

ESCASSEZ DE ÁGUA E MORTALIDADE INFANTIL: NOVAS EVIDÊNCIAS A PARTIR DE UM PAINEL QUANTÍLICO ROBUSTO

Área 10 - Economia Regional e Urbana

Erik Alencar de Figueiredo

Doutor em Economia pela UFRGS e Pós-Doutor pela University Of Tennessee – PPGE/UFPB

E-mail: eafigueiredo@gmail.com

Fernanda Leite Santana

Doutora em Economia Aplicada – PPGE/UFPB

E-mail: fleitesantana@gmail.com

Wallace Patrick Santos de Farias Souza

Doutorando em Economia Aplicada – PPGE-UFRGS

E-mail: wpsfarias@gmail.com

RESUMO: Propõe-se investigar os resultados de saúde expressos na taxa de mortalidade infantil para um painel de município-a-mês (1996-2010) do semiárido brasileiro, levando em consideração importantes efeitos: heterogeneidade não observada, a endogeneidade existente, bem como a confiabilidade dos dados de precipitação, até o momento negligenciado pela literatura. Os resultados, em conformidade com a literatura, mostram um efeito negativo da precipitação e um efeito positivo da seca sobre a taxa de mortalidade infantil dos municípios pertencentes ao semiárido brasileiro. Contudo, ressalta-se que os efeitos de tratamento aplicados ratifica o efeito negativo, porém, em magnitude inferior ao apresentado na literatura.

Palavras-chaves: Taxa de Mortalidade Infantil; Precipitação; Saúde; Semiárido Brasileiro

ABSTRACT: We propose to investigate health outcomes expressed in child mortality rate for a county-by-month panel (1996-2010) of the Brazilian semiarid, taking into account important effects: unobserved heterogeneity, the existing endogeneity and reliability of precipitation data so far neglected in the literature. The results, in accordance with the literature, show a negative effect of precipitation and a positive effect of drought on the child mortality rate in the municipalities belonging to the Brazilian semiarid region. However, it is noteworthy that the effects of treatment applied confirms the negative effect, however, to a lesser extent than shown in the literature.

Keywords: Infant Mortality Rate; Precipitation; Health; Brazilian semiarid

JEL: R10, C23, I10

1 INTRODUÇÃO

Há um grande número de pesquisas que investigam os impactos das condições iniciais sobre os determinantes dos resultados socioeconômicos dos indivíduos quando adultos (Barcker, 1998; Gluckman e Handson, 2005; Cunha e Heckman, 2007; Heckman, 2007). Dentro desse contexto, choques climáticos nos anos iniciais de vida, sobretudo relacionados às quantidades pluviométricas, tem desempenhado papel relevante nos estudos sobre o desenvolvimento humano (Maccini e Yang, 2006; Ugaz e Zanolini, 2011; Burguessn et. al, 2014).

Dentre as principais consequências, ressaltam-se o efeito negativo em: i) processos produtivos (Aguilar e Vicarelli, 2011; Bezabih, Di Falco e Mekonnen, 2014); ii) na saúde (Currie e Vogl, 2012; Currie et. al., 2013; Currie e Rossim-Slter, 2013; Rocha e Soares, 2015); iii) bem como na acumulação de capital humano (Shah e Steinberg, 2013). O pressuposto básico é de que os choques climáticos propiciam ambientes iniciais (des) favoráveis, uma vez que há efeito negativo/positivo na cadeia produtiva, na saúde e na qualidade de vida da região (Cirilo, 2008).

Nesse íterim, a questão pluviométrica vem ganhando destaque em estudos recentes. Por exemplo, Shah e Stenberg (2013) ressaltam que exposição a períodos de seca no início da vida tem efeitos prejudiciais nos salários, na saúde e na escolaridade dos indivíduos adultos da área rural da Índia. Burgess, et. al. (2014) ratificam esses resultados para a Índia em termos de mortalidade em populações rurais e urbanas, bem como altas temperaturas reduzindo salários agrícolas e reais. No Brasil, Rocha e Soares (2015) enfatizam a escassez da água na região semiárida do nordeste como fator determinante para a mortalidade infantil mais elevada, menor peso ao nascer, e períodos de gestação mais curtos.

Por outro lado, Maccini e Yang (2008) investigam o efeito das condições climáticas na época do nascimento sobre a saúde, educação e resultados socioeconômicos dos adultos indonésios e concluem que uma maior precipitação no início da vida tem grandes efeitos positivos sobre os resultados de mulheres adultas, mas não para homens. Este conjunto de evidências vem atribuindo determinado grau de importância para as questões pluviométricas no desenvolvimento humano destas regiões, contudo, é relevante ressaltar que estas estimativas podem estar viesadas se não forem observadas questões como confiabilidade de dados de clima e a endogeneidade existente nessas relações.

Em que pese o desenvolvimento recente da associação de clima no desenvolvimento humano dos indivíduos, estes questionamentos merecem especial atenção para não incorrer em possíveis sub ou superestimações de resultados, ou ainda, em erros de especificação. Tais tratamentos são importantes na definição do correto impacto das mudanças climáticas sob resultados de saúde, crucial para os formuladores de políticas públicas. Deste modo, o objetivo desta pesquisa é adicionar elementos à literatura existente no intuito de reduzir incertezas e fornecer *insights* sobre os resultados de saúde dos indivíduos enfatizando três conjunturas.

Primeiro, considera-se que as estimativas do efeito médio da variável dependente podem não retratar a realidade da maioria dos municípios. A explicação para tal é a existência de heterogeneidade entre os municípios que a regressão clássica de dados de painel não capta. Potencialmente, modelos de regressão quantílica poderiam explicar a heterogeneidade não observada e efeitos de localização, e, concomitantemente a inclusão dos efeitos fixos dos dados em painel comportando, assim, um modelo de regressão com estimadores mais robustos. Nesse sentido, propõe-se averiguar as implicações de restrições pluviométricas sobre os períodos iniciais de vida dos indivíduos ao longo da distribuição utilizando o método de regressões quantílicas em dados de painel.

Segundo, pretende-se ampliar a análise. O ponto central é a omissão de variáveis, uma vez que muitas pesquisas vêm atribuindo ao fator clima especial ênfase, e por ora deixando de levar em conta outros fatores que podem afetar a mudança dos padrões de saúde que não seja clima. Por exemplo, a distribuição geográfica e a sazonalidade de doenças transmitidas podem alterar resultados de saúde e não

necessariamente a causa disso estar nas mudanças climáticas. Para tratar esses problemas empregou-se o método de Harding e Lamarche (2009), que seja um estimador de regressão quantílica de um modelo de dados em painel com variáveis independentes endógenas, no qual permite que a variável endógena possa ser correlacionada com fatores não observados que afetam a variável resposta.

Para a consecução do método necessitar-se-á de um vetor que denote uma variável que afeta o nível de precipitação, mas não tem efeito direto sobre os resultados de saúde dos indivíduos. Pesquisas tem mostrado que as condições oceânicas, sobretudo a Temperatura da Superfície do Mar (TSM), apresentam-se como a principal variável física a influenciar as condições climáticas oriundas de dois fenômenos climáticos: El Niño¹ – Oscilação Sul (ENOS) e Dipolo Atlântico, tornando-se, portanto, o instrumento utilizado na pesquisa. Neste caso captam-se oscilações dos Oceanos Atlântico e Pacífico que afetam a ocorrência de chuvas no semiárido brasileiro.

Terceiro, e não menos importante, pretende-se averiguar a base de dados de clima, em cenário com dados observacionais², estes considerados na literatura meteorológica e climatológica como mais confiáveis, uma vez que se verifica que boa parte da literatura supracitada utiliza dados históricos de variações de chuvas obtidas por técnicas de reanálise³ – chamados de dados modelados. Embora ambas as variáveis reproduzam a climatologia e as variabilidades das regiões de interesse, há um grande debate⁴ acerca da precisão dos dados de reanálise em razão de sub/superestimações de intensidade de chuvas na média. Portanto, objetiva-se testar o impacto das chuvas nos primeiros anos de vida utilizando dados observacionais obtidos no Banco de Dados Meteorológicos para Ensino e Pesquisa (BDMEP) através de estimação de dados de painel com efeitos fixos.

Em síntese, a estratégia empírica adotada basear-se-á em três etapas. Primeiro, baseado em Canay (2011), estimar-se-á por regressão quantílica com dados em painel considerando a heterogeneidade não observada e os efeitos das covariáveis em ambientes com efeito fixo. Segundo, visando endogeneizar a quantidade de precipitações ocorridas no semiárido do Nordeste brasileiro, utilizar-se-á a Temperatura da Superfície do Mar (TSM), como instrumento em um modelo empírico baseado em Harding e Lamarche (2009). Por último, na terceira etapa, com a base dados obtidos de estações meteorológicas observacionais, realizar-se-á estimação para comparação entre as duas diferentes bases a fim de comprovar sub/superestimações dos resultados de Rocha e Soares (2015).

Além desta introdução, o restante do trabalho está organizado em mais quatro seções. A próxima seção apresenta a estratégia empírica, bem como a descrição dos dados e concepção das variáveis utilizadas nas estimações. Na terceira seção são discutidos os resultados apurados e a quarta parte reservada para as considerações finais.

2 METODOLOGIA E DADOS

Esta seção destina-se a apresentar a estratégia adotada neste estudo. Inicialmente, apresenta-se a contexto teórico e posteriormente o modelo empírico para avaliar o impacto das restrições pluviométricas nos primeiros anos de nascimento nos resultados de saúde dos indivíduos, nascidos no semiárido na década de 90. A seguir, discorre-se sobre dois métodos adotados. O primeiro, baseado em Canay (2011) visa apresentar modelo de regressão quantílica com dados em painel com efeitos fixos. Por sua vez, o

¹ Ver Kayano, Rao e Moura (1988), Rao, Lima e Franchito (1993), Xavier et. al. (2003).

² Dados observacionais são oriundos de estações meteorológicas de superfície. De acordo com Vieira e Piculli (2009), há dois tipos de estações: i) convencionais – onde os dados são coletados e organizados por um observador; e ii) automáticas – a coleta é feita através de sensores.

³ O processo de reanálise captam dados oriundos de satélites/estações, no qual acontece por duas etapas. Primeiro são interpolados através de modelos utilizados para prever tempo e, posteriormente, novos dados são interpolados para áreas em que não possuem dados disponíveis (Freitas, Franchito e Rao, 2010).

⁴ Ver Guedes et. al., 2005; Pinto et. al., 2009; Freitas, Franchito e Rao, 2010

segundo baseia-se em Harding e Lamarche (2009) numa abordagem de regressão quantílica para modelos de dados em painel com variáveis endógenas e efeitos individuais associados com as variáveis independentes. Por fim, expõem-se os dados utilizados, bem como a variável instrumental Temperatura da Superfície do Mar (TSM).

2.1 Contexto

O estudo da literatura da economia da saúde remonta a análise comportamental desenvolvido por Grossman (1972) que, análogo à literatura sobre o capital humano, considera a saúde como resultante de um processo de produção, portanto, uma variável de estoque. A partir deste escopo, desenvolvimentos recentes relacionados a literatura da psicologia, neurociência cognitiva, economia e epidemiologia tem destacado que a condição inicial depende das características genéticas, mas também das condições iniciais vivenciadas pelos indivíduos.

Nesse contexto, o modelo teórico de Heckman (2007) é aqui utilizado e adaptado para fins desta pesquisa, e tem como base a teoria do ciclo da vida, em que apresenta como principal característica o desenvolvimento de um processo dinâmico de múltiplos estágios, ou seja, o ciclo de vida dos indivíduos é analisado considerando vários períodos e cada qual com uma taxa de retorno diferente. A ideia é mostrar que o desenvolvimento da criança depende de um conjunto de fatores.

Conforme Heckman (2007) esta tecnologia é suficientemente rica para descrever a evidência de que a capacidade da criança em ser mais saudável afeta o resultados de saúde posteriores. Tal constatação é decorrente de duas características importantes da tecnologia: i) **autoprodutividade** – habilidades produzidas em um estágio aumentam as habilidades desenvolvidas nos estágios posteriores, e surge quando $\partial f_t(h, \theta_t, I_t)/\partial \theta_t > 0$; e ii) **complementariedade dinâmica** – habilidades produzidas em um estágio aumentam a produtividade do investimento em estágios subsequentes e surge quando $\partial^2 f_t(h, \theta_t, I_t)/\partial \theta_t \partial I'_t > 0$.

Além disso, os autores afirmam que esta tecnologia capta os períodos sensíveis e críticos postulados na literatura psicológica e da neurociência. Em outras palavras, o período de t^* é um período sensível em relação ao período de s se, ao mesmo nível de insumos, o investimento é mais produtivo na fase t^* do que em outro estágio $s \neq t^*$. A ideia é de que no **período sensível** haja certa estabilidade na conectividade dos neurônios decorrente da influência do ambiente, assim indivíduos nascidos em ambientes restritos tendem a produzir habilidades menos eficazes. Mesmo que se invistam períodos mais tardes, denominados de **períodos críticos**, este nunca será um substituto perfeito. Nesse contexto, o objetivo da pesquisa é investigar os efeitos das mudanças pluviométricas, corroboradas pelos longos períodos de seca no nordeste, nos resultados de saúde de jovens nascidos no semiárido brasileiro na década de 90.

2.2 Estratégia empírica

A. Modelo Empírico

O painel é composto de municípios por mês segundo a região semiárida nordestina, para o período de 1996-2010, construído com as variáveis de interesse. Assim, a análise dos impactos das variações pluviométricas na mortalidade infantil durante o período de gestação segue a especificação do painel de municípios por mês de nascimento abordado em Rocha e Soares (2015):

$$Y_{itm} = \alpha + \beta R_{itm} + \phi_{im} + \lambda_t + \varphi t_{gtm} + \pi T_{itm} + \epsilon_{itm} \quad (1)$$

onde Y_{itm} é um resultado de saúde médio (por município) para as crianças nascidas no município i , no ano t e mês m . A variável dependente utilizada é a mortalidade infantil. R_{itm} é a variável de precipitação; ϕ_{it} é um efeito-fixo para município i e mês m (com $m = 1, 2, \dots, 12$); λ_t é um efeito-fixo ano; T_{itm} é a temperatura média no município i no mesmo período de 12 meses antes do nascimento; t_{gtm} é a tendência temporal linear e ϵ_{itm} é o termo de erro aleatório.

Cabe aqui destacar que a variável relacionada ao clima, R_{itm} , é análoga ao realizada em Rocha e Soares (2015), ou seja, são concebidas de mesma forma, embora difiram na natureza do problema, uma vez que a fonte de dados e sua tabulação são diferentes entre si. Esta variável é construída de duas maneiras, ambas para captar os efeitos das secas no semiárido através da flutuação das chuvas durante o período de gestação dos indivíduos:

- I. **Log desvio das chuvas nos últimos 12 meses** – conforme os autores, esta variável por ser interpretada como uma porcentagem de desvio da precipitação média. Definida por:

$$R_{it}^1 = \ln \left(\sum_{t=\tau-11}^{\tau} r_{im} \right) - \ln(\bar{r}_i) \quad (2)$$

onde r_{im} indica a precipitação mensal no município i e mês m , \bar{r}_i é a precipitação anual média histórica no município i , e τ indica o mês de nascimento de um indivíduo. Diferente ao realizado em Rocha e Soares (2015), o cálculo do $\sum_{t=\tau-11}^{\tau} r_{im}$ será realizado apenas para os municípios pertencentes ao semiárido, uma vez que contabilizar estados onde tem muita chuva, por exemplo, os estado do Norte, pode superestimar esta covariável.

- II. **Seca dos últimos 12 meses** – uma variável *dummy* que assume o valor 1 se houve seca no período e 0 se o contrário. Em síntese esta variável busca captar os eventos extremos, calculado por:

$$R_{it}^2 = 1 \text{ se } \sum_{t=-11}^{\tau} r_{im} < (\bar{r}_i - r_i^{SD}), \text{ e } 0 \text{ para outros} \quad (3)$$

onde r_i^{SD} é o histórico anual do desvio padrão de chuvas para o município i (calculado sobre o período de 1961-1990). Em outras palavras, o $R_{it}^2 = 1$ indica que a precipitação ao longo dos 12 meses anteriores ao nascimento de um indivíduo era mais do que um desvio padrão abaixo da média histórica para o município i .

Em suma, a equação (1) estima os efeitos das variações pluviométricas, aqui representadas nas equações (2) e (3), sobre a taxa de mortalidade infantil em dados em painel de município-a-mês de 1996 a 2010.

B. Métodos Empíricos

Para estimar a equação (1) é necessário considerar que os efeitos fixos por município-mês controlam variações climáticas em determinados períodos do ano assim como diferenças entre os municípios, enquanto que o efeito fixo por ano tenta captar choques agregados e tendências que impactam toda a região. Por sua vez, a tendência linear controla a dinâmica da evolução das mudanças climáticas e socioeconômicas da região, e a temperatura é responsável por outras variações climáticas que podem estar relacionadas com a quantidade de chuvas (Rocha e Soares, 2015).

Uma grande literatura sobre estimação de regressões quantílicas com dados em painel tem surgido nos últimos anos, tal como Koenker (2004) que introduziu uma abordagem geral para estimação com

dados longitudinais. Controlando para os efeitos fixos municipais e a heterogeneidade das covariáveis, a estimação por *quantis* é uma abordagem mais flexível para a análise de dados em painel do que os estimadores de efeitos fixos e aleatórios gaussianos clássicos. Em adição, considerar os efeitos de tratamento em modelos quantílicos de dados em painel constitui mais recente robustez das análises desta natureza. Não obstante, é necessário ir além e também captar outros efeitos na média. Portanto, neste arcabouço, apresentam-se os dois estimadores aqui utilizados.

I - Estimador assintoticamente consistente de Canay (2011) – (QRFE)

Canay (2011) apresenta um estimador consistente em que utiliza uma simples transformação de dados para eliminar os efeitos fixos dos *quantis* condicionais, conforme o modelo:

$$y_{it} = x'_{it}\beta + \alpha_i + \varepsilon_{it} \quad E(\varepsilon_{it}/X_{it}, \alpha_i) = 0 \quad (4)$$

Onde $(y_{it}, x_{it}) \in \mathbb{R} \times \mathbb{R}^k$ são variáveis observáveis, $\theta(U_{it}) = \beta$ e $(U_{it}, \alpha_{it}) \in \mathbb{R} \times \mathbb{R}$ são não observados. $U_{it}/x_{it}, \alpha_{it} \approx U[0,1]$. A função $\tau \rightarrow x'\theta(\tau)$ é assumida como sendo estritamente crescente em $\tau \in (0,1)$ e o parâmetro de interesse é assumido para ser $\theta(\tau)$. A principal restrição é que as heterogeneidades individuais associadas a x_{it} deve ter um efeito local puro.

Em suma, o método versa na questão em que as condições (U_{it}, α_{it}) do parâmetro $\theta(\tau)$ podem ser identificadas e consistentemente calculadas a partir dos dados. Para estimar a equação (7), Canay (2011) propõe um estimador simples de dois passos, sendo $E_T(.) \equiv T^{-1} \sum_{t=1}^T(.)$ e $E_{nT}(\cdot) \equiv (nT)^{-1} \sum_{t=1}^T \sum_{i=1}^n(\cdot)$, como segue:

Passo 1. Deixe $\hat{\theta}_\mu$ para \sqrt{nT} – estimador consistente de θ_μ . Estimam-se os efeitos fixos como: $\hat{\alpha}_i \equiv E_T[y_{it} - x'_{it}\hat{\theta}_\mu]$.

Passo 2. Sendo $\hat{y}_{it} \equiv y_{it} - \hat{\alpha}_i$ e definir o estimar do dois-estágios $\hat{\theta}(\tau)$ como:

$$\hat{\theta}(\tau) \equiv \underset{\theta \in (\cdot)}{\operatorname{argmin}} E_{nT}[\rho_r(\hat{y}_{it} - x'_{it}\theta)]$$

O método consiste na obtenção de *insights* sobre os efeitos dos regressores na variável dependente. Isto porque permite uma leitura dos efeitos heterogêneos através da transformação dos dados, uma vez que em modelos de regressão padrão, que seguem a média, a heterogeneidade muitas vezes não é captada.

II - Estimador de regressões quantílicas de dados em painel com variáveis instrumentais (QRVI)

Alternativamente, considerar-se-á para efeitos de resultados, o estimador proposto por Harding e Lamarche (2009), no qual apresentam uma abordagem de regressão quantílica para modelos de dados em painel com variáveis endógenas. Este método é mais robusto do que o método de Canay (2011), uma vez que potencializa as estimativas ao considerar na regressão efeitos além da heterogeneidade não observada. O método é uma extensão do procedimento de Chernozhukov e Hansen (2008) permitindo a inclusão de efeitos fixos introduzidos em Koenker (2004). Denote o vetor de variável endógena, d , na equação (4):

$$y_{it} = d'_{it}\delta + x'_{it}\beta + \alpha_i + u_{it} \quad (5)$$

$$d_{it} = h(x_{it}, w_{it}, v_{it}) \quad (6)$$

$$\alpha_i = g(x_{i1}, \dots, x_{iT}, d_{i1}, \dots, d_{iT}, \varepsilon_{it}) \quad (7)$$

onde a equação (6) define a variável endógena d relacionada a um vetor de instrumentos w que estão estocasticamente independente de u . Já a variável v é estocasticamente dependente de u . A equação (7) expressa a correlação entre as variáveis e os efeitos individuais. Neste caso, assume-se que ε é independente de u e v . Considere a função objetivo para o relacionamento quantílico instrumental condicional:

$$R(\tau, \delta, \beta, \gamma, \alpha, \mu) = \sum_{t=1}^T \sum_{i=1}^N \rho_{\tau}(y_{it} - d'_{it}\delta - x'_{it}\beta - \alpha_i - z'_{it}\alpha - \hat{w}'_{it}\gamma) \quad (8)$$

Onde \hat{w}_{it} é um vetor de instrumentos e $\rho_{\tau} = u(\tau - 1(u \leq 0))$ é uma função perda de regressão quantílica. Seguindo o procedimento de dois passos de Chernozhukov e Hansen (2006, 2008), temos:

Passo 1. Minimizar a função objetivo (8) para β, α, μ e γ em função de τ e δ :

$$\{\hat{\beta}(\tau, \delta), \hat{\gamma}(\tau, \delta), \hat{\alpha}(\tau, \delta)\} = \arg \min_{\beta, \gamma, \alpha} R(\tau, \delta, \beta, \gamma, \alpha) \quad (9)$$

Passo 2. Em seguida, estima-se o coeficiente da variável endógena para encontrar o valor de δ , no qual minimiza a função distância ponderada definida em y :

$$\hat{\delta}(\tau) = \arg \min_{\delta} \hat{\gamma}(\tau, \delta)' A \hat{\gamma}(\tau, \delta) \quad (10)$$

Onde A é uma matriz definida positiva. De acordo com Harding e Lamarche (2009) este estimador é assintoticamente normal podendo ser executado para mais de um quantil simultaneamente.

2.3 Descrição dos dados

A. Dados de Clima

O uso de variáveis relacionadas ao clima é empregado para suportar uma ampla gama de estudos de investigação sobre o impacto destas variáveis sobre os resultados de saúde dos indivíduos. Neste sentido, o uso de dados de precipitação e temperatura vem ganhando espaço da literatura, pois a variabilidade destas afetam diversas atividades econômicas (Freitas, Franchito e Rao, 2010) e costumam representar a variabilidade de sinais geofísicos e regimes de precipitação (Guedes et. al., 2004), sobretudo na literatura climatológica. O Quadro 1 apresenta algumas fontes e tipo de análise.

Na literatura meteorológica e climatológica, destacam-se três tipos de dados: i) dados observacionais - estações; ii) dados de reanálise e iii) dados de sensoriamento. O primeiro é oriundo de estações compostas por sensores isolados que registram continuamente os parâmetros meteorológicos que são lidos e anotados por um observador a cada intervalo. Contudo, apesar de reconhecer a fidedignidade deste tipo de observação, questões geográficas e econômicas tem dificultado a produção de séries temporárias de longo prazo (Pinto et. al., 2009).

Quadro 1 – Diversas Fontes e Metodologias de dados de Precipitação

Tipos de Dados	Fonte	Grade	Autores que utilizaram os dados
Observacionais	ANEEL	-	Freitas, Franchito e Rao, 2010;
	SUDENE	-	Menezes, et. al (2008)
	BDMEP	-	Rao et. al., 1996;
	DNOCS	-	Rao, Lima e Franchito (1993)
Reanálise	NCEP/NCAR	~1,9° x 1,9°	Guedes et. al., 2005; Pinto et. al., 2009; Freitas, Franchito e Rao, 2010;
	CPTEC	0,4° x 0,4°	Pinto et. al., 2009;
	CPC/NOOA		
	Terrestrial Air Temperature and Terrestrial Precipitation	0,5° x 0,5°	Rocha e Soares (2015)
Sensoriamento	GPCP	2,5° x 2,5°	Pinto et. al., 2009; Freitas, Franchito e Rao, 2010;
	CMAP	2,5° x 2,5°	Pinto et. al., 2009;
	TRMM	1,0° x 1,0°	Pinto et. al., 2009;

Fonte: Elaboração dos autores.

Por sua vez, dados de reanálise são muito convenientes para completar as lacunas onde não possuem informações, como por exemplo, onde não há estação meteorológica, ou mesmo, devido a problema de obtenção de dados. No entanto, as séries históricas são geradas a partir de modelos e são construídas com base em algum tipo de observação em superfície, inclusive de estações. Por fim, os dados de sensoriamento remoto combinam dados de superfície com estimativas feitas por satélite para calibrar o banco de dados.

É possível perceber que existe um *trade-off* no uso de variáveis clima. Se por um lado, os dados observacionais são fontes mais confiáveis, a ausência de informações e as quebras das séries de longo prazo tem dificultado as investigações climatológicas. Todavia, os dados de reanálise dependem de fontes primárias confiáveis, modelos bem ajustados e técnicas de obtenção robustas. Neste sentido, Silva, *et. al.* (2007), considerando as novas calibrações efetuadas no modelo de obtenção dos dados de reanálise da CPC/NCAR, realizado em 2005, constataram através dos diagramas de dispersão tendências para valores mais altos da CPC em relação aos dados observacionais. Não obstante, os autores ainda ressaltam, consistentemente aos estudos anteriores, que as comparações indicam menor número de dias secos, maior número de dias de baixa precipitação, e menos eventos extremos de precipitação por grade analisada, apesar da boa correlação existente entre as mesmas.

Outros estudos também visam comparar alguns produtos de precipitação. Por exemplo, Freitas, Franchito e Rao (2010) que compararam os dados ANEEL (observacionais); GPCP (sensoriamento) e NCEP/NCAR (reanálise) e concluíram que os dados de reanálise NCEP/NCAR superestimam a precipitação no Nordeste Brasileiro e que os dados observacionais ANEEL, apesar da ausência de dados e falhas reproduziram melhor a climatologia. Guedes *et. al.* (2005) em um estudo para Fortaleza-CE compararam dados NCEP/NCAR com os dados das estações e concluem que os dados observados e de reanálise são comparáveis, mas o ciclo anual da reanálise parece estar um mês adiantado em relação aos dados observados e que os desvios padrões mensais apresentam grandes diferenças. Tais resultados são corroborados em Pinto *et. al.* (2009) ao concluir que os dados de reanálise CPTEC não são capazes de representar os principais regimes de precipitação existentes no continente, com forte tendência a superestimar a intensidade média no interior do Nordeste Brasileiro.

Nesse arcabouço, recentemente Rocha e Soares (2015) utilizaram os dados *Terrestrial Precipitation: 1900-2010 Gridded Montly Times Series* (versão 3.02). Este conjunto de séries temporais é

procedente das estações do *Global Historical Climatology Network* (GHCN) e *Global Surface Summary of day* (GSOD) através de interpolação espacial. Portanto, trata-se de dados de reanálise e que de acordo com os estudos supracitados podem-se incorrer em superestimações de valores de precipitação e temperatura para o semiárido brasileiro. Para fins de comparação com os resultados recentes dos autores e em virtude da ausência de séries mais longas, neste artigo empregam-se os dados observacionais do Banco de Dados Meteorológicos para Ensino e Pesquisa (BDMEP), obtidos através do Instituto Nacional de Meteorologia (INMET).

A ideia é verificar se os impactos das restrições pluviométricas na época de nascimento do indivíduo são realmente relevantes nos resultados de saúde quando adultos em cenário com dados mais confiáveis de precipitação, embora isso acarrete na redução da amostra. Isto porque este banco não possui estações para todos os municípios pertencentes do semiárido brasileiro como também os inícios das séries históricas divergem entre elas.

Neste escopo foram criadas as duas variáveis expressas nas equações (1) e (2) utilizando os dados obtidos na BDMEP, primeiro localizando o município através da latitude e longitude; e depois, gerando os dados mensais de precipitação para cada município que compõe o semiárido brasileiro para o período de 1977 à 2010. No total, foram geradas 8.370 observações distribuídas nos 49 municípios analisados.

B. Variável Instrumental

Para empregar o método de Harding e Lamarche (2009), que seja um estimador de regressão quantílica de dados em painel com variáveis instrumentais necessitar-se-á de um vetor que denote uma variável que afeta o nível de precipitação, mas não tem efeito direto sobre os resultados de saúde dos indivíduos. Pesquisas tem mostrado que as condições oceânicas, sobretudo a Temperatura da Superfície do Mar (TSM), apresentam-se como a principal variável física a influenciar as condições climáticas.

Um fenômeno climático de interação oceano-atmosfera bastante investigado é o *El Niño* – Oscilação Sul (ENOS). Esta variabilidade climática é cíclica, oriunda do oceano Pacífico, com periodicidade que varia de 3 a 7 anos, duração igual ou superior a cinco meses consecutivos (Trenberth, 1997), e podem ser detectadas através das anomalias persistentes da TSM. Anos de *El Niño* (ou *El Niña*) estão normalmente associados à escassez (ou chuvas abundantes). A influência dos ENOS, sobretudo no Nordeste Brasileiro, tem sido objeto de estudo a destaca Kayano, Rao e Moura (1988); Rao, Lima e Franchito (1993); Xavier et. al. (2003); Alves e Souza (2003); Santos e Brito (2007), entre outros.

Outro fenômeno oceano-atmosfera é o dipolo do Atlântico. Destacam-se os estudos seminais de Hastenrath e Heller (1977), Moura e Slukla (1981), e mais recentemente, Aragão (1998), Moraes Neto e Araújo (2007), Andreoli e Kayano (2007), entre outros. Identifica-se o Dipolo quando as águas do Atlântico tropical estão mais quentes e do Atlântico Equatorial e Tropical Sul estão mais frias culminando para movimentos descendentes transportando o frio e seco inibindo a formação de nuvens e diminuindo a precipitação (Aragão, 1998).

Com base na literatura climatológica, optou-se pelos dados de temperatura dos Oceanos Pacíficos e Atlântico, conforme as pesquisas retromencionadas, uma vez que tanto os ENOS quanto o Dipolo Atlântico afetam a ocorrência de secas no semiárido brasileiro, mas não afetam diretamente a taxa de mortalidade infantil nem as condições de saúde ao nascer. Desta feita, os dados de TSM observados no Oceano Pacífico e Atlântico foram extraídos do *International Comprehensive Ocean–Atmosphere Dataset* (ICOADS) e *National Ocean and Atmospheric Administration* (NOAA), respectivamente.

O primeiro⁵ arquivo contém dados mensais em pontos de grade de 2° x 2° de latitude-longitude para um período de 1854 – até o presente composto por quatro arquivos a seguir: i) Extremo Oriente -

⁵ Ver detalhes são encontrados em Huang, et. al (2014) e Liu, et. al. (2014).

Niño 1 + 2 (0-10S; 90W-80W); ii) Leste Tropical - Niño 3 (5N-5S; 150W-90W); iii) Leste Central Tropical - Niño 3.4 (5N-5S; 170W-120W); e iv) Tropical Central – Niño 4 (5N-5S; 160E – 150W). Por sua vez, o segundo⁶ arquivo são dados com resolução 0,25° x 0,25° obtidos através de uma interpolação ótima composto de dois tipos: i) Atlântico Norte (5N-20N; 60W-30W); e ii) Atlântico Norte (0-20S; 30W-10E). A Tabela 1 apresenta os coeficientes de correlação de *Spearman* para os dados de precipitação e os dados de TSM.

Tabela 1 – Correlação de *Spearman* entre dados de Precipitação de TSM

DADOS PRECIPITAÇÃO/TSM	OCEANO PACÍFICO				OCEANO ATLÂNTICO	
	Niño 1 + 2	Niño 3	Niño 3.4	Niño 4	Atlântico Norte	Atlântico Sul
BMEP	0,229 (0,000)*	0,0047 (0,673)	0,239 (0,000)*	0,0098 (0,377)	-0,461 (0,000)*	0,469 (0,000)*
Rocha e Soares (2015)	0,211 (0,000)*	-0,0211 (0,056)	0,219 (0,000)*	-0,019 (0,089)	-0,480 (0,000)*	0,472 (0,000)*

Fonte Elaboração dos Autores

* Estatisticamente significativo a um nível de significância de 5%.

Para fins desta pesquisa, utilizar-se-á os dados de TSM com maior correlação com os dados de precipitação. Logo, observa-se na Tabela 1 que o Niño 3.4 correspondente ao leste central tropical do Oceano Pacífico apresenta maior correlação com as duas bases de dados de precipitação e sendo o motivo pelo qual utilizamos na pesquisa.

3 RESULTADOS EMPÍRICOS

Para este estudo, os dados foram coletados a partir de áreas geográficas dos quais as chuvas são bastante irregulares e o solo é raso. O semiárido nordestino brasileiro possui essas características, e embora exista ocorrência de chuvas, historicamente são permeados por longos períodos de secas, com características edafoclimáticas e socioeconômicas restritivas à população em razão de sua escassez. O saldo global dos efeitos sobre a saúde é provável que seja negativo, principalmente quando acometem indivíduos no período intrauterino ou nos primeiros anos de nascimento (Backer, 1998; Cunha e Heckman, 2007; Heckman, 2007), onde tendem a ser particularmente vulneráveis aos efeitos adversos.

Dentro desse contexto, a análise da relação entre as condições externas nos anos iniciais de vida, especificamente em ambientes com escassez de água, sobre os resultados de saúde dos adultos – aqui expressos pela taxa de mortalidade infantil, é objeto desta seção dos quais se consideram três importantes inovações avaliadas a seguir.

3.1 Heterogeneidade não-observada

Ainda que de difícil identificação por não serem diretamente observáveis ou mensuráveis, a heterogeneidade não podem ser negligenciada nas estimações. Isso se deve ao fato de que como as estimativas se tratam de efeito médio, portanto, é plausível admitir que estas não retratem em sua totalidade a realidade dos 1.048 municípios analisados. Além do mais, a literatura epigenética (Fiori, 2008), na condução dos estudos de interações gene-ambiente, ratifica a clara distinção entre os atributos adquiridos ao longo da vida e aqueles relacionados aos genes, muitas vezes negligenciados na literatura da economia da saúde.

⁶ Ver detalhes em Reynolds, et. al (2007).

Assim sendo, características não observáveis correlacionadas com a probabilidade de exposição a ambientes desfavoráveis poderiam ser responsáveis por uma parcela dos resultados de saúde. Outrora, condições inerentes aos indivíduos são particularmente associadas aos seus resultados de saúde independente do ambiente em que convive. Não obstante, é relevante considerar, também, que os indivíduos podem responder a choques externos como a seca de forma distinta entre eles. Ademais, questões como distribuição geográfica e sazonalidade dever-se-iam também ser consideradas.

Uma alternativa para contornar estes problemas é o método de regressão quantílica com dados em painel proposto por Canay (2011) – denominado aqui de estimador QR(FE), apresentados na Tabela 2. Para obter as estimativas, cabe aqui enfatizar que foram utilizadas as seguintes variáveis de controle: *i*) temperatura média dos últimos doze meses; *ii*) tendência temporal baseada na grade formada pelos municípios pertencentes ao quadrado 5° x 5° de latitude/longitude; *iii*) precipitação dos 13 a 24 meses antes do nascimento; *iv*) precipitação nos doze meses posteriores ao nascimento dos indivíduos. Não obstante, é relevante destacar que para cada painel construiu-se as variáveis de precipitação e seca, expressos nas equações (2) e (3) deste estudo e apresentados no Painel (A) e (B) das Tabelas, respectivamente.

Tabela 2 - Resultados das estimações com método de Canay (2011) do período de 1996-2010 da Região Semiárida do Nordeste

	FE			QR(FE)		
	(1)	(2)	(3)	(4)	(5)	(6)
Painel A – Precipitação						
Precipitação antes do nascimento	-3,25* (1,775)	-5,48*** (1,507)	-6,14*** (1,989)	1,09*** (0,129)	-2,08*** (0,134)	-3,02*** (0,141)
Precipitação 13-24 meses antes do nascimento			-1,49 (2,218)			-0,91*** (0,127)
Precipitação 1-12 meses após o nascimento			-2,74 (2,339)			-2,88*** (0,142)
Painel B – Seca						
Seca antes do nascimento	2,54 (1,801)	3,34*** (1,175)	3,73*** (1,227)	18,83*** (0,231)	2,67*** (0,147)	2,95*** (0,161)
Precipitação 13-24 meses antes do nascimento			-0,980 (2,154)			-1,15*** (0,125)
Precipitação 1-12 meses após o nascimento			-1,812 (2,176)			-2,69*** (0,132)
Observações	182.458	180.659	168.267	182.458	182.458	169.946
Número de Municípios	1.048	1.037	1.037	1.048	1.048	1.048

Fonte: Elaboração com base nas estimações.

Notas: Erros padrão robustos entre parênteses. ***p-valor < 0,01. ** p-valor < 0,05. * p-valor < 0,10. As colunas (7) e (10) não considera nas estimações a temperatura antes do nascimento.

A Tabela 2 apresenta os resultados utilizando o método de dados em painel com efeitos fixos (FE) expressos nas colunas (1), (2) e (3) e o estimador robusto de Canay (2011) de regressão quantílica com dados em painel (QRFE) apresentados nas colunas (4), (5) e (6), onde a variável dependente é a mortalidade infantil por cada 1000 nascimentos, mensurados por mês de nascimento em termos de nível municipal. O Painel A interpreta a precipitação como uma porcentagem da precipitação média, enquanto que o Painel B como uma *dummy* para seca.

Vislumbrando apenas o Painel A tem-se na Coluna 1 que, em termos de magnitude, isto representa que o aumento no nível de precipitação no período gestacional reduz em cerca de 3% a taxa de mortalidade infantil. Tais valores aumentam quando controlado pela temperatura média durante a gravidez (-5,48%) e nível de precipitação antes da concepção e depois do nascimento (-6,14%), conforme a Coluna (2) e (3), respectivamente.

Posteriormente, ao comparar-se os estimadores FE(RS) e QR(FE), cabem aqui alguns apontamentos. Primeiro, verifica-se na coluna (3) – FE(RS) - que os controles utilizados não foram estatisticamente significativos corroborando para a conclusão de Rocha e Soares (2015) de que níveis de chuva em outros períodos que não seja o gestacional não afetam a mortalidade infantil. O segundo, diz respeito a implicação do primeiro, uma vez que afirmar que as variáveis de controle não exercem influência sob a mortalidade infantil seria concluir que nutrição materna no ano antes da concepção, bem como a nutrição e doenças no primeiro ano de vida não são afetados pelas chuvas. Tal conclusão é justamente oposta a literatura vigente (Glewwe e King, 2001; Cunha e Heckman, 2007, Kudamatsu et al., 2010).

Terceiro, tais evidências são refutadas com o estimador QR(FE), mostrando que as variáveis de controle são estatisticamente significativas e negativas e a principal implicação é a redução da magnitude da variável resposta. Além do mais, verifica-se através da coluna (6) que o efeito da precipitação de 1-12 meses após o nascimento (-2,88) é maior do que o efeito da precipitação de 13-24 meses antes do nascimento (-0,914), corroborando para as evidências da literatura de que os primeiros anos de vida dos indivíduos são períodos sensíveis⁷ e respondem mais precisamente a choques adversos (Cunha e Heckman, 2007; Heckman, 2007). Por exemplo, privação nutricional no início da vida decorrentes das secas pode retrair indicadores antropométricos persistindo até a vida adulta (Aguilar e Vicarelli, 2011).

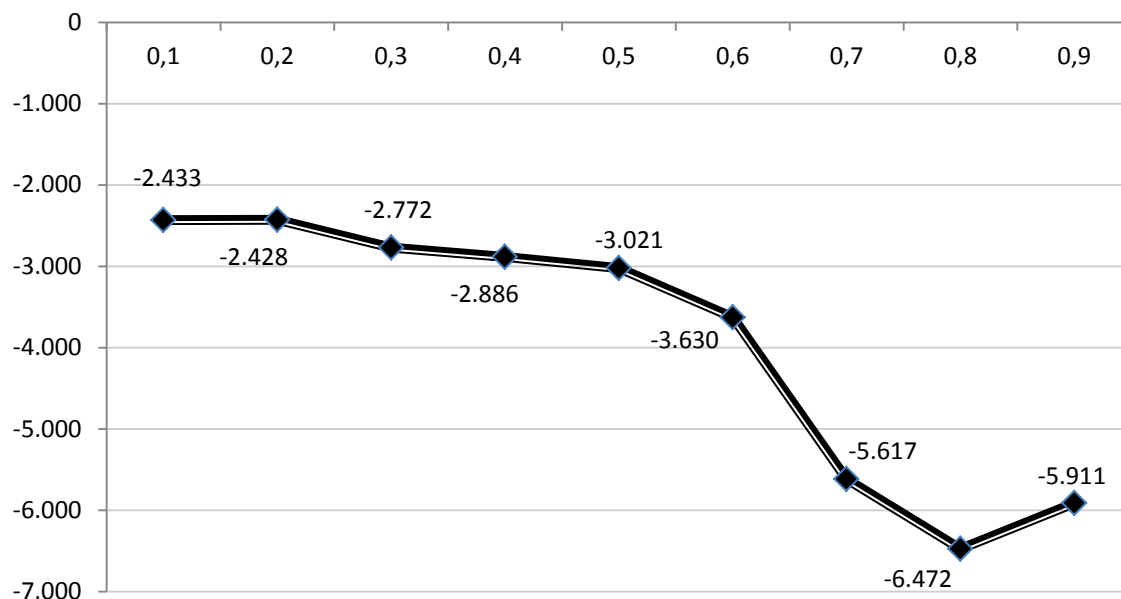
Quarto, e não menos importante, é notório que a falta de rigor metodológico imprime um papel demasiadamente relevante para o clima nos resultados de Rocha e Soares (2015). Notem que as colunas (1), (2) e (3) são, em números, quase o dobro das colunas (4), (5) e (6). Em síntese, tem-se um efeito superestimado da precipitação, o que nos leva a concluir *a priori* que existe condição endógena entre as variáveis relacionadas à precipitação que afetam resultados de saúde dos indivíduos, muito embora seja consensual de que esse efeito é negativo.

Sob outra ótica, observam-se resultados qualitativamente semelhantes para o Painel B que apresentam estimativas considerando uma *dummy* de seca. Por conseguinte, os resultados corroboram-se com os resultados do Painel A uma vez que mostra que as secas tem impacto positivo sob a taxa de mortalidade infantil dos quais tentam captar a gravidade dos choques destas. No geral nossos resultados confirmam os de Rocha e Soares (2015) de que independente de como se mensura a variável precipitação, são estatisticamente significativas e negativas. No entanto, considerando a heterogeneidade não observada verificou-se que as estimativas dos autores estavam superestimadas.

Outro ponto importante a destacar é que o método de Canay (2015) nos permite ir além das estimativas condicionais a média e conhecer um pouco da distribuição condicional através dos *quantis*. O Gráfico 1 apresenta as estimativas controladas pela temperatura e precipitação nos 12 meses antes 24 meses após do nascimento ao longo da distribuição, no qual ressalta-se que ambas apresentaram-se estatisticamente significativas.

⁷ De acordo com Cunha e Heckman (2007) e Heckman (2007) mostram que etapas importantes no desenvolvimento físico dos indivíduos pode ocorrer na infância e ter consequências para os resultados de saúde quando adultos. São os chamados períodos sensíveis que segundo os autores decorrem da primeira infância que vai até os primeiro quatro anos de idade.

Grafico 1 – Estimativas do impacto das secas sob os resultados dos indivíduos ao longo da distribuição



Fonte: Elaboração Própria

Nota: Todas os estimadores são estatisticamente significativos.

Verifica-se que o impacto da porcentagem da precipitação média cresce negativamente na medida em que cresce a taxa de mortalidade infantil. Em outras palavras, é possível perceber que o efeito é negativo, com tendência a aumentar na medida em que avançamos através dos *quantis*. É interessante ressaltar que sob o ponto de vista de políticas públicas, esta evidência tem papel importante ao identificar que choques externos tem efeito de longo prazo e com intensidade diferenciada (Maccini, 2009).

As evidências também implicam em dizer que além dos resultados diretos em saúde, poder-se-ia também alcançar efeitos indiretos em educação comprometendo as habilidades cognitivas dos indivíduos (Heckman, 2007). Embora as funções cognitivas possam ser afetadas negativamente por herança epigenética, a baixa cognição pode ocorrer, também, pela ausência de condições favoráveis durante a gravidez e na infância (Currie, 2006). Por exemplo, a escassez de água pode reduzir a produção agrícola e a ingestão suficiente de nutrientes afetando os resultados de saúde, e a falta de água potável pode gerar maior incidência de doenças infecciosas comprometendo a qualidade de vida da região (Rocha e Soares, 2015).

Com isso, políticas voltadas a atenuar os efeitos da seca na vida intrauterina tendem a melhorar o bem-estar dos indivíduos no longo prazo, e, quanto maior o *quantil* pertence o município, melhor seria o efeito desta política pública tanto em termos de saúde quando em educação no futuro dos indivíduos. Esta conclusão abre espaço para uma agenda de pesquisa em economia da saúde e da educação que foge ao escopo deste artigo.

3.2 Endogeneidade

Com o intuito de ampliar a análise esta subseção baseia-se no método proposto por Harding e Lamarche (2009) visando endogeneizar a quantidade de precipitações através de um modelo empírico de regressão quantílica de dados em painel com variáveis instrumentais, ora aqui chamado de estimador QR(IV). Os resultados encontram-se na Tabela 3. A inserção de variáveis instrumentais dado que estas não pertencem à equação explicativa, mas está correlacionada com a variável de precipitação é aqui justificada pelos fenômenos oceano-atmosfera ENOS (Rao, Lima e Franchito, 1993) e do Dipolo Atlântico (Aragão, 1998), expressos na variável Temperatura da Superfície do Mar (TSM).

Tabela 3 - Resultados das estimações com método de Harding e Lamarche (2009) do período de 1996-2010 da Região Semiárida do Nordeste

	QR(FE)			QR(VI)		
	(7)	(8)	(9)	(10)	(11)	(12)
Painel A – Precipitação						
Precipitação antes do nascimento	1,09*** (0,129)	-2,08*** (0,134)	-3,02*** (0,141)	0,65*** (0,134)	-2,49*** (0,132)	-2,63*** (0,133)
Precipitação 13-24 meses antes do nascimento			-0,91*** (0,127)			-1,94*** (0,126)
Precipitação 1-12 meses após o nascimento			-2,88*** (0,142)			-1,98*** (0,131)
Painel B – Seca						
Seca antes do nascimento	18,83*** (0,231)	2,67*** (0,147)	2,95*** (0,161)	18,81*** (0,2358)	2,83*** (0,152)	2,86*** (0,156)
Precipitação 13-24 meses antes do nascimento			-1,15*** (0,125)			-2,03*** (0,127)
Precipitação 1-12 meses após o nascimento			-2,69*** (0,132)			-1,79*** (0,120)
Observações	182.458	182.458	169.946	182.458	182.458	169.946
Número de Municípios	1.048	1.048	1.048	1.048	1.048	1.048

Fonte: Elaboração dos autores com base nas estimações.

Nota 1: Erros padrão robustos entre parênteses. ***p-valor < 0,01. ** p-valor < 0,05. * p-valor < 0,10.

Nota 2: Os resultados por quantis encontram-se no anexo I.

São transportados para a Tabela 3, colunas (7), (8) e (9) os resultados já apresentados nas colunas (4), (5) e (6) da Tabela 2 no que concernem as estimativas obtidas através de Canay (2011). Por sua vez, as colunas (10), (11) e (12) mostram os resultados para o estimador EQD(IV), que intuitivamente tem as mesmas conclusões do estimador QR(FE), exceto pelas magnitudes dos parâmetros que diminuem em absoluto. Enquanto que Rocha e Soares (2015) concluíram que aumentos de precipitação nos primeiros anos de vida pode reduzir em torno de 6% a taxa de mortalidade infantil, o estimador EQD(IV) demonstra que seu poder redutor resume-se em apenas 2,6%. Tais evidências são encontradas também por *quantil*, conforme Anexo I, no qual as variáveis instrumentais foram significativas.

Sob outro prisma, o Painel B apresenta uma variante do nível de precipitação. Denota-se a variável seca uma *dummy* que quando assumido valor 1 significa que a precipitação ao longo dos 12 meses anteriores ao nascimento era mais do que um desvio padrão abaixo da média histórica. Com esta medida a interpretação se inverte, espera-se sinal positivo desta relação, conforme se verifica nos resultados dos estimadores aqui utilizados. Contudo, análogo à análise anterior, constata-se que os resultados de Rocha e Soares (2015) também foram superestimados provavelmente pelos problemas de endogeneidade não contabilizados.

Em suma, apesar dos resultados entre os estimadores serem semelhante qualitativamente, o que corrobora para a aplicação de políticas públicas voltadas a amenizar os efeitos das secas, é relevante destacar que as evidências correspondentes apontam um trato demasiadamente elevado ao *status* clima nos resultados de saúde dos indivíduos. Ademais, suspeita-se que esse *quórum* superestimado dos parâmetros possa, além do método inadequado, também ser decorrente do uso inapropriado de dados de precipitação, conforme analisado na próxima subseção.

3.3 Dados Observacionais

É consensual que dados históricos de chuvas e temperatura são boas medidas para verificar seu impacto sobre os resultados de saúde dos indivíduos, principalmente em regiões como a semiárida brasileira que é acometida por longos períodos de secas, e, que, conseqüentemente deve gerar ônus para sua população residente. Uma das justificativas para seu uso é que a variabilidade destas afetam diversas atividades econômicas (Freitas, Franchito e Rao, 2010), o que potencializa seu impacto, principalmente em regiões subdesenvolvidas e dependentes da agricultura (Cirilo, 2008).

Um dos desafios centrais da literatura ao analisar a relação entre a precipitação no início da vida e os próprios resultados de saúde dos indivíduos é que os dados de precipitação utilizados são suscetíveis a erro. Mas como isso ocorre? Isto porque o nível de precipitação é obtido através de estações meteorológicas que em virtude das dificuldades geográficas e do custo de mantê-las, muitas vezes não abrangem um contingente territorial representativo. Além disso, mesmo em locais onde existam estações meteorológicas, pelas mais diversas razões, pode-se gerar problemas que as tornem cronologicamente descontínua (Pinto et. al., 2009). Para contornar este problema, surgem na literatura às chamadas séries de reanálise baseados em modelos matemáticos, como por exemplo, interpolação linear, entre outros. No entanto, a contra gota é que essa medida também contém erros que podem superestimar as variabilidades climatológicas estudadas.

Os dados de reanálise são modelados com a média entre dois pontos de grade distanciados entre si em diferentes resoluções espaciais baseados nas distâncias de latitude e longitude (equivalente a km). Quanto maior a resolução, maior a distância entre os pontos. No entanto, as diferenças entre Observacionais e Reanálise poderiam decorrer das diferentes representatividades dos dados utilizados, ora aqui questionado. Para consolidar as conclusões aqui expostas, utilizou-se os dados de reanálise a *Terrestrial Air Temperature and Terrestrial Precipitation: 1900–2010 Gridded Monthly Time Series*, versões 3.01 e 3.02, respectivamente (Matsuura and Willmott, 2009) utilizados em Rocha e Soares (2015), considerando apenas os municípios do semiárido em que tem informações na base de dados observacional.

No que tange a escolha dos dados observacionais, optou-se pelos dados obtidos no BDMEP, para testar se erro de medição clássica na variável precipitação no início da vida leva a estimativa dos coeficientes atenuados. Muito embora a escolha da base implique em redução da amostra estudada, não há embargo uma vez que estamos comparando mesmos municípios. Portanto, resume-se em 47 municípios analisados em uma amostra de 8.070 observações. No que se refere a questão metodológica, optou-se pelo método de Harding e Lamarche (2009) uma vez já comprovados nas seções anteriores como mais adequado para estimar a relação aqui investigada em virtude deste captar a heterogeneidade não observada e tratar a endogeneidade existente.

A sinopse dos resultados é apresentada na Tabela 4, dos quais as colunas (13), (14) e (15) com estimativas dos dados de reanálise, enquanto que as colunas (16), (17) e (18) expõe os dados observacionais. Para ambos os resultados, considerou-se a Temperatura do Oceano Atlântico (TSM) como variável instrumental. Observa-se que a maioria dos resultados foram estatisticamente significativos e negativos, ratificando os resultados anteriores, contudo, merece atenção alguns fatos.

A primeira conclusão, *a priori*, é que de fato variáveis oriundas de reanálise tendem a superestimar os parâmetros desta relação. Notem que para cada coluna o parâmetro estimado com dados de reanálise são maiores do que aqueles estimados com dados de precipitação observacionais. Com isso, vale aqui destacar duas importantes reflexões. Primeiro, chama-se atenção para a Tabela 4, coluna (15), que o controle precipitação de 13-24 meses antes do nascimento não é estatisticamente significativo nos dois estimadores (variável precipitação e seca), dando apenas importância para a precipitação após a

concepção dos indivíduos. Parece que reduzir a amostra potencializa erro de medição da variável comprometendo os resultados.

Tabela 4 - Resultados das estimações de comparação de dados de Reanálise e Observacionais do período de 1996-2010 da Região Semiárida do Nordeste

	Dados de Reanálise			Dados Observacionais		
	(13)	(14)	(15)	(16)	(17)	(18)
Painel A – Precipitação						
Precipitação antes do nascimento	-5,79*** (0,897)	-4,60*** (0,958)	-4,26*** (1,020)	-4,42*** (0,924)	4,56*** (0,972)	-3,80*** (1,040)
Precipitação 13-24 meses antes do nascimento			0,27 (0,956)			-2,83** (1,124)
Precipitação 1-12 meses após o nascimento			-2,52*** (0,944)			-7,72*** (1,239)
Painel B – Seca						
Seca antes do nascimento	20,44*** (1,361)	4,36*** (1,194)	5,52*** (1,683)	10,72*** (1,477)	4,35*** (1,454)	4,09*** (1,385)
Precipitação 13-24 meses antes do nascimento			0,34 (0,981)			-2,73** (1,126)
Precipitação 1-12 meses após o nascimento			-1,77* (0,950)			-7,52*** (1,178)
Observações	8.070	8.070	7.541	8.070	8.070	7.541
Número de Municípios	47	47	47	47	47	47

Fonte: Elaboração dos autores com base nas estimações.

Notas: Erros padrão robustos entre parênteses. ***p-valor < 0,01. ** p-valor < 0,05. * p-valor < 0,10. As colunas (13) e (16) não considera a temperatura antes do nascimento nas estimações.

A segunda reflexão se concentra na coluna (17), em que apresenta coeficiente positivo. Quando não controlado pelas variáveis de precipitação antes e depois do nascimento, o resultado é contraditório. Seria plausível pensar que como estamos em um número reduzido de municípios, talvez para os municípios pertencentes a amostra, esse resultado fosse amplamente aceito, no entanto, a coluna (18) mostra ser conivente com a literatura e com os demais resultados aqui explorados de que aumento de chuvas no período de nascimento tendem a reduzir a taxa de mortalidade infantil.

Com vistas a esses resultados, frisa-se que os efeitos de tratamento desse estudo ratifica o efeito negativo entre precipitação e mortalidade infantil, contudo, em magnitude inferior ao apresentado em Rocha e Soares (2015).

4 CONSIDERAÇÕES FINAIS

O objetivo do trabalho foi avaliar o impacto dos resultados pluviométricos e de períodos de seca na taxa de mortalidade infantil no semiárido nordestino, a fim de ratificar ou não os resultados encontrados na literatura. Para tanto, utilizou-se três abordagens: *i*) método de regressão quantílica com dados em painel de efeitos fixos relatado em Canay (2011); *ii*) método de regressão quantílica com dados em painel com variáveis instrumentais proporcionado em Harding e Lamarche (2009) utilizando a Temperatura do Oceano como variável instrumental; *iii*) dados observacionais para contrapor os de reanálise.

Os resultados indicam um efeito negativo da precipitação e um efeito positivo da seca na taxa de mortalidade infantil, para as três estratégias. No entanto, a ausência de tratamento no que concerne considerar a heterogeneidade não observada e a endogeneidade, bem como o uso de dados modelados,

revelam possíveis superestimações. Dessa forma, os efeitos de tais fenômenos naturais na redução da mortalidade infantil existem, mas o seu impacto pode ser menor do que o comumente aceito, sendo a maior parte do impacto devido a outros fatores, tais como as condições de saúde dos municípios, entre outros.

Logo, as principais conclusões são: *i)* quanto maior a precipitação antes da concepção, durante o período gestacional e nos primeiros anos de vida dos indivíduos nascidos no semiárido menor a taxa de mortalidade infantil nesta região; *ii)* quando controlado a heterogeneidade não observada e a endogeneidade existente nesta relação, a precipitação tem menor peso indicando que existem outras variáveis que afetam os resultados de saúde; *iii)* o método utilizado em Rocha e Soares (2015) não é adequado para estimar essa relação. Sem embargo, o Painel B, que considera a variável seca, apresentam as mesmas conclusões sob ângulo oposto.

Em síntese, temos que bebês nascidos nas zonas áridas que sofrem secas quando no útero de suas mães residentes no semiárido enfrentam risco maior de morte, especialmente se nasceram quando as secas comprometem os seus primeiros anos de vida e quando a escassez de água afeta a nutrição materna antes da concepção desses indivíduos. Trabalhos como o de Maccini e Yang (2009), Kudamtsu et. al (2010); Aguilar e Vicarelli(2011), Shah e Steberg (2013), Burgess et. Al (2014) e Rocha e Soares (2015) tem ressaltado os efeitos das questões climáticas nos resultados de saúde, contudo, este estudo vem a contribuir ao ressaltar a importância de escolher métodos adequados que tratem os problemas existentes, bem como a escolha de dados confiáveis e que sejam fidedignos da relação causal a ser estudada.

5 REFERÊNCIAS

ALVES, José M. Brabo; SOUZA, R. O.; CAMPOS, José Nilson B. Previsão da anomalia de temperatura da superfície do mar (tsm) no atlântico tropical, com a equação da difusão de temperatura. **Revista Climanalise**, ano, v. 3, p. 163-172, 2003.

ANDREOLI, Rita Valéria; KAYANO, Mary Toshie. A importância relativa do atlântico tropical sul e pacífico leste na variabilidade de precipitação do Nordeste do Brasil. **Revista Brasileira de Meteorologia**, v. 22, n. 1, p. 63-74, 2007.

AGUILAR, Arturo; VICARELLI, Marta. El Nino and Mexican children: medium-term effects of early-life weather shocks on cognitive and health outcomes. **Cambridge, United States: Harvard University, Department of Economics. Manuscript**, 2011.

ARAGÃO, José Oribe Rocha. O Impacto do ENSO e do Dipolo do Atlântico no Nordeste do Brasil. **Bull. Inst. fr. études andines**, v. 27, n. 3, p. 839-844, 1998.

ARAÚJO COSTA, José. O FENÔMENO EL NIÑO E AS SECAS NO NORDESTE DO BRASIL. **EDUCTE: Revista Científica do Instituto Federal de Alagoas**, v. 3, n. 1, 2014.

BARKER, D. J. P. **Mothers, Babies and Health in Later Life**. Editora Churchill Livingstone, edição nº 2, ISBN-10: 0443061653, 1998.

BRAHMANANDA RAO, V.; DE LIMA, Marley C.; FRANCHITO, S. H. Seasonal and interannual variations of rainfall over eastern northeast Brazil. **Journal of Climate**, v. 6, n. 9, p. 1754-1763, 1993.

BEZABIH, Mintewab; DI FALCO, Salvatore; MEKONNEN, Alemu. **On the Impact of Weather Variability and Climate Change on Agriculture: Evidence from Ethiopia**. 2014.

BURGESS, Robin et al. The unequal effects of weather and climate change: Evidence from mortality in india. **Unpublished working paper**, 2013.

CANAY, Ivan A. A simple approach to quantile regression for panel data. **The Econometrics Journal**, v. 14, n. 3, p. 368-386, 2011.

CIRILO, José Almir. Public water resources policy for the semi-arid region. **estudos avançados**, v. 22, n. 63, p. 61-82, 2008.

CUNHA, Flavio; HECKMAN, James. **The technology of skill formation**. National Bureau of Economic Research, 2007.

CURRIE, Janet. **Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development**. National Bureau of Economic Research, 2008.

CURRIE, Janet et al. Something in the water: Contaminated drinking water and infant health. **Canadian Journal of Economics/Revue canadienne d'économie**, v. 46, n. 3, p. 791-810, 2013.

CURRIE, Janet; VOGL, Tom. **Early-life health and adult circumstance in developing countries**. National Bureau of Economic Research, 2012.

CURRIE, Janet; ROSSIN-SLATER, Maya. Weathering the storm: Hurricanes and birth outcomes. **Journal of health economics**, v. 32, n. 3, p. 487-503, 2013.

FIORI, N. As neurociências cognitivas. Tradução de Sonia M. S. Fuhrmann. Petrópolis RJ:Vozes, 2008.

FREITAS, Ana Carolina Vasques; FRANCHITO, Sérgio Henrique; RAO, Vadlamudi Brahmananda. Analise dos dados de precipitacao provenientes de diferentes fontes sobre a américa do sul, com enfase no brasil. **CLIMEP-Climatologia e Estudos da Paisagem, Rios Claros, PR, Brasil**, v. 5, n. 1, 2010.

GALVAO, Antonio F.; LAMARCHE, Carlos; LIMA, Luiz Renato. Estimation of censored quantile regression for panel data with fixed effects. **Journal of the American Statistical Association**, v. 108, n. 503, p. 1075-1089, 2013.

GLEWWE, Paul; KING, Elizabeth M. The impact of early childhood nutritional status on cognitive development: Does the timing of malnutrition matter?. **The world bank economic review**, v. 15, n. 1, p. 81-113, 2001.

GLUCKMAN, Peter. e HANSON, Mark. **The fetal Matrix: Evolution, Development, and Disease**. Cambridge: Cambridge University Press, 2005.

GROSSMAN, Michael. The demand for health: a theoretical and empirical investigation. **NBER Books**, 1972.

GUEDES, R et al. Série temporal de precipitação mensal de Fortaleza, Brasil: Comparação entre observações e dados de reanálise do NCEP/NCAR. **Revista Brasileira de Meteorologia**, v. 20, p. 83-92, 2005.

HARDING, Matthew; LAMARCHE, Carlos. A quantile regression approach for estimating panel data models using instrumental variables. **Economics Letters**, v. 104, n. 3, p. 133-135, 2009.

HASTENRATH, S.; HELLER, L. Dynamics of climatic hazards in north-east Brazil. *Quartely Journal of the Royal Meteorological Society*, v. 110, p. 411-425, 1977.

HECKMAN, James. The technology and neuroscience of skill formation. **Invest in Kids Working Group, Center for Economic Development, Partnership for America's Economic Success, July**, v. 17, 2007.

KAYANO, Mary Toshie; RAO, Vadlamudi Brahmananda; MOURA, Antonio Divino. Tropical circulations and the associated rainfall anomalies during two contrasting years. **Journal of Climatology**, v. 8, n. 5, p. 477-488, 1988.

KOENKER, Roger. Quantile regression for longitudinal data. **Journal of Multivariate Analysis**, v. 91, n. 1, p. 74-89, 2004.

KUDAMATSU, Masayuki; PERSSON, Torsten; STRÖMBERG, David. Weather and Infant Mortality in Africa. 2012. Manuscrito não publicado.

MACCINI, Sharon L.; YANG, Dean. **Under the weather: Health, schooling, and economic consequences of early-life rainfall**. National Bureau of Economic Research, 2008.

MORAES NETO, João M.; BARBOSA, Marx P.; DE ARAÚJO, Alexandre E. Efeito dos eventos ENOS e das TSM na variação pluviométrica do semi-árido paraibano. **R. Bras. Eng. Agríc. Ambiental**, v. 11, n. 1, p. 61-66, 2007.

MOURA, A. D.; SHUKLA, J. On the dynamics of the droughts in Northeast Brazil: observations, theory and numerical experiments with a general circulation model. *Journal Atmospheric Science*, v. 38, n. 12, p. 2653-2673, 1981.

PINTO, L. et al. Comparação de produtos de precipitação para a América do sul. **Revista Brasileira de Meteorologia**, v. 24, n. 4, p. 461-472, 2009.

ROCHA, Rudi; SOARES, Rodrigo R. Water scarcity and birth outcomes in the Brazilian semiarid. **Journal of Development Economics**, v. 112, p. 72-91, 2015.

SANTOS, C.; BRITO, J. Análise dos índices de extremos para o semi-árido do Brasil e suas relações com TSM e IVDN. **Revista Brasileira de Meteorologia**, v. 22, n. 3, p. 303-312, 2007.

SHAH, Manisha; STEINBERG, Bryce Millett. Could droughts improve human capital? evidence from india. **Unpublished manuscript, last modified**, 2012.

SILVA, Viviane BS et al. An improved gridded historical daily precipitation analysis for Brazil. **Journal of Hydrometeorology**, v. 8, n. 4, p. 847-861, 2007.

TRENBERTH, K. E. The definition of El Niño. **Bulletin of the American Meteorological Society**, v.78, n.12, p.2771-2777. 1997.

UGAZ, Jorge; ZANOLINI, Arianna. Effects of extreme weather shocks during pregnancy and early life on later health outcomes: the case of philippines? typhoons. 2011.

XAVIER, T. M. B. S.; XAVIER, A. F. S.; SILVA-DIAS, M. A. F.; SILVA-DIAS, P. L. Interrelações entre Eventos ENOS (ENSO), a ZCIT (ITCZ) no Atlântico e a Chuva nas Bacias Hidrográficas do Ceará, **Revista Brasileira Recursos Hídricos**, v. 8, n. 2, p. 111-126. 2003.

Anexo I – Resultados por *quantil* do Estimador QRD(FE)

<i>Quantil</i>	<i>Coefficiente</i>	<i>IV atlântico</i>	<i>IV pacífico</i>
0.1	-1,521*** (0,134)	-0,315*** (0,033)	0,212*** (0,028)
0.2	-1,729*** (0,124)	-0,316*** (0,030)	0,210*** (0,026)
0.3	-2,279*** (0,108)	-0,355*** (0,029)	0,214*** (0,024)
0.4	-2,405*** (0,110)	-0,368*** (0,029)	0,213*** (0,023)
0.5	-2,627*** (0,133)	-0,414*** (0,033)	0,220*** (0,026)
0.6	-3,457*** (0,172)	-0,455*** (0,042)	0,180*** (0,034)
0.7	-6,310*** (0,311)	-0,648*** (0,082)	0,075 (0,071)
0.8	-8,143*** (0,736)	-1,604*** (0,219)	-0,486*** (0,166)
0.9	-4,965*** (1,531)	-3,186*** (0,487)	-1,907*** (0,352)

Fonte: Elaboração dos autores.