_ext_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2, i(id_municipio) q(0.50)

xtqreg obito_ext_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14
semiarido_2, i(id_municipio) q(0.75)

```
xtqreg obito_ext_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14  
semiarido_2, i(id_municipio) q(0.90)
```

** Taxa de distorção idade-série no Ensino Médio

```
xtqreg tdi_em w_prop fne_pc f2000-f2018 primario secundario terciario x14 semiarido_2,  
i(id_municipio) q(0.10)
```

```
xtqreg tdi_em w_prop fne_pc f2000-f2018 primario secundario terciario x14 semiarido_2,  
i(id_municipio) q(0.25)
```

```
xtqreg tdi_em w_prop fne_pc f2000-f2018 primario secundario terciario x14 semiarido_2,  
i(id_municipio) q(0.50)
```

```
xtqreg tdi_em w_prop fne_pc f2000-f2018 primario secundario terciario x14 semiarido_2,  
i(id_municipio) q(0.75)
```

```
xtqreg tdi_em w_prop fne_pc f2000-f2018 primario secundario terciario x14 semiarido_2,  
i(id_municipio) q(0.90)
```

** Obtendo os parâmetros de cada quantil para construção dos gráficos no Excel

```
xtqreg ln_obito_inf_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14  
semiarido_2, i(id_municipio) q(0.05(0.05)0.95)
```

```
xtreg ln_obito_inf_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14  
semiarido_2, fe vce(bootstrap)
```

```
xtqreg ln_obito_mat_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14  
semiarido_2, i(id_municipio) q(0.05(0.05)0.95)
```

```
xtreg ln_obito_mat_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14  
semiarido_2, fe vce(bootstrap)
```

```
xtqreg ln_obito_ext_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14  
semiarido_2, i(id_municipio) q(0.05(0.05)0.95)
```

```
xtreg ln_obito_ext_pc w_prop fne_pc f2000-f2018 primario secundario terciario x14  
semiarido_2, fe vce(bootstrap)
```

```
xtqreg tdi_em w_prop fne_pc f2000-f2018 primario secundario terciario x14 semiarido_2,  
i(id_municipio) q(0.05(0.05)0.95)
```

```
xtreg tdi_em w_prop fne_pc f2000-f2018 primario secundario terciario x14 semiarido_2, fe  
vce(bootstrap)
```

* Análise descritiva da base de dados

** Análise de estatísticas descritivas para subsidiar os resultados de efetividade

*** Transformando a variável de recursos do FNE (divisão por mil reais)

```
xtset id_municipio ano
```

```
gen mfne=fne/1000
```

```
replace mfne=0 if mfne==.
```

*** Criando a variável estado (UF)

```
gen estado = .
```

```
tostring estado, replace
```

```
replace estado = "Maranhão" if id_uf == 21
```

```
replace estado = "Piauí" if id_uf == 22
```

```
replace estado = "Ceará" if id_uf == 23
```

```
replace estado = "Rio Grande do Norte" if id_uf == 24
```

```
replace estado = "Paraíba" if id_uf == 25
```

```
replace estado = "Pernambuco" if id_uf == 26
```

```
replace estado = "Alagoas" if id_uf == 27
```

```
replace estado = "Sergipe" if id_uf == 28
```

```
replace estado = "Bahia" if id_uf == 29
```

```
replace estado = "Minas Gerais" if id_uf == 31
```

```
replace estado = "Espírito Santo" if id_uf == 32
```

*** Definindo o volume médio de recursos por municípios considerando os estados ao longo dos anos

```
tabstat mfne if ano==2006, by(estado)
```

```
tabstat mfne if ano==2007, by(estado)
```

```
tabstat mfne if ano==2008, by(estado)
```

```
tabstat mfne if ano==2009, by(estado)
```

```
tabstat mfne if ano==2010, by(estado)
```

```
tabstat mfne if ano==2011, by(estado)
```

```
tabstat mfne if ano==2012, by(estado)
```

```
tabstat mfne if ano==2013, by(estado)
```

```
tabstat mfne if ano==2014, by(estado)
```

```
tabstat mfne if ano==2015, by(estado)
```

```
tabstat mfne if ano==2016, by(estado)
```

```
tabstat mfne if ano==2017, by(estado)
```

```
tabstat mfne if ano==2018, by(estado)
```

```
by estado ano, sort: sum mfne
```

*** Definindo o volume de recursos por municípios considerando a região semiárida ao longo dos anos

```
tabstat mfne if ano==2006, by(w_semi)
tabstat mfne if ano==2007, by(w_semi)
tabstat mfne if ano==2008, by(w_semi)
tabstat mfne if ano==2009, by(w_semi)
tabstat mfne if ano==2010, by(w_semi)
tabstat mfne if ano==2011, by(w_semi)
tabstat mfne if ano==2012, by(w_semi)
tabstat mfne if ano==2013, by(w_semi)
tabstat mfne if ano==2014, by(w_semi)
tabstat mfne if ano==2015, by(w_semi)
tabstat mfne if ano==2016, by(w_semi)
tabstat mfne if ano==2017, by(w_semi)
tabstat mfne if ano==2018, by(w_semi)
```

```
by w_semi ano, sort: sum mfne
```

*** Volume de recursos por designação de tratamento (tratados vs não tratados)

```
by w_prop ano, sort: sum mfne
```

*** Tranformando a variável de recursos do FNE por setor (divisão por mil reais)

```
gen agricola = fnes1/1000
replace agricola = 0 if agricola ==.
gen agroindustria = fnes2/1000
replace agroindustria=0 if agroindustria ==.
gen industria=fnes3/1000
replace industria=0 if industria==.
gen infraestrutura=fnes4/1000
replace infraestrutura=0 if infraestrutura==.
gen pecuaria=fnes5/1000
replace pecuaria=0 if pecuaria==.
gen turismo=fnes6/1000
replace turismo=0 if turismo==.
```

*** Obtando as médias do volume de recursos de cada setor nos estados

```
by estado ano, sort: sum agricola
```

```
by estado ano, sort: sum agroindustria
```

```
by estado ano, sort: sum industria
```

```
by estado ano, sort: sum infraestrutura
```

```
by estado ano, sort: sum pecuaria
```

by estado ano, sort: sum turismo

*** Obtendo as médias do volume de recursos de cada setor para tratados e não tratados

by w_prop ano, sort: sum agricola

by w_prop ano, sort: sum agroindustria

by w_prop ano, sort: sum industria

by w_prop ano, sort: sum infraestrutura

by w_prop ano, sort: sum pecuaria

by w_prop ano, sort: sum turismo

*** Obtendo as médias do volume de recursos de cada setor considerando o semiário

by w_semi ano, sort: sum agricola

by w_semi ano, sort: sum agroindustria

by w_semi ano, sort: sum industria

by w_semi ano, sort: sum infraestrutura

by w_semi ano, sort: sum pecuaria

by w_semi ano, sort: sum turismo

*** Obtendo a proporção do FNE em relação ao PIB

gen pfne=(fne/pib_def)*100

by estado ano, sort: sum pfne

by w_prop ano, sort: sum pfne

by w_semi ano, sort: sum pfne

*** Obtendo as participações de cada setor nos recursos do FNE

gen p_agricola = (agricola/mfne)*100

gen p_agroindustria=(agroindustria/mfne)*100

gen p_industria=(industria/mfne)*100

gen p_infraestrutura=(infraestrutura/mfne)*100

gen p_pecuaria=(pecuaria/mfne)*100
gen p_turismo=(turismo/mfne)*100

*** Obtendo a média das participações de cada setor nos recursos do FNE por estado

by estado ano, sort: sum p_agricola

by estado ano, sort: sum p_agroindustria

by estado ano, sort: sum p_industria

by estado ano, sort: sum p_infraestrutura

by estado ano, sort: sum p_pecuaria

by estado ano, sort: sum p_turismo

*** Obtendo a média das participações de cada setor nos recursos do FNE

by ano, sort: sum p_agricola

by ano, sort: sum p_agroindustria

by ano, sort: sum p_industria

by ano, sort: sum p_infraestrutura

by ano, sort: sum p_pecuaria

by ano, sort: sum p_turismo

*** Obtendo a média das participações de cada setor nos recursos do FNE por designação de tratamento

by w_prop ano, sort: sum p_agricola

by w_prop ano, sort: sum p_agroindustria

by w_prop ano, sort: sum p_industria

by w_prop ano, sort: sum p_infraestrutura

by w_prop ano, sort: sum p_pecuaria

by w_prop ano, sort: sum p_turismo

*** Obtendo a média das participações de cada setor nos recursos do FNE por localização no semiárido

by w_semi ano, sort: sum p_agricola

by w_semi ano, sort: sum p_agroindustria

by w_semi ano, sort: sum p_industria

by w_semi ano, sort: sum p_infraestrutura

by w_semi ano, sort: sum p_pecuaria

by w_semi ano, sort: sum p_turismo

** Estatísticas descritivas de variáveis de resultados

*** PIB per capita

by estado ano, sort: sum pib_pc

by ano, sort: sum pib_pc

by w_prop ano, sort: sum pib_pc

by w_semi ano, sort: sum pib_pc

*** Saldo de empregos

by estado ano, sort: sum saldo_pc

by ano, sort: sum saldo_pc

by w_prop ano, sort: sum saldo_pc

by w_semi ano, sort: sum saldo_pc

*** Valor adicionado - Agricultura

by estado ano, sort: sum va_agro_pc

by ano, sort: sum va_agro_pc

by w_prop ano, sort: sum va_agro_pc

by w_semi ano, sort: sum va_agro_pc

*** Valor adicionado - Indústria

by estado ano, sort: sum va_ind_pc

by ano, sort: sum va_ind_pc

by w_prop ano, sort: sum va_ind_pc

by w_semi ano, sort: sum va_ind_pc

*** Valor adicionado - Serviços

by estado ano, sort: sum va_ser_pc

by ano, sort: sum va_ser_pc

by w_prop ano, sort: sum va_ser_pc

by w_semi ano, sort: sum va_ser_pc

*** Valor adicionado - Administração Pública

by estado ano, sort: sum va_ade_pc

by ano, sort: sum va_ade_pc

by w_prop ano, sort: sum va_ade_pc

by w_semi ano, sort: sum va_ade_pc

*** Óbitos infantis por mil habitantes

gen oi_percapita=obitos_inf/(populacao/1000)

by estado ano, sort: sum oi_percapita

by ano, sort: sum oi_percapita

by w_prop ano, sort: sum oi_percapita

by w_semi ano, sort: sum oi_percapita

*** Óbitos maternos por mil habitantes

gen om_percapita=obitos_mat/(populacao/1000)

by estado ano, sort: sum om_percapita

by ano, sort: sum om_percapita

by w_prop ano, sort: sum om_percapita

by w_semi ano, sort: sum om_percapita

*** Óbitos por causas externas por mil habitantes

gen oe_percapita=obitos_ext/(populacao/1000)

by estado ano, sort: sum oe_percapita

by ano, sort: sum oe_percapita

by w_prop ano, sort: sum oe_percapita

by w_semi ano, sort: sum oe_percapita

*** Taxa de distorção idade-série no Ensino Médio

by estado ano, sort: sum tdi_em

by ano, sort: sum tdi_em

by w_prop ano, sort: sum tdi_em

by w_semi ano, sort: sum tdi_em

* Analisando a distribuição de renda

findit inequality /// instale o sg30

** Estimando o índice de Gini para a renda no Maranhão

inequal pib_pc if id_uf == 21 & ano==2006

inequal pib_pc if id_uf == 21 & ano==2007

inequal pib_pc if id_uf == 21 & ano==2008

inequal pib_pc if id_uf == 21 & ano==2009

inequal pib_pc if id_uf == 21 & ano==2010

inequal pib_pc if id_uf == 21 & ano==2011

inequal pib_pc if id_uf == 21 & ano==2012

inequal pib_pc if id_uf == 21 & ano==2013

inequal pib_pc if id_uf == 21 & ano==2014

inequal pib_pc if id_uf == 21 & ano==2015

inequal pib_pc if id_uf == 21 & ano==2016

inequal pib_pc if id_uf == 21 & ano==2017

inequal pib_pc if id_uf == 21 & ano==2018

** Estimando o índice de Gini para a renda no Piauí

inequal pib_pc if id_uf == 22 & ano==2006
inequal pib_pc if id_uf == 22 & ano==2007
inequal pib_pc if id_uf == 22 & ano==2008
inequal pib_pc if id_uf == 22 & ano==2009
inequal pib_pc if id_uf == 22 & ano==2010
inequal pib_pc if id_uf == 22 & ano==2011
inequal pib_pc if id_uf == 22 & ano==2012
inequal pib_pc if id_uf == 22 & ano==2013
inequal pib_pc if id_uf == 22 & ano==2014
inequal pib_pc if id_uf == 22 & ano==2015
inequal pib_pc if id_uf == 22 & ano==2016
inequal pib_pc if id_uf == 22 & ano==2017
inequal pib_pc if id_uf == 22 & ano==2018

** Estimando o índice de Gini para a renda no Ceará

inequal pib_pc if id_uf == 23 & ano==2006
inequal pib_pc if id_uf == 23 & ano==2007
inequal pib_pc if id_uf == 23 & ano==2008
inequal pib_pc if id_uf == 23 & ano==2009
inequal pib_pc if id_uf == 23 & ano==2010
inequal pib_pc if id_uf == 23 & ano==2011
inequal pib_pc if id_uf == 23 & ano==2012
inequal pib_pc if id_uf == 23 & ano==2013
inequal pib_pc if id_uf == 23 & ano==2014
inequal pib_pc if id_uf == 23 & ano==2015
inequal pib_pc if id_uf == 23 & ano==2016
inequal pib_pc if id_uf == 23 & ano==2017
inequal pib_pc if id_uf == 23 & ano==2018

** Estimando o índice de Gini para a renda no Rio Grande do Norte

inequal pib_pc if id_uf == 24 & ano==2006
inequal pib_pc if id_uf == 24 & ano==2007
inequal pib_pc if id_uf == 24 & ano==2008
inequal pib_pc if id_uf == 24 & ano==2009
inequal pib_pc if id_uf == 24 & ano==2010
inequal pib_pc if id_uf == 24 & ano==2011
inequal pib_pc if id_uf == 24 & ano==2012
inequal pib_pc if id_uf == 24 & ano==2013
inequal pib_pc if id_uf == 24 & ano==2014
inequal pib_pc if id_uf == 24 & ano==2015
inequal pib_pc if id_uf == 24 & ano==2016

inequal pib_pc if id_uf == 24 & ano==2017
inequal pib_pc if id_uf == 24 & ano==2018

** Estimando o índice de Gini para a renda na Paraíba

inequal pib_pc if id_uf == 25 & ano==2006
inequal pib_pc if id_uf == 25 & ano==2007
inequal pib_pc if id_uf == 25 & ano==2008
inequal pib_pc if id_uf == 25 & ano==2009
inequal pib_pc if id_uf == 25 & ano==2010
inequal pib_pc if id_uf == 25 & ano==2011
inequal pib_pc if id_uf == 25 & ano==2012
inequal pib_pc if id_uf == 25 & ano==2013
inequal pib_pc if id_uf == 25 & ano==2014
inequal pib_pc if id_uf == 25 & ano==2015
inequal pib_pc if id_uf == 25 & ano==2016
inequal pib_pc if id_uf == 25 & ano==2017
inequal pib_pc if id_uf == 25 & ano==2018

** Estimando o índice de Gini para a renda em Pernambuco

inequal pib_pc if id_uf == 26 & ano==2006
inequal pib_pc if id_uf == 26 & ano==2007
inequal pib_pc if id_uf == 26 & ano==2008
inequal pib_pc if id_uf == 26 & ano==2009
inequal pib_pc if id_uf == 26 & ano==2010
inequal pib_pc if id_uf == 26 & ano==2011
inequal pib_pc if id_uf == 26 & ano==2012
inequal pib_pc if id_uf == 26 & ano==2013
inequal pib_pc if id_uf == 26 & ano==2014
inequal pib_pc if id_uf == 26 & ano==2015
inequal pib_pc if id_uf == 26 & ano==2016
inequal pib_pc if id_uf == 26 & ano==2017
inequal pib_pc if id_uf == 26 & ano==2018

** Estimando o índice de Gini para a renda em Alagoas

inequal pib_pc if id_uf == 27 & ano==2006
inequal pib_pc if id_uf == 27 & ano==2007
inequal pib_pc if id_uf == 27 & ano==2008
inequal pib_pc if id_uf == 27 & ano==2009
inequal pib_pc if id_uf == 27 & ano==2010
inequal pib_pc if id_uf == 27 & ano==2011
inequal pib_pc if id_uf == 27 & ano==2012
inequal pib_pc if id_uf == 27 & ano==2013
inequal pib_pc if id_uf == 27 & ano==2014
inequal pib_pc if id_uf == 27 & ano==2015

inequal pib_pc if id_uf == 27 & ano==2016
inequal pib_pc if id_uf == 27 & ano==2017
inequal pib_pc if id_uf == 27 & ano==2018

** Estimando o índice de Gini para a renda no Sergipe

inequal pib_pc if id_uf == 28 & ano==2006
inequal pib_pc if id_uf == 28 & ano==2007
inequal pib_pc if id_uf == 28 & ano==2008
inequal pib_pc if id_uf == 28 & ano==2009
inequal pib_pc if id_uf == 28 & ano==2010
inequal pib_pc if id_uf == 28 & ano==2011
inequal pib_pc if id_uf == 28 & ano==2012
inequal pib_pc if id_uf == 28 & ano==2013
inequal pib_pc if id_uf == 28 & ano==2014
inequal pib_pc if id_uf == 28 & ano==2015
inequal pib_pc if id_uf == 28 & ano==2016
inequal pib_pc if id_uf == 28 & ano==2017
inequal pib_pc if id_uf == 28 & ano==2018

** Estimando o índice de Gini para a renda na Bahia

inequal pib_pc if id_uf == 29 & ano==2006
inequal pib_pc if id_uf == 29 & ano==2007
inequal pib_pc if id_uf == 29 & ano==2008
inequal pib_pc if id_uf == 29 & ano==2009
inequal pib_pc if id_uf == 29 & ano==2010
inequal pib_pc if id_uf == 29 & ano==2011
inequal pib_pc if id_uf == 29 & ano==2012
inequal pib_pc if id_uf == 29 & ano==2013
inequal pib_pc if id_uf == 29 & ano==2014
inequal pib_pc if id_uf == 29 & ano==2015
inequal pib_pc if id_uf == 29 & ano==2016
inequal pib_pc if id_uf == 29 & ano==2017
inequal pib_pc if id_uf == 29 & ano==2018

** Estimando o índice de Gini para a renda em Minas Gerais

inequal pib_pc if id_uf == 31 & ano==2006
inequal pib_pc if id_uf == 31 & ano==2007
inequal pib_pc if id_uf == 31 & ano==2008
inequal pib_pc if id_uf == 31 & ano==2009
inequal pib_pc if id_uf == 31 & ano==2010
inequal pib_pc if id_uf == 31 & ano==2011
inequal pib_pc if id_uf == 31 & ano==2012
inequal pib_pc if id_uf == 31 & ano==2013
inequal pib_pc if id_uf == 31 & ano==2014

```
inequal pib_pc if id_uf == 31 & ano==2015  
inequal pib_pc if id_uf == 31 & ano==2016  
inequal pib_pc if id_uf == 31 & ano==2017  
inequal pib_pc if id_uf == 31 & ano==2018
```

** Estimando o índice de Gini para a renda no Espírito Santo

```
inequal pib_pc if id_uf == 32 & ano==2006  
inequal pib_pc if id_uf == 32 & ano==2007  
inequal pib_pc if id_uf == 32 & ano==2008  
inequal pib_pc if id_uf == 32 & ano==2009  
inequal pib_pc if id_uf == 32 & ano==2010  
inequal pib_pc if id_uf == 32 & ano==2011  
inequal pib_pc if id_uf == 32 & ano==2012  
inequal pib_pc if id_uf == 32 & ano==2013  
inequal pib_pc if id_uf == 32 & ano==2014  
inequal pib_pc if id_uf == 32 & ano==2015  
inequal pib_pc if id_uf == 32 & ano==2016  
inequal pib_pc if id_uf == 32 & ano==2017  
inequal pib_pc if id_uf == 32 & ano==2018
```

** Estimando o índice de Gini para a renda em toda a amostra

```
inequal pib_pc if ano==2006  
inequal pib_pc if ano==2007  
inequal pib_pc if ano==2008  
inequal pib_pc if ano==2009  
inequal pib_pc if ano==2010  
inequal pib_pc if ano==2011  
inequal pib_pc if ano==2012  
inequal pib_pc if ano==2013  
inequal pib_pc if ano==2014  
inequal pib_pc if ano==2015  
inequal pib_pc if ano==2016  
inequal pib_pc if ano==2017  
inequal pib_pc if ano==2018
```



IPPDS
Instituto de Políticas Públicas e
Desenvolvimento Sustentável

* Analisando a distribuição de recursos do FNE

```
gen fne_percapita = fne/populacao
```

** Estimando o índice de Gini para os recursos do FNE no Maranhão

```
inequal fne_percapita if id_uf == 21 & ano==2006  
inequal fne_percapita if id_uf == 21 & ano==2007  
inequal fne_percapita if id_uf == 21 & ano==2008  
inequal fne_percapita if id_uf == 21 & ano==2009
```

inequal fne_percapita if id_uf == 21 & ano==2010
inequal fne_percapita if id_uf == 21 & ano==2011
inequal fne_percapita if id_uf == 21 & ano==2012
inequal fne_percapita if id_uf == 21 & ano==2013
inequal fne_percapita if id_uf == 21 & ano==2014
inequal fne_percapita if id_uf == 21 & ano==2015
inequal fne_percapita if id_uf == 21 & ano==2016
inequal fne_percapita if id_uf == 21 & ano==2017
inequal fne_percapita if id_uf == 21 & ano==2018

** Estimando o índice de Gini para os recursos do FNE no Piauí

inequal fne_percapita if id_uf == 22 & ano==2006
inequal fne_percapita if id_uf == 22 & ano==2007
inequal fne_percapita if id_uf == 22 & ano==2008
inequal fne_percapita if id_uf == 22 & ano==2009
inequal fne_percapita if id_uf == 22 & ano==2010
inequal fne_percapita if id_uf == 22 & ano==2011
inequal fne_percapita if id_uf == 22 & ano==2012
inequal fne_percapita if id_uf == 22 & ano==2013
inequal fne_percapita if id_uf == 22 & ano==2014
inequal fne_percapita if id_uf == 22 & ano==2015
inequal fne_percapita if id_uf == 22 & ano==2016
inequal fne_percapita if id_uf == 22 & ano==2017
inequal fne_percapita if id_uf == 22 & ano==2018

** Estimando o índice de Gini para os recursos do FNE no Ceará

inequal fne_percapita if id_uf == 23 & ano==2006
inequal fne_percapita if id_uf == 23 & ano==2007
inequal fne_percapita if id_uf == 23 & ano==2008
inequal fne_percapita if id_uf == 23 & ano==2009
inequal fne_percapita if id_uf == 23 & ano==2010
inequal fne_percapita if id_uf == 23 & ano==2011
inequal fne_percapita if id_uf == 23 & ano==2012
inequal fne_percapita if id_uf == 23 & ano==2013
inequal fne_percapita if id_uf == 23 & ano==2014
inequal fne_percapita if id_uf == 23 & ano==2015
inequal fne_percapita if id_uf == 23 & ano==2016
inequal fne_percapita if id_uf == 23 & ano==2017
inequal fne_percapita if id_uf == 23 & ano==2018

** Estimando o índice de Gini para os recursos do FNE no Rio Grande do Norte

inequal fne_percapita if id_uf == 24 & ano==2006
inequal fne_percapita if id_uf == 24 & ano==2007
inequal fne_percapita if id_uf == 24 & ano==2008

inequal fne_percapita if id_uf == 24 & ano==2009
inequal fne_percapita if id_uf == 24 & ano==2010
inequal fne_percapita if id_uf == 24 & ano==2011
inequal fne_percapita if id_uf == 24 & ano==2012
inequal fne_percapita if id_uf == 24 & ano==2013
inequal fne_percapita if id_uf == 24 & ano==2014
inequal fne_percapita if id_uf == 24 & ano==2015
inequal fne_percapita if id_uf == 24 & ano==2016
inequal fne_percapita if id_uf == 24 & ano==2017
inequal fne_percapita if id_uf == 24 & ano==2018

** Estimando o índice de Gini para os recursos do FNE na Paraíba

inequal fne_percapita if id_uf == 25 & ano==2006
inequal fne_percapita if id_uf == 25 & ano==2007
inequal fne_percapita if id_uf == 25 & ano==2008
inequal fne_percapita if id_uf == 25 & ano==2009
inequal fne_percapita if id_uf == 25 & ano==2010
inequal fne_percapita if id_uf == 25 & ano==2011
inequal fne_percapita if id_uf == 25 & ano==2012
inequal fne_percapita if id_uf == 25 & ano==2013
inequal fne_percapita if id_uf == 25 & ano==2014
inequal fne_percapita if id_uf == 25 & ano==2015
inequal fne_percapita if id_uf == 25 & ano==2016
inequal fne_percapita if id_uf == 25 & ano==2017
inequal fne_percapita if id_uf == 25 & ano==2018

** Estimando o índice de Gini para os recursos do FNE em Pernambuco

inequal fne_percapita if id_uf == 26 & ano==2006
inequal fne_percapita if id_uf == 26 & ano==2007
inequal fne_percapita if id_uf == 26 & ano==2008
inequal fne_percapita if id_uf == 26 & ano==2009
inequal fne_percapita if id_uf == 26 & ano==2010
inequal fne_percapita if id_uf == 26 & ano==2011
inequal fne_percapita if id_uf == 26 & ano==2012
inequal fne_percapita if id_uf == 26 & ano==2013
inequal fne_percapita if id_uf == 26 & ano==2014
inequal fne_percapita if id_uf == 26 & ano==2015
inequal fne_percapita if id_uf == 26 & ano==2016
inequal fne_percapita if id_uf == 26 & ano==2017
inequal fne_percapita if id_uf == 26 & ano==2018

** Estimando o índice de Gini para os recursos do FNE em Alagoas

inequal fne_percapita if id_uf == 27 & ano==2006
inequal fne_percapita if id_uf == 27 & ano==2007

inequal fne_percapita if id_uf == 27 & ano==2008
inequal fne_percapita if id_uf == 27 & ano==2009
inequal fne_percapita if id_uf == 27 & ano==2010
inequal fne_percapita if id_uf == 27 & ano==2011
inequal fne_percapita if id_uf == 27 & ano==2012
inequal fne_percapita if id_uf == 27 & ano==2013
inequal fne_percapita if id_uf == 27 & ano==2014
inequal fne_percapita if id_uf == 27 & ano==2015
inequal fne_percapita if id_uf == 27 & ano==2016
inequal fne_percapita if id_uf == 27 & ano==2017
inequal fne_percapita if id_uf == 27 & ano==2018

** Estimando o índice de Gini para os recursos do FNE no Sergipe

inequal fne_percapita if id_uf == 28 & ano==2006
inequal fne_percapita if id_uf == 28 & ano==2007
inequal fne_percapita if id_uf == 28 & ano==2008
inequal fne_percapita if id_uf == 28 & ano==2009
inequal fne_percapita if id_uf == 28 & ano==2010
inequal fne_percapita if id_uf == 28 & ano==2011
inequal fne_percapita if id_uf == 28 & ano==2012
inequal fne_percapita if id_uf == 28 & ano==2013
inequal fne_percapita if id_uf == 28 & ano==2014
inequal fne_percapita if id_uf == 28 & ano==2015
inequal fne_percapita if id_uf == 28 & ano==2016
inequal fne_percapita if id_uf == 28 & ano==2017
inequal fne_percapita if id_uf == 28 & ano==2018

** Estimando o índice de Gini para os recursos do FNE na Bahia

inequal fne_percapita if id_uf == 29 & ano==2006
inequal fne_percapita if id_uf == 29 & ano==2007
inequal fne_percapita if id_uf == 29 & ano==2008
inequal fne_percapita if id_uf == 29 & ano==2009
inequal fne_percapita if id_uf == 29 & ano==2010
inequal fne_percapita if id_uf == 29 & ano==2011
inequal fne_percapita if id_uf == 29 & ano==2012
inequal fne_percapita if id_uf == 29 & ano==2013
inequal fne_percapita if id_uf == 29 & ano==2014
inequal fne_percapita if id_uf == 29 & ano==2015
inequal fne_percapita if id_uf == 29 & ano==2016
inequal fne_percapita if id_uf == 29 & ano==2017
inequal fne_percapita if id_uf == 29 & ano==2018

** Estimando o índice de Gini para os recursos do FNE em Minas Gerais

inequal fne_percapita if id_uf == 31 & ano==2006

```
inequal fne_percapita if id_uf == 31 & ano==2007  
inequal fne_percapita if id_uf == 31 & ano==2008  
inequal fne_percapita if id_uf == 31 & ano==2009  
inequal fne_percapita if id_uf == 31 & ano==2010  
inequal fne_percapita if id_uf == 31 & ano==2011  
inequal fne_percapita if id_uf == 31 & ano==2012  
inequal fne_percapita if id_uf == 31 & ano==2013  
inequal fne_percapita if id_uf == 31 & ano==2014  
inequal fne_percapita if id_uf == 31 & ano==2015  
inequal fne_percapita if id_uf == 31 & ano==2016  
inequal fne_percapita if id_uf == 31 & ano==2017  
inequal fne_percapita if id_uf == 31 & ano==2018
```

** Estimando o índice de Gini para os recursos do FNE no Espírito Santo

```
inequal fne_percapita if id_uf == 32 & ano==2006  
inequal fne_percapita if id_uf == 32 & ano==2007  
inequal fne_percapita if id_uf == 32 & ano==2008  
inequal fne_percapita if id_uf == 32 & ano==2009  
inequal fne_percapita if id_uf == 32 & ano==2010  
inequal fne_percapita if id_uf == 32 & ano==2011  
inequal fne_percapita if id_uf == 32 & ano==2012  
inequal fne_percapita if id_uf == 32 & ano==2013  
inequal fne_percapita if id_uf == 32 & ano==2014  
inequal fne_percapita if id_uf == 32 & ano==2015  
inequal fne_percapita if id_uf == 32 & ano==2016  
inequal fne_percapita if id_uf == 32 & ano==2017  
inequal fne_percapita if id_uf == 32 & ano==2018
```

** Estimando o índice de Gini para os recursos do FNE em toda a amostra

```
inequal fne_percapita if ano==2006  
inequal fne_percapita if ano==2007  
inequal fne_percapita if ano==2008  
inequal fne_percapita if ano==2009  
inequal fne_percapita if ano==2010  
inequal fne_percapita if ano==2011  
inequal fne_percapita if ano==2012  
inequal fne_percapita if ano==2013  
inequal fne_percapita if ano==2014  
inequal fne_percapita if ano==2015  
inequal fne_percapita if ano==2016  
inequal fne_percapita if ano==2017  
inequal fne_percapita if ano==2018
```

* Analisando a distribuição de óbitos infantis

** Estimando o índice de Gini para os óbitos infantis no Maranhão

inequal oi_percapita if id_uf == 21 & ano==2006
inequal oi_percapita if id_uf == 21 & ano==2007
inequal oi_percapita if id_uf == 21 & ano==2008
inequal oi_percapita if id_uf == 21 & ano==2009
inequal oi_percapita if id_uf == 21 & ano==2010
inequal oi_percapita if id_uf == 21 & ano==2011
inequal oi_percapita if id_uf == 21 & ano==2012
inequal oi_percapita if id_uf == 21 & ano==2013
inequal oi_percapita if id_uf == 21 & ano==2014
inequal oi_percapita if id_uf == 21 & ano==2015
inequal oi_percapita if id_uf == 21 & ano==2016
inequal oi_percapita if id_uf == 21 & ano==2017
inequal oi_percapita if id_uf == 21 & ano==2018

** Estimando o índice de Gini para os óbitos infantis no Piauí

inequal oi_percapita if id_uf == 22 & ano==2006
inequal oi_percapita if id_uf == 22 & ano==2007
inequal oi_percapita if id_uf == 22 & ano==2008
inequal oi_percapita if id_uf == 22 & ano==2009
inequal oi_percapita if id_uf == 22 & ano==2010
inequal oi_percapita if id_uf == 22 & ano==2011
inequal oi_percapita if id_uf == 22 & ano==2012
inequal oi_percapita if id_uf == 22 & ano==2013
inequal oi_percapita if id_uf == 22 & ano==2014
inequal oi_percapita if id_uf == 22 & ano==2015
inequal oi_percapita if id_uf == 22 & ano==2016
inequal oi_percapita if id_uf == 22 & ano==2017
inequal oi_percapita if id_uf == 22 & ano==2018

** Estimando o índice de Gini para os óbitos infantis no Ceará

inequal oi_percapita if id_uf == 23 & ano==2006
inequal oi_percapita if id_uf == 23 & ano==2007
inequal oi_percapita if id_uf == 23 & ano==2008
inequal oi_percapita if id_uf == 23 & ano==2009
inequal oi_percapita if id_uf == 23 & ano==2010
inequal oi_percapita if id_uf == 23 & ano==2011
inequal oi_percapita if id_uf == 23 & ano==2012
inequal oi_percapita if id_uf == 23 & ano==2013
inequal oi_percapita if id_uf == 23 & ano==2014
inequal oi_percapita if id_uf == 23 & ano==2015
inequal oi_percapita if id_uf == 23 & ano==2016
inequal oi_percapita if id_uf == 23 & ano==2017
inequal oi_percapita if id_uf == 23 & ano==2018

** Estimando o índice de Gini para os óbitos infantis no Rio Grande do Norte

inequal oi_percapita if id_uf == 24 & ano==2006
inequal oi_percapita if id_uf == 24 & ano==2007
inequal oi_percapita if id_uf == 24 & ano==2008
inequal oi_percapita if id_uf == 24 & ano==2009
inequal oi_percapita if id_uf == 24 & ano==2010
inequal oi_percapita if id_uf == 24 & ano==2011
inequal oi_percapita if id_uf == 24 & ano==2012
inequal oi_percapita if id_uf == 24 & ano==2013
inequal oi_percapita if id_uf == 24 & ano==2014
inequal oi_percapita if id_uf == 24 & ano==2015
inequal oi_percapita if id_uf == 24 & ano==2016
inequal oi_percapita if id_uf == 24 & ano==2017
inequal oi_percapita if id_uf == 24 & ano==2018

** Estimando o índice de Gini para os óbitos infantis na Paraíba

inequal oi_percapita if id_uf == 25 & ano==2006
inequal oi_percapita if id_uf == 25 & ano==2007
inequal oi_percapita if id_uf == 25 & ano==2008
inequal oi_percapita if id_uf == 25 & ano==2009
inequal oi_percapita if id_uf == 25 & ano==2010
inequal oi_percapita if id_uf == 25 & ano==2011
inequal oi_percapita if id_uf == 25 & ano==2012
inequal oi_percapita if id_uf == 25 & ano==2013
inequal oi_percapita if id_uf == 25 & ano==2014
inequal oi_percapita if id_uf == 25 & ano==2015
inequal oi_percapita if id_uf == 25 & ano==2016
inequal oi_percapita if id_uf == 25 & ano==2017
inequal oi_percapita if id_uf == 25 & ano==2018

** Estimando o índice de Gini para os óbitos infantis em Pernambuco

inequal oi_percapita if id_uf == 26 & ano==2006
inequal oi_percapita if id_uf == 26 & ano==2007
inequal oi_percapita if id_uf == 26 & ano==2008
inequal oi_percapita if id_uf == 26 & ano==2009
inequal oi_percapita if id_uf == 26 & ano==2010
inequal oi_percapita if id_uf == 26 & ano==2011
inequal oi_percapita if id_uf == 26 & ano==2012
inequal oi_percapita if id_uf == 26 & ano==2013
inequal oi_percapita if id_uf == 26 & ano==2014
inequal oi_percapita if id_uf == 26 & ano==2015
inequal oi_percapita if id_uf == 26 & ano==2016
inequal oi_percapita if id_uf == 26 & ano==2017

inequal oi_percapita if id_uf == 26 & ano==2018

** Estimando o índice de Gini para os óbitos infantis em Alagoas

inequal oi_percapita if id_uf == 27 & ano==2006
inequal oi_percapita if id_uf == 27 & ano==2007
inequal oi_percapita if id_uf == 27 & ano==2008
inequal oi_percapita if id_uf == 27 & ano==2009
inequal oi_percapita if id_uf == 27 & ano==2010
inequal oi_percapita if id_uf == 27 & ano==2011
inequal oi_percapita if id_uf == 27 & ano==2012
inequal oi_percapita if id_uf == 27 & ano==2013
inequal oi_percapita if id_uf == 27 & ano==2014
inequal oi_percapita if id_uf == 27 & ano==2015
inequal oi_percapita if id_uf == 27 & ano==2016
inequal oi_percapita if id_uf == 27 & ano==2017
inequal oi_percapita if id_uf == 27 & ano==2018

** Estimando o índice de Gini para os óbitos infantis no Sergipe

inequal oi_percapita if id_uf == 28 & ano==2006
inequal oi_percapita if id_uf == 28 & ano==2007
inequal oi_percapita if id_uf == 28 & ano==2008
inequal oi_percapita if id_uf == 28 & ano==2009
inequal oi_percapita if id_uf == 28 & ano==2010
inequal oi_percapita if id_uf == 28 & ano==2011
inequal oi_percapita if id_uf == 28 & ano==2012
inequal oi_percapita if id_uf == 28 & ano==2013
inequal oi_percapita if id_uf == 28 & ano==2014
inequal oi_percapita if id_uf == 28 & ano==2015
inequal oi_percapita if id_uf == 28 & ano==2016
inequal oi_percapita if id_uf == 28 & ano==2017
inequal oi_percapita if id_uf == 28 & ano==2018

** Estimando o índice de Gini para os óbitos infantis na Bahia

inequal oi_percapita if id_uf == 29 & ano==2006
inequal oi_percapita if id_uf == 29 & ano==2007
inequal oi_percapita if id_uf == 29 & ano==2008
inequal oi_percapita if id_uf == 29 & ano==2009
inequal oi_percapita if id_uf == 29 & ano==2010
inequal oi_percapita if id_uf == 29 & ano==2011
inequal oi_percapita if id_uf == 29 & ano==2012
inequal oi_percapita if id_uf == 29 & ano==2013
inequal oi_percapita if id_uf == 29 & ano==2014
inequal oi_percapita if id_uf == 29 & ano==2015
inequal oi_percapita if id_uf == 29 & ano==2016

inequal oi_percapita if id_uf == 29 & ano==2017
inequal oi_percapita if id_uf == 29 & ano==2018

** Estimando o índice de Gini para os óbitos infantis em Minas Gerais

inequal oi_percapita if id_uf == 31 & ano==2006
inequal oi_percapita if id_uf == 31 & ano==2007
inequal oi_percapita if id_uf == 31 & ano==2008
inequal oi_percapita if id_uf == 31 & ano==2009
inequal oi_percapita if id_uf == 31 & ano==2010
inequal oi_percapita if id_uf == 31 & ano==2011
inequal oi_percapita if id_uf == 31 & ano==2012
inequal oi_percapita if id_uf == 31 & ano==2013
inequal oi_percapita if id_uf == 31 & ano==2014
inequal oi_percapita if id_uf == 31 & ano==2015
inequal oi_percapita if id_uf == 31 & ano==2016
inequal oi_percapita if id_uf == 31 & ano==2017
inequal oi_percapita if id_uf == 31 & ano==2018

** Estimando o índice de Gini para os óbitos infantis no Espírito Santo

inequal oi_percapita if id_uf == 32 & ano==2006
inequal oi_percapita if id_uf == 32 & ano==2007
inequal oi_percapita if id_uf == 32 & ano==2008
inequal oi_percapita if id_uf == 32 & ano==2009
inequal oi_percapita if id_uf == 32 & ano==2010
inequal oi_percapita if id_uf == 32 & ano==2011
inequal oi_percapita if id_uf == 32 & ano==2012
inequal oi_percapita if id_uf == 32 & ano==2013
inequal oi_percapita if id_uf == 32 & ano==2014
inequal oi_percapita if id_uf == 32 & ano==2015
inequal oi_percapita if id_uf == 32 & ano==2016
inequal oi_percapita if id_uf == 32 & ano==2017
inequal oi_percapita if id_uf == 32 & ano==2018

** Estimando o índice de Gini para os óbitos infantis em toda a amostra

inequal oi_percapita if ano==2006
inequal oi_percapita if ano==2007
inequal oi_percapita if ano==2008
inequal oi_percapita if ano==2009
inequal oi_percapita if ano==2010
inequal oi_percapita if ano==2011
inequal oi_percapita if ano==2012
inequal oi_percapita if ano==2013
inequal oi_percapita if ano==2014
inequal oi_percapita if ano==2015

inequal oi_percapita if ano==2016
inequal oi_percapita if ano==2017
inequal oi_percapita if ano==2018

* Analisando a distribuição da taxa de distorção idade-série no Ensino Médio

** Estimando o índice de Gini para a taxa de distorção no Maranhão

inequal tdi_em if id_uf == 21 & ano==2006
inequal tdi_em if id_uf == 21 & ano==2007
inequal tdi_em if id_uf == 21 & ano==2008
inequal tdi_em if id_uf == 21 & ano==2009
inequal tdi_em if id_uf == 21 & ano==2010
inequal tdi_em if id_uf == 21 & ano==2011
inequal tdi_em if id_uf == 21 & ano==2012
inequal tdi_em if id_uf == 21 & ano==2013
inequal tdi_em if id_uf == 21 & ano==2014
inequal tdi_em if id_uf == 21 & ano==2015
inequal tdi_em if id_uf == 21 & ano==2016
inequal tdi_em if id_uf == 21 & ano==2017
inequal tdi_em if id_uf == 21 & ano==2018

** Estimando o índice de Gini para a taxa de distorção no Piauí

inequal tdi_em if id_uf == 22 & ano==2006
inequal tdi_em if id_uf == 22 & ano==2007
inequal tdi_em if id_uf == 22 & ano==2008
inequal tdi_em if id_uf == 22 & ano==2009
inequal tdi_em if id_uf == 22 & ano==2010
inequal tdi_em if id_uf == 22 & ano==2011
inequal tdi_em if id_uf == 22 & ano==2012
inequal tdi_em if id_uf == 22 & ano==2013
inequal tdi_em if id_uf == 22 & ano==2014
inequal tdi_em if id_uf == 22 & ano==2015
inequal tdi_em if id_uf == 22 & ano==2016
inequal tdi_em if id_uf == 22 & ano==2017
inequal tdi_em if id_uf == 22 & ano==2018

** Estimando o índice de Gini para a taxa de distorção no Ceará

inequal tdi_em if id_uf == 23 & ano==2006
inequal tdi_em if id_uf == 23 & ano==2007
inequal tdi_em if id_uf == 23 & ano==2008
inequal tdi_em if id_uf == 23 & ano==2009
inequal tdi_em if id_uf == 23 & ano==2010
inequal tdi_em if id_uf == 23 & ano==2011
inequal tdi_em if id_uf == 23 & ano==2012

inequal tdi_em if id_uf == 23 & ano==2013
inequal tdi_em if id_uf == 23 & ano==2014
inequal tdi_em if id_uf == 23 & ano==2015
inequal tdi_em if id_uf == 23 & ano==2016
inequal tdi_em if id_uf == 23 & ano==2017
inequal tdi_em if id_uf == 23 & ano==2018

** Estimando o índice de Gini para a taxa de distorção no Rio Grande do Norte

inequal tdi_em if id_uf == 24 & ano==2006
inequal tdi_em if id_uf == 24 & ano==2007
inequal tdi_em if id_uf == 24 & ano==2008
inequal tdi_em if id_uf == 24 & ano==2009
inequal tdi_em if id_uf == 24 & ano==2010
inequal tdi_em if id_uf == 24 & ano==2011
inequal tdi_em if id_uf == 24 & ano==2012
inequal tdi_em if id_uf == 24 & ano==2013
inequal tdi_em if id_uf == 24 & ano==2014
inequal tdi_em if id_uf == 24 & ano==2015
inequal tdi_em if id_uf == 24 & ano==2016
inequal tdi_em if id_uf == 24 & ano==2017
inequal tdi_em if id_uf == 24 & ano==2018

** Estimando o índice de Gini para a taxa de distorção na Paraíba

inequal tdi_em if id_uf == 25 & ano==2006
inequal tdi_em if id_uf == 25 & ano==2007
inequal tdi_em if id_uf == 25 & ano==2008
inequal tdi_em if id_uf == 25 & ano==2009
inequal tdi_em if id_uf == 25 & ano==2010
inequal tdi_em if id_uf == 25 & ano==2011
inequal tdi_em if id_uf == 25 & ano==2012
inequal tdi_em if id_uf == 25 & ano==2013
inequal tdi_em if id_uf == 25 & ano==2014
inequal tdi_em if id_uf == 25 & ano==2015
inequal tdi_em if id_uf == 25 & ano==2016
inequal tdi_em if id_uf == 25 & ano==2017
inequal tdi_em if id_uf == 25 & ano==2018

** Estimando o índice de Gini para a taxa de distorção em Pernambuco

inequal tdi_em if id_uf == 26 & ano==2006
inequal tdi_em if id_uf == 26 & ano==2007
inequal tdi_em if id_uf == 26 & ano==2008
inequal tdi_em if id_uf == 26 & ano==2009
inequal tdi_em if id_uf == 26 & ano==2010
inequal tdi_em if id_uf == 26 & ano==2011

inequal tdi_em if id_uf == 26 & ano==2012
inequal tdi_em if id_uf == 26 & ano==2013
inequal tdi_em if id_uf == 26 & ano==2014
inequal tdi_em if id_uf == 26 & ano==2015
inequal tdi_em if id_uf == 26 & ano==2016
inequal tdi_em if id_uf == 26 & ano==2017
inequal tdi_em if id_uf == 26 & ano==2018

** Estimando o índice de Gini para a taxa de distorção em Alagoas

inequal tdi_em if id_uf == 27 & ano==2006
inequal tdi_em if id_uf == 27 & ano==2007
inequal tdi_em if id_uf == 27 & ano==2008
inequal tdi_em if id_uf == 27 & ano==2009
inequal tdi_em if id_uf == 27 & ano==2010
inequal tdi_em if id_uf == 27 & ano==2011
inequal tdi_em if id_uf == 27 & ano==2012
inequal tdi_em if id_uf == 27 & ano==2013
inequal tdi_em if id_uf == 27 & ano==2014
inequal tdi_em if id_uf == 27 & ano==2015
inequal tdi_em if id_uf == 27 & ano==2016
inequal tdi_em if id_uf == 27 & ano==2017
inequal tdi_em if id_uf == 27 & ano==2018

** Estimando o índice de Gini para a taxa de distorção no Sergipe

inequal tdi_em if id_uf == 28 & ano==2006
inequal tdi_em if id_uf == 28 & ano==2007
inequal tdi_em if id_uf == 28 & ano==2008
inequal tdi_em if id_uf == 28 & ano==2009
inequal tdi_em if id_uf == 28 & ano==2010
inequal tdi_em if id_uf == 28 & ano==2011
inequal tdi_em if id_uf == 28 & ano==2012
inequal tdi_em if id_uf == 28 & ano==2013
inequal tdi_em if id_uf == 28 & ano==2014
inequal tdi_em if id_uf == 28 & ano==2015
inequal tdi_em if id_uf == 28 & ano==2016
inequal tdi_em if id_uf == 28 & ano==2017
inequal tdi_em if id_uf == 28 & ano==2018

** Estimando o índice de Gini para a taxa de distorção na Bahia

inequal tdi_em if id_uf == 29 & ano==2006
inequal tdi_em if id_uf == 29 & ano==2007
inequal tdi_em if id_uf == 29 & ano==2008
inequal tdi_em if id_uf == 29 & ano==2009
inequal tdi_em if id_uf == 29 & ano==2010

inequal tdi_em if id_uf == 29 & ano==2011
inequal tdi_em if id_uf == 29 & ano==2012
inequal tdi_em if id_uf == 29 & ano==2013
inequal tdi_em if id_uf == 29 & ano==2014
inequal tdi_em if id_uf == 29 & ano==2015
inequal tdi_em if id_uf == 29 & ano==2016
inequal tdi_em if id_uf == 29 & ano==2017
inequal tdi_em if id_uf == 29 & ano==2018

** Estimando o índice de Gini para a taxa de distorção em Minas Gerais

inequal tdi_em if id_uf == 31 & ano==2006
inequal tdi_em if id_uf == 31 & ano==2007
inequal tdi_em if id_uf == 31 & ano==2008
inequal tdi_em if id_uf == 31 & ano==2009
inequal tdi_em if id_uf == 31 & ano==2010
inequal tdi_em if id_uf == 31 & ano==2011
inequal tdi_em if id_uf == 31 & ano==2012
inequal tdi_em if id_uf == 31 & ano==2013
inequal tdi_em if id_uf == 31 & ano==2014
inequal tdi_em if id_uf == 31 & ano==2015
inequal tdi_em if id_uf == 31 & ano==2016
inequal tdi_em if id_uf == 31 & ano==2017
inequal tdi_em if id_uf == 31 & ano==2018

** Estimando o índice de Gini para a taxa de distorção no Espírito Santo

inequal tdi_em if id_uf == 32 & ano==2006
inequal tdi_em if id_uf == 32 & ano==2007
inequal tdi_em if id_uf == 32 & ano==2008
inequal tdi_em if id_uf == 32 & ano==2009
inequal tdi_em if id_uf == 32 & ano==2010
inequal tdi_em if id_uf == 32 & ano==2011
inequal tdi_em if id_uf == 32 & ano==2012
inequal tdi_em if id_uf == 32 & ano==2013
inequal tdi_em if id_uf == 32 & ano==2014
inequal tdi_em if id_uf == 32 & ano==2015
inequal tdi_em if id_uf == 32 & ano==2016
inequal tdi_em if id_uf == 32 & ano==2017
inequal tdi_em if id_uf == 32 & ano==2018

** Estimando o índice de Gini para a taxa de distorção em toda a amostra

inequal tdi_em if ano==2006
inequal tdi_em if ano==2007
inequal tdi_em if ano==2008
inequal tdi_em if ano==2009

```
inequal tdi_em if ano==2010  
inequal tdi_em if ano==2011  
inequal tdi_em if ano==2012  
inequal tdi_em if ano==2013  
inequal tdi_em if ano==2014  
inequal tdi_em if ano==2015  
inequal tdi_em if ano==2016  
inequal tdi_em if ano==2017  
inequal tdi_em if ano==2018
```

* Calculando o índice de Gini por designação de tratamento

** Estimando o índice de Gini para o PIB per capita

```
inequal pib_pc if w_prop == 0 & ano==2006  
inequal pib_pc if w_prop == 0 & ano==2007  
inequal pib_pc if w_prop == 0 & ano==2008  
inequal pib_pc if w_prop == 0 & ano==2009  
inequal pib_pc if w_prop == 0 & ano==2010  
inequal pib_pc if w_prop == 0 & ano==2011  
inequal pib_pc if w_prop == 0 & ano==2012  
inequal pib_pc if w_prop == 0 & ano==2013  
inequal pib_pc if w_prop == 0 & ano==2014  
inequal pib_pc if w_prop == 0 & ano==2015  
inequal pib_pc if w_prop == 0 & ano==2016  
inequal pib_pc if w_prop == 0 & ano==2017  
inequal pib_pc if w_prop == 0 & ano==2018
```

```
inequal pib_pc if w_prop == 1 & ano==2006  
inequal pib_pc if w_prop == 1 & ano==2007  
inequal pib_pc if w_prop == 1 & ano==2008  
inequal pib_pc if w_prop == 1 & ano==2009  
inequal pib_pc if w_prop == 1 & ano==2010  
inequal pib_pc if w_prop == 1 & ano==2011  
inequal pib_pc if w_prop == 1 & ano==2012  
inequal pib_pc if w_prop == 1 & ano==2013  
inequal pib_pc if w_prop == 1 & ano==2014  
inequal pib_pc if w_prop == 1 & ano==2015  
inequal pib_pc if w_prop == 1 & ano==2016  
inequal pib_pc if w_prop == 1 & ano==2017  
inequal pib_pc if w_prop == 1 & ano==2018
```

** Estimando o índice de Gini para a distribuição do FNE per capita

```
inequal fne_percapita if w_prop == 0 & ano==2006  
inequal fne_percapita if w_prop == 0 & ano==2007  
inequal fne_percapita if w_prop == 0 & ano==2008
```

inequal fne_percapita if w_prop == 0 & ano==2009
inequal fne_percapita if w_prop == 0 & ano==2010
inequal fne_percapita if w_prop == 0 & ano==2011
inequal fne_percapita if w_prop == 0 & ano==2012
inequal fne_percapita if w_prop == 0 & ano==2013
inequal fne_percapita if w_prop == 0 & ano==2014
inequal fne_percapita if w_prop == 0 & ano==2015
inequal fne_percapita if w_prop == 0 & ano==2016
inequal fne_percapita if w_prop == 0 & ano==2017
inequal fne_percapita if w_prop == 0 & ano==2018

inequal fne_percapita if w_prop == 1 & ano==2006
inequal fne_percapita if w_prop == 1 & ano==2007
inequal fne_percapita if w_prop == 1 & ano==2008
inequal fne_percapita if w_prop == 1 & ano==2009
inequal fne_percapita if w_prop == 1 & ano==2010
inequal fne_percapita if w_prop == 1 & ano==2011
inequal fne_percapita if w_prop == 1 & ano==2012
inequal fne_percapita if w_prop == 1 & ano==2013
inequal fne_percapita if w_prop == 1 & ano==2014
inequal fne_percapita if w_prop == 1 & ano==2015
inequal fne_percapita if w_prop == 1 & ano==2016
inequal fne_percapita if w_prop == 1 & ano==2017
inequal fne_percapita if w_prop == 1 & ano==2018

** Estimando o índice de Gini para os óbitos infantis por mil habitantes

inequal oi_percapita if w_prop == 0 & ano==2006
inequal oi_percapita if w_prop == 0 & ano==2007
inequal oi_percapita if w_prop == 0 & ano==2008
inequal oi_percapita if w_prop == 0 & ano==2009
inequal oi_percapita if w_prop == 0 & ano==2010
inequal oi_percapita if w_prop == 0 & ano==2011
inequal oi_percapita if w_prop == 0 & ano==2012
inequal oi_percapita if w_prop == 0 & ano==2013
inequal oi_percapita if w_prop == 0 & ano==2014
inequal oi_percapita if w_prop == 0 & ano==2015
inequal oi_percapita if w_prop == 0 & ano==2016
inequal oi_percapita if w_prop == 0 & ano==2017
inequal oi_percapita if w_prop == 0 & ano==2018

inequal oi_percapita if w_prop == 1 & ano==2006
inequal oi_percapita if w_prop == 1 & ano==2007
inequal oi_percapita if w_prop == 1 & ano==2008
inequal oi_percapita if w_prop == 1 & ano==2009
inequal oi_percapita if w_prop == 1 & ano==2010
inequal oi_percapita if w_prop == 1 & ano==2011

```
inequal oi_percapita if w_prop == 1 & ano==2012  
inequal oi_percapita if w_prop == 1 & ano==2013  
inequal oi_percapita if w_prop == 1 & ano==2014  
inequal oi_percapita if w_prop == 1 & ano==2015  
inequal oi_percapita if w_prop == 1 & ano==2016  
inequal oi_percapita if w_prop == 1 & ano==2017  
inequal oi_percapita if w_prop == 1 & ano==2018
```

** Estimando o índice de Gini para a taxa de distorção de idade-série no Ensino Médio

```
inequal tdi_em if w_prop == 0 & ano==2006  
inequal tdi_em if w_prop == 0 & ano==2007  
inequal tdi_em if w_prop == 0 & ano==2008  
inequal tdi_em if w_prop == 0 & ano==2009  
inequal tdi_em if w_prop == 0 & ano==2010  
inequal tdi_em if w_prop == 0 & ano==2011  
inequal tdi_em if w_prop == 0 & ano==2012  
inequal tdi_em if w_prop == 0 & ano==2013  
inequal tdi_em if w_prop == 0 & ano==2014  
inequal tdi_em if w_prop == 0 & ano==2015  
inequal tdi_em if w_prop == 0 & ano==2016  
inequal tdi_em if w_prop == 0 & ano==2017  
inequal tdi_em if w_prop == 0 & ano==2018
```

```
inequal tdi_em if w_prop == 1 & ano==2006  
inequal tdi_em if w_prop == 1 & ano==2007  
inequal tdi_em if w_prop == 1 & ano==2008  
inequal tdi_em if w_prop == 1 & ano==2009  
inequal tdi_em if w_prop == 1 & ano==2010  
inequal tdi_em if w_prop == 1 & ano==2011  
inequal tdi_em if w_prop == 1 & ano==2012  
inequal tdi_em if w_prop == 1 & ano==2013  
inequal tdi_em if w_prop == 1 & ano==2014  
inequal tdi_em if w_prop == 1 & ano==2015  
inequal tdi_em if w_prop == 1 & ano==2016  
inequal tdi_em if w_prop == 1 & ano==2017  
inequal tdi_em if w_prop == 1 & ano==2018
```

* Calculando o índice de Gini por localização no semiárido

** Estimando o índice de Gini para o PIB per capita

```
inequal pib_pc if w_semi == 0 & ano==2006  
inequal pib_pc if w_semi == 0 & ano==2007  
inequal pib_pc if w_semi == 0 & ano==2008  
inequal pib_pc if w_semi == 0 & ano==2009  
inequal pib_pc if w_semi == 0 & ano==2010
```

inequal pib_pc if w_semi == 0 & ano==2011
inequal pib_pc if w_semi == 0 & ano==2012
inequal pib_pc if w_semi == 0 & ano==2013
inequal pib_pc if w_semi == 0 & ano==2014
inequal pib_pc if w_semi == 0 & ano==2015
inequal pib_pc if w_semi == 0 & ano==2016
inequal pib_pc if w_semi == 0 & ano==2017
inequal pib_pc if w_semi == 0 & ano==2018

inequal pib_pc if w_semi == 1 & ano==2006
inequal pib_pc if w_semi == 1 & ano==2007
inequal pib_pc if w_semi == 1 & ano==2008
inequal pib_pc if w_semi == 1 & ano==2009
inequal pib_pc if w_semi == 1 & ano==2010
inequal pib_pc if w_semi == 1 & ano==2011
inequal pib_pc if w_semi == 1 & ano==2012
inequal pib_pc if w_semi == 1 & ano==2013
inequal pib_pc if w_semi == 1 & ano==2014
inequal pib_pc if w_semi == 1 & ano==2015
inequal pib_pc if w_semi == 1 & ano==2016
inequal pib_pc if w_semi == 1 & ano==2017
inequal pib_pc if w_semi == 1 & ano==2018

** Estimando o índice de Gini para a distribuição do FNE per capita

inequal fne_percapita if w_semi == 0 & ano==2006
inequal fne_percapita if w_semi == 0 & ano==2007
inequal fne_percapita if w_semi == 0 & ano==2008
inequal fne_percapita if w_semi == 0 & ano==2009
inequal fne_percapita if w_semi == 0 & ano==2010
inequal fne_percapita if w_semi == 0 & ano==2011
inequal fne_percapita if w_semi == 0 & ano==2012
inequal fne_percapita if w_semi == 0 & ano==2013
inequal fne_percapita if w_semi == 0 & ano==2014
inequal fne_percapita if w_semi == 0 & ano==2015
inequal fne_percapita if w_semi == 0 & ano==2016
inequal fne_percapita if w_semi == 0 & ano==2017
inequal fne_percapita if w_semi == 0 & ano==2018

inequal fne_percapita if w_semi == 1 & ano==2006
inequal fne_percapita if w_semi == 1 & ano==2007
inequal fne_percapita if w_semi == 1 & ano==2008
inequal fne_percapita if w_semi == 1 & ano==2009
inequal fne_percapita if w_semi == 1 & ano==2010
inequal fne_percapita if w_semi == 1 & ano==2011
inequal fne_percapita if w_semi == 1 & ano==2012
inequal fne_percapita if w_semi == 1 & ano==2013

```
inequal fne_percapita if w_semi == 1 & ano==2014  
inequal fne_percapita if w_semi == 1 & ano==2015  
inequal fne_percapita if w_semi == 1 & ano==2016  
inequal fne_percapita if w_semi == 1 & ano==2017  
inequal fne_percapita if w_semi == 1 & ano==2018
```

** Estimando o índice de Gini para os óbitos infantis por mil habitantes

```
inequal oi_percapita if w_semi == 0 & ano==2006  
inequal oi_percapita if w_semi == 0 & ano==2007  
inequal oi_percapita if w_semi == 0 & ano==2008  
inequal oi_percapita if w_semi == 0 & ano==2009  
inequal oi_percapita if w_semi == 0 & ano==2010  
inequal oi_percapita if w_semi == 0 & ano==2011  
inequal oi_percapita if w_semi == 0 & ano==2012  
inequal oi_percapita if w_semi == 0 & ano==2013  
inequal oi_percapita if w_semi == 0 & ano==2014  
inequal oi_percapita if w_semi == 0 & ano==2015  
inequal oi_percapita if w_semi == 0 & ano==2016  
inequal oi_percapita if w_semi == 0 & ano==2017  
inequal oi_percapita if w_semi == 0 & ano==2018
```

```
inequal oi_percapita if w_semi == 1 & ano==2006  
inequal oi_percapita if w_semi == 1 & ano==2007  
inequal oi_percapita if w_semi == 1 & ano==2008  
inequal oi_percapita if w_semi == 1 & ano==2009  
inequal oi_percapita if w_semi == 1 & ano==2010  
inequal oi_percapita if w_semi == 1 & ano==2011  
inequal oi_percapita if w_semi == 1 & ano==2012  
inequal oi_percapita if w_semi == 1 & ano==2013  
inequal oi_percapita if w_semi == 1 & ano==2014  
inequal oi_percapita if w_semi == 1 & ano==2015  
inequal oi_percapita if w_semi == 1 & ano==2016  
inequal oi_percapita if w_semi == 1 & ano==2017  
inequal oi_percapita if w_semi == 1 & ano==2018
```

** Estimando o índice de Gini para a taxa de distorção de idade-série no Ensino Médio

```
inequal tdi_em if w_semi == 0 & ano==2006  
inequal tdi_em if w_semi == 0 & ano==2007  
inequal tdi_em if w_semi == 0 & ano==2008  
inequal tdi_em if w_semi == 0 & ano==2009  
inequal tdi_em if w_semi == 0 & ano==2010  
inequal tdi_em if w_semi == 0 & ano==2011  
inequal tdi_em if w_semi == 0 & ano==2012  
inequal tdi_em if w_semi == 0 & ano==2013  
inequal tdi_em if w_semi == 0 & ano==2014
```

inequal tdi_em if w_semi == 0 & ano==2015
inequal tdi_em if w_semi == 0 & ano==2016
inequal tdi_em if w_semi == 0 & ano==2017
inequal tdi_em if w_semi == 0 & ano==2018

inequal tdi_em if w_semi == 1 & ano==2006
inequal tdi_em if w_semi == 1 & ano==2007
inequal tdi_em if w_semi == 1 & ano==2008
inequal tdi_em if w_semi == 1 & ano==2009
inequal tdi_em if w_semi == 1 & ano==2010
inequal tdi_em if w_semi == 1 & ano==2011
inequal tdi_em if w_semi == 1 & ano==2012
inequal tdi_em if w_semi == 1 & ano==2013
inequal tdi_em if w_semi == 1 & ano==2014
inequal tdi_em if w_semi == 1 & ano==2015
inequal tdi_em if w_semi == 1 & ano==2016
inequal tdi_em if w_semi == 1 & ano==2017
inequal tdi_em if w_semi == 1 & ano==2018



IPPDS
Instituto de Políticas Públicas e
Desenvolvimento Sustentável

ANEXO 4. ROTINAS DE PROGRAMAÇÃO EMPREGADAS NO PRODUTO 8

O Produto 8 da Meta 2, intitulado *Avaliação do retorno econômico e social do FNE*, lançou mão de ferramentas de programação matemática para a solução de problemas de otimização e a consequente obtenção dos resultados desejados. Em específico, foi utilizado o sistema de modelagem matemática *GAMS*. Abaixo são apresentadas as linhas de programação que formam a rotina empregada para a análise empírica da efetividade do FNE.

```

$title Modelo FNE em MPSGE
*** Data in R$ millions
$if not set ano $set ano 2008
$setglobal gdx      .\data_gdx\dados\
$set lstdir lst
$set resdir resultados
$if not dexist "%lstdir%"      $CALL mkdir "%lstdir%"
$if not dexist "%resdir%"     $CALL mkdir "%resdir%"

* Incluir dados:
$include read_data_model.gms

* Here the model in MPSGE starts:

$ontext
$model:sudene

$sectors:
    c(r)          ! Consumption
    g(r)          ! Government demand
    inv(r)        ! Investments
    y(i,r)$vom(i,r) ! Supply
    D(i,r)$vomd(i,r) ! Domestic output
    m(i,r)$vxmd(i,r) ! Imports from abroad
    x(i,r)$vxm(i,r) ! Exports to abroad
    mm(i,r)$vim(i,r) ! Interstate trade

$commodities:
    pc(r)          ! Private consumption price index
    pg(r)          ! Public consumption price index
    pinv(r)        ! Investment price index
    py(j,r)$vom(j,r) ! Output price index
    pd(j,r)$vomd(j,r) ! Domestic output price index
    pm(j,r)$vxmd(j,r) ! Import price index
    pf(f,r)$evom(f,r) ! Primary factors rent

```

px(j,r)\$v _{xm} (j,r)	! Export price index
ptii(i,r)\$fne(i,r) or fne_(i,r)	! FNE price index
px	! Exchange rate
pmm(i,r)	! Import price index among Brazilian regions
\$consumers:	
hh(r)	! Representative household in Brazil
govt(r)	! Representative government
\$auxiliary:	
unemp(r)\$u ₀ (r)	! Unemployment
KSTOCK(r)\$steady	! Capital stock
ocios(r)\$i ₀ (r)	! Capital unemployed
\$prod:y(j,r)\$vom(j,r) t:eta(j) s:esubd(j)	
o:py(j,r)	q:(vom(j,r) - vxm(j,r))
o:px(j,r)\$v _{xm} (j,r)	q:v _{xm} (j,r)
i:pd(j,r)	q:vomd(j,r)
i:pm(j,r)	q:v _{xm} d(j,r)
\$prod:D(j,r)\$vomd(j,r) t:0 s:0 i.tl:esubd(i) va:esubva(j)	
o:ptii(j,r)\$fne_(j,r)	q:(0.01*vomd(j,r))
o:pd(j,r)	q:vomd(j,r) a:govt(r) t:rto(r,j)
i:py(i,r)	q:vdfm(i,j,r) p:(1+rtfd0(i,j,r)) i.tl: a:govt(r) t:rtfd(i,j,r)
i:pmm(i,r)	q:(sum(s, vifm(i,j,s,r))) p:(1+rtfd0(i,j,r)) i.tl: a:govt(r) t:rtfd(i,j,r)
i:pf(f,r)	q:vfm(f,j,r) p:(1+rtf0(f,j,r)) va: a:govt(r) t:rtf(f,j,r)
i:ptii(j,r)\$fne(j,r)	q:(0.01*vomd(j,r))
\$prod:mm(i,r)\$vim(i,r) s:esubm(i)	
o:pmm(i,r)	q:vim(i,r)
i:py(i,s)	q:(vipm(i,s,r) + vigm(i,s,r) + viim(i,s,r) + sum(j, vifm(i,j,s,r)))
\$prod:c(r) s:1 i.tl:esubd(i)	
o:pc(r)	q:vpm(r)
i:py(i,r)	q:vdpm(i,r) i.tl: p:(1+rtpd0(i,r)) a:govt(r) t:rtpd(i,r)
i:pmm(i,r)	q:(sum(s, vipm(i,s,r))) i.tl: p:(1+rtpi0(i,r)) a:govt(r) t:rtpi(i,r)
\$prod:g(r) s:0 i.tl:esubd(i)	
o:pg(r)	q:vgm(r)
i:py(i,r)	q:vdgm(i,r) i.tl: p:(1+rtgd0(i,r)) a:govt(r) t:rtgd(i,r)
i:pmm(i,r)	q:(sum(s, vigm(i,s,r))) i.tl: p:(1+rtgd0(i,r)) a:govt(r) t:rtgd(i,r)
\$prod:inv(r) s:0 i.tl:esubd(i)	
o:pinv(r)	q:vinm(r)
i:py(i,r)	q:vdim(i,r) i.tl: p:(1+rtid0(i,r)) a:govt(r) t:rtid(i,r)

i:pmm(i,r) q:(sum(s, viim(i,s,r))) i.tl: p:(1+rtii0(i,r)) a:govt(r) t:rtii(i,r)

\$prod:m(i,r)\$vxmd(i,r)

o:pm(i,r) q:vxmd(i,r)
i:px q:vxmd(i,r)

\$prod:x(i,r)\$vxm(i,r)

o:px q:vxm(i,r)
i:px(i,r) q:vxm(i,r)

* Private household:

\$demand:hh(r)

d:pc(r) q:vpm(r)
e:pinv(r) q:(-vinm(r)) r:kstock(r)\$steady
e:pf(f,r)\$((u0(r) eq 0) and (i0(r) eq 0)) q:(evom(f,r)) r:kstock(r)\$sameas(f,"cap") and steady)

e:pf("cap",r)\$u0(r) and (i0(r) eq 0) q:evom("cap",r) r:kstock(r)\$steady
e:pf("cap",r)\$i0(r) q:(evom("cap",r)/(1-i0(r)))
e:pf("cap",r)\$i0(r) q:(-evom("cap",r)/(1-i0(r))) r:ocios(r)\$i0(r)
e:pf("lab",r)\$u0(r) q:(evom("lab",r)/(1-u0(r)))
e:pf("lab",r)\$u0(r) q:(-evom("lab",r)/(1-u0(r))) r:unemp(r)\$u0(r)
e:pc(r) q:(-r_d_f(r))
e:px q:(-trade_d(r))
e:px q:(-trade_i(r))
e:ptii(i,r)\$fne_(i,r) q:(-fne_(i,r))
e:ptii(i,r)\$fne(i,r) q:(fne(i,r))

* Government:

\$demand:govt(r)

d:pg(r) q:vgm(r)
e:pc(r) q:(-r_d_g(r))

\$constraint:unemp(r)\$u0(r)

pf("lab",r) =e= pc(r);

\$constraint:kstock(r)\$steady

* pf("cap",r) =e= pf("lab",r)/pc(r);
pf("cap",r) =e= pinv(r)/pc(r);

\$constraint:ocios(r)\$i0(r)

pf("cap",r) =e= pc(r);

\$report:

v:vom_(i,r) o:py(i,r) prod:y(i,r)

```

v:vxm_(i,r)      o:px(i,r)   prod:y(i,r)
v:vxmmd_(i,r)   i:pm(i,r)   prod:y(i,r)

* v:vbam_(i,r)   o:pt      prod:d(i,r)
  v:vomd_(i,r)   i:pd(i,r)   prod:y(i,r)
  v:vfmm_(i,f,r) i:pf(f,r)   prod:d(i,r)
  v:vbami(i,r)$fne(i,r) o:ptii(i,r) prod:d(i,r)

v:vim_(i,r)      o:pmm(i,r)   prod:mm(i,r)
v:vxxm_(i,s,r)   i:py(i,s)   prod:mm(i,r)

v:vpm_(r)        o:pc(r)     prod:c(r)
v:vdpm_(i,r)     i:py(i,r)   prod:c(r)
v:vdpmm_(i,r)    i:pmm(i,r)   prod:c(r)

v:vgm_(r)        o:pg(r)     prod:g(r)
v:vdgm_(i,r)     i:py(i,r)   prod:g(r)
v:vdgmm_(i,r)    i:pmm(i,r)   prod:g(r)

v:vinm_(r)       o:pinv(r)    prod:inv(r)
v:vdim_(i,r)     i:py(i,r)    prod:inv(r)
v:vdimm_(i,r)    i:pmm(i,r)    prod:inv(r)

v:vxmd_(i,r)     o:pm(i,r)    prod:m(i,r)
v:vxmm_(i,r)     i:px(i,r)    prod:x(i,r)

* v:lab_fl_(r)$lab_fl(r) i:pf("lab",r) prod:cc(r)

$offtext
$sysinclude mpsgeset sudene

***!!! Início do loop

*loop(t_,

fne(i,r) = 0;
fne_(i,r) = 0;

* Acrescentar desemprego:

u0(r) = 0.1;
*u0(r) = desoc(t_,"a")/100;

i0(r) = 0;

*u0(r) = 0;

```

```
steady = 1;
```

```
unemp.l(r)$u0(r) = u0(r);
```

```
KSTOCK.L(r) = 1;
```

```
KSTOCK.FX(r)$ (NOT STEADY) = 1;
```

```
*unemp.lo(r)$u0(r) = u0(r)*3/4;
```

```
*4/5;
```

* Acrescentar trade-off lazer x trabalho

```
lab_fl(r) = 0;
```

```
lab_lei(r) = 0;
```

* Considerando o desemprego do IBGE de 7,8% em 2008:

```
lab_lei(r) = 0.621;
```

```
lab_l(r) = 0;
```

* Desempregados considerando a taxa de desemprego do IBGE de 7,8% em 2008:

```
lab_l(r)$lab_fl(r) = evom("lab",r)/11.82;
```

* Incluir dados:

```
*$include dados_.gms
```

```
*evom("lab",r)$lab_fl(r) = evom("lab",r) + lab_l(r);
```

```
sudene.workspace = 128;
```

```
sudene.iterlim = 0;
```

```
$include sudene.gen
```

```
solve sudene using mcp;
```

* Clean-up run:

```
sudene.ITERLIM = 60000;
```

```
$INCLUDE sudene.GEN
```

```
SOLVE sudene USING MCP;
```

```
display vpm_.l, vgm_.l, vom_.l, vxm_.l;
```

* Store initial value of some variables:

```
vom0(i,r) = vom_.l(i,r);
```

```
vxm0(i,r) = vxm_.l(i,r);
```

```
vxmmd0(i,r) = vxmmd_.l(i,r);
```

```

*vbam0(i,r) = vbam_.l(i,r);
vomd0(i,r) = vomd_.l(i,r);
vfm0(i,f,r) = vfm_.l(i,f,r);
vbam0(i,r) = vbami.l(i,r);

vim0(i,r) = vim_.l(i,r);
vxxm0(i,s,r) = vxxm_.l(i,s,r);

vpm0(r) = vpm_.l(r);
vdpm0(i,r) = vdpm_.l(i,r);
vdpmm0(i,r) = vdpmm_.l(i,r);

vgm0(r) = vgm_.l(r);
vdgm0(i,r) = vdgm_.l(i,r);
vdgmm0(i,r) = vdgmm_.l(i,r);

vinm0(r) = vinm_.l(r);
vdim0(i,r) = vdim_.l(i,r);
vdimm0(i,r) = vdimm_.l(i,r);

vxmd0(i,r) = vxmd_.l(i,r);
vxmm0(i,r) = vxmm_.l(i,r);

fne_rep(r,i,"FNE (R$ Milhoes)","%ano%") = fne_gt(r,i,"%ano%")*vom(i,r);
fne_rep(r,"tot","FNE (R$ Milhoes)","%ano%") = sum(i, fne_gt(r,i,"%ano%")*vom(i,r));
fne_rep(r,i,"FNE (% VBP)","%ano%") = fne_gt(r,i,"%ano%")*100;
fne_rep(r,"tot","FNE (% VBP)","%ano%") = fne_rep(r,"tot","FNE (R$
Milhoes)","%ano%")/
sum(i, vom(i,r))*100;

pib(r,"va_1") = sum((i,f), vfm0(i,f,r))+sum(j, rto(r,j)*vomd(j,r));
pib("bra","va_1") = sum(r, pib(r,"va_1"));

fne_rep(r,"tot","FNE (% PIB)","%ano%") = sum(i, fne_rep(r,i,"FNE (R$
Milhoes)","%ano%"))
/pib(r,"va_1")*100;
fne_rep("bra","tot","FNE (% PIB)","%ano%") = sum((r,i), fne_rep(r,i,"FNE (R$
Milhoes)","%ano%"))
/sum(r, pib(r,"va_1"))*100;

fne_rep_pib(r,"FNE (% PIB)","%ano%") = fne_rep(r,"tot","FNE (% PIB)","%ano%");
fne_rep_pib("bra","FNE (% PIB)","%ano%") = fne_rep("bra","tot","FNE (%
PIB)","%ano%");

display fne_rep, fne_rep_pib, pib;

```

```

*#   Aplicar o choque aqui:
*   Acrescentar o montante recebido do FNE:
fne_(i,nr) = 0.01*vomd(i,nr)*(1+fne_gt(nr,i,"%ano%"));

$ontext
fne_gt(r,i)$stat(r,"pib") = (fne_gt(r,i))/1000*pib0(r,"df")/stat(r,"pib");
fne_ch(r,i)$vomd(i,r) = (fne_gt(r,i))/vomd(i,r);

fne_rep(r,i,"fne_$") = fne_gt(r,i);
fne_rep(r,"tot","fne_$") = sum(i, fne_gt(r,i));
fne_rep(r,i,"VP") = vomd(i,r);
fne_rep(r,"tot","VP") = sum(i, vomd(i,r));
fne_rep(r,i,"%fne/VP")$vomd(i,r) = fne_gt(r,i)/vomd(i,r)*100;
fne_rep(r,"tot","%fne/VP") = sum(i, fne_gt(r,i))/sum(i,vomd(i,r))*100;
fne_rep(r,i,"VA") = sum(f, vfm(f,i,r));
fne_rep(r,"tot","VA") = sum((i,f), vfm(f,i,r));
fne_rep(r,i,"%fne/VA")$sum(f, vfm(f,i,r)) = fne_gt(r,i)/sum(f, vfm(f,i,r))*100;
fne_rep(r,"tot","%fne/VA") = sum(i, fne_gt(r,i))/sum((i,f), vfm(f,i,r))*100;
display fne_rep;
$offtext

*   Solve the policy case:

*unemp.lo(r)$u0(r) = u0(r)*3/4;
*unemp.lo(r)$u0(r) = 0;
*kstock.up(r)$steady = 1.2;

$include sudene.gen
    solve sudene using mcp;

pib(r,"va_0") = pib0(r,"va");
pib(r,"va_1") = sum((i,f), vfm0(i,f,r))+sum(j, rto(r,j)*vomd(j,r));
pib(r,"va_f") = sum((i,f), vfm_l(i,f,r)) + vgm_l(r) + r_d_g(r) ;
pib(r,"va_ch") = (pib(r,"va_f")/pib(r,"va_0") - 1)*100;
pib(r,"df_0") = pib0(r,"df");
pib("bra","df_0") = sum(r, pib(r,"df_0"));
pib(r,"df_1") = vpm0(r)+vgm0(r)+vinm0(r)+sum(i, vxm0(i,r))
    +sum((i,s), vxxm0(i,r,s))
    - sum(i, vxmmd0(i,r))
    - sum((i,s), vxxm0(i,s,r));
pib(r,"df_f") = vpm_l(r)+vgm_l(r)+vinm_l(r)+sum(i, vxm_l(i,r))
    +sum((i,s), vxxm_l(i,r,s))
    - sum(i, vxmmd_l(i,r))
    - sum((i,s), vxxm_l(i,s,r));
pib("bra","df_f") = sum(r, pib(r,"df_f"));

```

```
pib(r,"df_ch") = (pib(r,"df_f") / pib(r,"df_0") - 1)*100;
```

```
display pib;
```

```
ev(r,"ch_w_%") = round(100 * (C.L(r)-1),3);
ev(r,"ch_w_bi$") = round(vpm(r) * (C.L(r)-1),3);
ev("bra","ch_w_bi$") = sum(r, ev(r,"ch_w_bi$"));
ev("bra","ch_w_%") = round(100*( ev("bra","ch_w_bi$")/sum(r, vpm(r))),2);
ych(j,r) = round(100 * (d.l(j,r) - 1),3);
gch(r) = round(100 * (g.l(r) - 1),3);
```

```
display ych, ev, gch;
```

* Calculate change in trade flows:

```
pcttr(i,r)$vxmd0(i,r) = round(100*(vxmd_.l(i,r)/vxmd0(i,r) - 1));
brexp(i,s,r) = vxxm_.l(i,s,r);
brimp(i,s,r) = vxxm_.l(i,r,s);
```

* Changes in total exports and imports

```
tpctexp(i,r)$vxmm0(i,r) = (vxmm_.l(i,r)/vxmm0(i,r) - 1)*100;
tpctimp(i,r)$vxmd0(i,r) = (vxmd_.l(i,r)/vxmd0(i,r) - 1)*100;
```

* Change in PIB (PIB = private consumption + public consumption + investments + exports - imports):

```
chpib(r) = round(100*(((pc.l(r)*vpm_.l(r) + pg.l(r)*vgm_.l(r) + pinv.l(r)*vinm_.l(r)
+ sum(i, px.l(i,r)*vxm_.l(i,r)) + sum((i,s), py.l(i,r)*vxxm_.l(i,r,s))
- sum(i, pm.l(i,r)*vxmmd_.l(i,r)) - sum((i,s), py.l(i,r)*vxxm_.l(i,s,r)))
/ pc.l(r)) /
(vpm0(r)+vgm0(r)+vinm0(r)+sum(i, vxm0(i,r))
+sum((i,s), vxxm0(i,r,s))
- sum(i, vxmmd0(i,r))
- sum((i,s), vxxm0(i,s,r)))-1),2);
```

```
chpib_r(r,"%ch_C") = round(100*((pc.l(r)*vpm_.l(r)/pc.l(r)) / vpm0(r) - 1), 3);
chpib_r(r,"%ch_G") = round(100*((pg.l(r)*vgm_.l(r)/pg.l(r)) / vgm0(r) - 1), 3);
chpib_r(r,"%ch_I.") = round(100*((pinv.l(r)*vinm_.l(r)/pc.l(r)) / vinm0(r) - 1), 3);
chpib_r(r,"%ch_X.") = round(100*(((sum(i, px.l(i,r)*vxm_.l(i,r)) + sum((i,s),
py.l(i,r)*vxxm_.l(i,r,s))) / pc.l(r))
/ (sum(i, vxm0(i,r))+sum((i,s), vxxm0(i,r,s)))-1), 3);
chpib_r(r,"%ch_M.") = round(100*(((sum(i, pm.l(i,r)*vxmmd_.l(i,r)) + sum((i,s),
py.l(i,r)*vxxm_.l(i,s,r))) / pc.l(r))
/ (sum(i, vxmmd0(i,r)) + sum((i,s), vxxm0(i,s,r)))-1), 3);
chpib_r(r,"%ch_PIB") = chpib(r);
```

```

chpib_r("C: Private Consumption", "%ch_PIB")= eps;
chpib_r("G: Government Consumption", "%ch_PIB")= eps;
chpib_r("I: Investment", "%ch_PIB")= eps;
chpib_r("X: Exports", "%ch_PIB")= eps;
chpib_r("M: Imports", "%ch_PIB")= eps;

pibr(r, "Bs_C") = round(vpm0(r), 2);
pibr(r, "Bs_G") = round(vgm0(r), 2);
pibr(r, "Bs_I") = round(vinm0(r), 2);
pibr(r, "Bs_X") = round(sum(i, vxm0(i,r))
+ sum((i,s), vxxm0(i,r,s)), 2);
pibr(r, "Bs_M") = round(sum(i, vxmmd0(i,r))
+ sum((i,s), vxxm0(i,s,r)), 2);
pibr(r, "Bs_PIB") = round(pibr(r, "Bs_C")
+ pibr(r, "Bs_G") + pibr(r, "Bs_I")
+ pibr(r, "Bs_X") - pibr(r, "Bs_M"), 2);
pibr(r, "Up_C") = round(vpm_l(r), 2);
pibr(r, "Up_G") = round(pg.l(r)*vgm_l(r)/pc.l(r), 2);
pibr(r, "Up_I") = round(pinv.l(r)*vinm_l(r)/pc.l(r), 2);
pibr(r, "Up_X") = round(((sum(i, px.l(i,r)*vxm_l(i,r)) + sum((i,s), py.l(i,r)*vxxm_l(i,r,s)))
/ pc.l(r)), 2);
pibr(r, "Up_M") = round(((sum(i, pm.l(i,r)*vxmmd_l(i,r)) + sum((i,s),
py.l(i,r)*vxxm_l(i,s,r))) / pc.l(r)), 2);
pibr(r, "Up_PIB") = round(pibr(r, "Up_C")
+ pibr(r, "Up_G") + pibr(r, "Up_I")
+ pibr(r, "Up_X") - pibr(r, "Up_M"), 2);
pibr("Bs: base data", "Up_PIB") = eps;
pibr("Up: updated data", "Up_PIB") = eps;
pibr("C: Private Consumption", "Up_PIB")= eps;
pibr("G: Government Consumption", "Up_PIB")= eps;
pibr("I: Investment", "Up_PIB")= eps;
pibr("X: Exports", "Up_PIB")= eps;
pibr("M: Imports", "Up_PIB")= eps;
pibr("* Data in 2004 US$ bi", "Up_PIB")= eps;

vom_rep(r,i,"Bs($bi)") = round(vomd0(i,r), 3);
vom_rep(r,i,"Up($bi)") = round(vomd_l(i,r), 3);
vom_rep(r,i,"%ch")$vomd0(i,r) = round((vomd_l(i,r)/vomd0(i,r) - 1)*100, 3);

pcch("r",r) = round(100*(pc.l(r)/pc.l(r)-1),3);
pcch_(r) = round(100*(pc.l(r)-1),3);
pcch("n",r) = pcch_(r);

psych("r",j,r) = round(100*(py.l(j,r)/pc.l(r)-1),3);
psych_(j,r) = round(100*(py.l(j,r)-1),3);
psych("n",j,r) = psych_(j,r);

```

```

pfch(f,"r",r) = round(100*((pf.l(f,r)/pc.l(r))-1),3);
pfch_(f,r) = round(100*(pf.l(f,r)-1),3);
pfch(f,"n",r) = pfch_(f,r);

pmch("r",i,r) = round(100*(pm.l(i,r)/pc.l(r)-1),3);
pmch_(i,r) = round(100*(pm.l(i,r)-1),3);
pmch("n",i,r) = pmch_(i,r);

ocup(r,i,"sal0") = vfm("lab",i,r);
ocup(r,"tot","sal0") = sum(i, vfm("lab",i,r));
ocup(r,i,"sal1") = pf.l("lab",r)/pc.l(r)*vfm_.l(i,"lab",r);
ocup(r,"tot","sal1") = sum(i, ocup(r,i,"sal1"));
ocup(r,i,"ch_sal%")$vfm0(i,"lab",r) = (ocup(r,i,"sal1")/ocup(r,i,"sal0") - 1)*100;
ocup(r,"tot","ch_sal%") = (ocup(r,"tot","sal1")/ocup(r,"tot","sal0") - 1)*100;
*ocup(r,i,"mdo0") = vad(r,"pes",i);
*ocup(r,"tot","mdo0") = sum(i, vad(r,"pes",i));
ocup(r,i,"ch_mdo_$") = vfm_.l(i,"lab",r) - vfm("lab",i,r);
ocup(r,i,"ch_mdo_$%") = 100*(vfm_.l(i,"lab",r)/vfm("lab",i,r) - 1);
ocup(r,"tot","ch_mdo_$") = (evom("lab",r)/(1-u0(r)))*unemp.l(r);
ocup(r,"tot","ch_mdo%")$u0(r) = -100*(unemp.l(r)/u0(r)-1) ;
ocup(r,"tot","unempl") = unemp.l(r);
*ocup(r,"tot","ch_mdo%") = ocup(r,"tot","ch_sal%") - (pf.l("lab",r)-1)*100;
*ocup(r,i,"mdo1") = ocup(r,i,"mdo0")*(1+ocup(r,i,"ch_mdo%")/100);
*ocup(r,"tot","mdo1") = sum(i, ocup(r,i,"mdo1"));
*ocup(r,i,"ch_mdo_ab") = ocup(r,i,"mdo1") - ocup(r,i,"mdo0");
*ocup(r,"tot","ch_mdo_ab") = ocup(r,"tot","mdo1") - ocup(r,"tot","mdo0");

ev(r,"ch_pib%") = chpib(r);
ev("bra","ch_pib%") = round((sum(r, pib(r,"df_f"))/sum(r, pib(r,"df_0"))) - 1)*100,3);
ev(r,"ch_Gov%") = chpib_r(r,"%ch_G");

gov_rep("chg",i,r) = vdgm_.l(i,r) - vdgm0(i,r);
gov_rep("chg","tot",r) = sum(i, vdgm_.l(i,r)) - sum(i, vdgm0(i,r));
gov_rep("chg","tot2",r) = vgm_.l(r) - vgm(r);
gov_rep("val_chg",i,r) = py.l(i,r)*vdgm_.l(i,r) - vdgm0(i,r);
gov_rep("val_chg","tot",r) = sum(i, py.l(i,r)*vdgm_.l(i,r)) - sum(i, vdgm0(i,r));
gov_rep("val_chg","tot2",r) = pg.l(r)*vgm_.l(r) - vgm0(r);

y_mult(r,i)$fne_rep(r,i,"FNE (R$ Milhoes)","%ano%") = (vomd_.l(i,r)-vomd0(i,r))/fne_rep(r,i,"FNE (R$ Milhoes)","%ano%");
pib_mult(r)$sum(i, fne_rep(r,i,"FNE (R$ Milhoes)","%ano%")) = (pibr(r,"Up_PIB")-pibr(r,"Bs_PIB")) / sum(i, fne_rep(r,i,"FNE (R$ Milhoes)","%ano%"));

option ev:3, ych:3, gch:3;
display ev, ych, gch;
display tpctexp, tpctimp, brexp, brimp, chpib, ocup, gov_rep;

```

```
display pcch, pcch_, pych, pych_, pfch, pfch_, pmch, pmch_;
display chpib_r, pibr, y_mult, pib_mult;
```

****## Parametros para o relatorio do FNO**

```
tabela_1(r,i,"FNE (R$ Milhoes)") = fne_gt(r,i,"%ano%")*vom(i,r);
tabela_1(r,"Tot","FNE (R$ Milhoes)") = fne_rep(r,"tot","FNE (R$ Milhoes)","%ano%");
tabela_2(r,i,"FNE/VBP (%)") = fne_rep(r,i,"FNE (% VBP)","%ano%");
tabela_2(r,"Tot","FNE/VBP (%)") = fne_rep(r,"tot","FNE (% VBP)","%ano%");
tabela_3(r,"FNO/PIB estadual (%)") = fne_rep(r,"tot","FNE (% PIB)","%ano%");
tabela_4(r,i,"Impacto FNO na producao (%)") = ych(i,r);
tabela_5(r,i,"Impacto FNO no preco (%)") = pych("r",i,r);
tabela_6(f,r,"R$ milhoes","Impacto FNO fatores") = pf.l(f,r)/pc.l(r)*sum(i, vfm_.l(i,f,r)) -
sum(i, vfm(f,i,r));
tabela_6(f,r,"%","Impacto FNO fatores") = 100*(pf.l(f,r)/pc.l(r)*sum(i, vfm_.l(i,f,r))/sum(i,
vfm(f,i,r))-1);
tabela_7(r,"R$ milhoes","Impacto FNO PIB") = pibr(r,"Up_PIB") - pibr(r,"Bs_PIB");
tabela_7(r,"%","Impacto FNO PIB") = chpib(r);
tabela_8(r,"Multiplicador PIB/FNO") = pib_mult(r);
tabela_9(r,"R$ milhoes","Impacto FNO bem-estar") = ev(r,"ch_w_bi$");
tabela_9(r,"%","Impacto FNO bem-estar") = ev(r,"ch_w_%");
tabela_10(r,"Impacto FNO arrecadacao governo") = chpib_r(r,"%ch_G");

execute_unload "resultados\%ano%.gdx" tabela_1, tabela_2, tabela_3, tabela_4, tabela
```