

O DIFERENCIAL URBANO-RURAL NA FERTILIDADE TOTAL DAS MULHERES BRASILEIRAS[†]

Mateus Pereira Lavorato
DER/UFV
mateus.lavorato@ufv.br
João Eustáquio de Lima
DER/UFV
jelima@ufv.br
Marcelo José Braga
DER/UFV
mjbraga@ufv.br

Resumo: Neste artigo se buscou analisar a diferença no nível de fertilidade total das mulheres brasileiras domiciliadas nos meios urbano e rural. Duas perspectivas teóricas deste diferencial foram empiricamente analisadas, quais sejam a hipótese composicional e a hipótese contextual. Para tanto, primeiro foi modelada a fertilidade total das mulheres brasileiras, utilizando-se um modelo hurdle loglog complementar binominal negativo truncado em zero. Posteriormente, o diferencial urbano-rural da fertilidade total foi decomposto através de uma adaptação do modelo de Blinder-Oaxaca para funções não-lineares. Os principais resultados apontaram que, em geral, os fatores socioeconômicos analisados afetam de modo heterogêneo a fertilidade total das mulheres dos meios urbano e rural. Além disso, a decomposição do diferencial de fertilidade forneceu subsídios empíricos a favor das duas hipóteses consideradas, evidenciando que tanto as diferenças socioeconômicas quanto as especificidades dos meios de vida urbano e rural são capazes de explicar o distinto padrão de fertilidade total observado.

Palavras-Chave: fertilidade total; diferencial; urbano; rural; Brasil

Abstract: We analyzed the perceived difference on the completed fertility level of urban and rural Brazilian women. Two distinct theoretical perspectives of this differential were empirically analyzed: compositional and contextual hypothesis. On this purpose, we firstly modeled the completed fertility of Brazilian women through a negative binomial complementary loglog hurdle model. After this, we decomposed the completed fertility urban-rural differential through an adaptation of the Blinder-Oaxaca model for non-linear functions. Main results pointed out that, in general, the analyzed socioeconomic factors affect urban and rural completed fertility on a heterogeneous way. Furthermore, the decomposition provided empirical subsidies in favor of the two hypotheses considered, showing that both the socioeconomic differences and the specificities of urban and rural livelihoods are able to explain the distinct pattern of total fertility between urban and rural women.

Keywords: completed fertility; differential; urban; rural; Brazil.

Código JEL: J13; O18

[†] O(s) autor(es) agradecem o apoio financeiro da Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG).

1. Introdução

Há muito se estuda as diferenças socioeconômicas e demográficas existentes entre as populações residentes nos meios urbano e rural. Uma das pautas mais marcantes dessa agenda de pesquisa é aquela que versa sobre os diferentes padrões de fecundidade existentes entre estes dois grupos populacionais. Historicamente e, como comprovado empiricamente, não por coincidência, as mulheres do meio rural possuem maiores níveis de fecundidade e piores indicadores socioeconômicos que os de suas contrapartes domiciliadas no meio urbano (e.g., ROBINSON, 1961; MIRO; MERTENS; DAVIS, 1968; TROVATO; GRINDSTAFF, 1980).

Como destacado por Berquó e Cavenaghi (2006), o nível de fecundidade do país se apresenta como um indicador capaz de orientar os *policy makers* na formulação de políticas públicas, tendo em vista as implicações relacionadas a taxas de fecundidades destoantes do nível de reposição (2,1 filhos por mulher): valores baixos podem levar à redução do volume da população e seu envelhecimento, enquanto padrões de alta fecundidade podem evidenciar a existência de serviços de saúde reprodutiva precários e um processo falho de transmissão de informações à população. No caso brasileiro, a taxa de fecundidade vêm se reduzindo, sistematicamente, desde a década de 1960, chegando em 2010 ao patamar médio de 1,9 filho por mulher, valor inferior ao nível de reposição (BERQUÓ; CAVENAGHI, 2014).

Berquó e Cavenaghi (2005) apontam que, apesar da expressiva queda geral da taxa de fecundidade no Brasil, a redução não foi uniforme em diversos segmentos socioeconômicos da população. Como demonstrado pelas autoras, a queda mais acentuada, entre os anos de 1991 e 2000, foi observada em grupos com características socioeconômicas específicas: mulheres mais pobres, negras, com menor nível educacional, domiciliadas no meio rural e residentes nas regiões Norte e Nordeste. Além disso, também é destacado que, neste mesmo período, a queda de aproximadamente 21% ocorrida na fecundidade rural foi a grande responsável pela redução verificada no Brasil de modo geral, tendo em vista que para o meio urbano o declínio foi de apenas pouco mais de 4%.

Tendo em vista os diferenciais socioeconômicos existentes entre as populações dos meios urbano e rural no Brasil e, principalmente, a divergência existente entre os regimes de fecundidade observados para as mulheres destes dois grupos, a presente pesquisa tem o objetivo de analisar, a partir de microdados da Pesquisa Nacional por Amostra de Domicílios (PNAD) 2015, os determinantes socioeconômicos da fecundidade total no Brasil e o diferencial urbano-rural. Utiliza-se a noção de fertilidade total, de modo a se considerar apenas aquelas mulheres que tenham completado seu período fértil. A estratégia empírica tem por base um modelo *hurdle*, em que a probabilidade de a mulher ser mãe é modelada a partir de uma função log-log complementar e o número de filhos das mães brasileiras é modelado por uma função binomial negativa truncada em zero. O diferencial da fertilidade total é decomposto pelo método de Blinder-Oaxaca. A opção pelo primeiro procedimento é baseada na hipótese de que a mulher decida pelo seu próprio regime de fecundidade em duas fases distintas: primeiro é decidido entre ter ou não filhos e, posteriormente, determina-se o número de filhos desejados.

Esta pesquisa contribui para a literatura através da análise do diferencial da fertilidade completa entre as mulheres residentes nos meios urbano e rural do Brasil, tendo em vista que, pelo conhecimento dos autores, pesquisa semelhante ainda não foi realizada para o País. Ademais, os resultados encontrados podem vir a subsidiar a formulação de políticas públicas que, possivelmente, poderiam ser implementadas de modo diferenciado entre os meios urbano e rural.

Além desta introdução e das conclusões, o trabalho encontra-se organizado em mais quatro seções. Primeiro são apresentadas as teorias que embasam esta pesquisa. Posteriormente é apresentada a metodologia, destacando-se o modelo *hurdle* empregado. O tópico seguinte, por sua vez, busca descrever e analisar os principais resultados encontrados. Por fim são apresentadas as principais conclusões.

2. Hipóteses Teóricas sobre o diferencial de fecundidade

Do ponto de vista teórico, busca-se avaliar a eficácia empírica de duas hipóteses distintas que procuram explicar o diferencial de fecundidade entre os meios urbano e rural. A hipótese composicional sugere que a existência de diferentes níveis de fertilidade é explicada pelo fato de tipos distintos de mulheres residirem nos meios urbano e rural, enquanto a hipótese contextual argumenta que o grande volume de características normativas e sociais abarcadas no estilo de vida rural é de crítica importância na explicação do diferencial urbano-rural (KULU, 2011).

A hipótese composicional, como previamente exposto, procura explicar a diferença na fertilidade dos meios urbano e rural com base nas diferenças observadas para variáveis demográficas e socioeconômicas, como educação e renda. Argumenta-se, sob essa perspectiva, que as mulheres do meio rural apresentam piores níveis para estas variáveis, de modo a apresentarem um regime de fertilidade mais elevado. De fato, em diversos países os níveis de fertilidade tendem a se diferenciar de acordo o nível educacional da mulher, sendo observada uma relação inversa entre esses níveis (HOEM, 2005). O fato de a população do meio rural tender a possuir menor nível educacional leva então a um maior nível de fertilidade. Pode-se destacar também a maior presença de mulheres casadas no meio rural, em comparação ao meio urbano, o que, obviamente, favorece o aumento do nível de fertilidade local (KULU, 2011).

A hipótese contextual se baseia na ideia de que a fertilidade é diretamente influenciada por fatores culturais e econômicos relacionados ao meio em que a mulher reside. Neste sentido, teoria neoclássica da fertilidade (BECKER; LEWIS, 1973), conflui para a abordagem composicional, ao ressaltar a existência de um trade-off entre quantidade e qualidade dos filhos, dado o fato de ser mais custoso criar um filho no meio urbano do que no meio rural (BECKER, 1991). A vida no meio rural pode ser considerada como mais simples do que no meio urbano, dada a existência de menores pressões normativas sobre os pais, com destaque para o custo de oportunidade relacionado ao mercado de trabalho no meio urbano. Além das normas ligadas aos aspectos econômicos, deve-se destacar a existência de um modo de vida tradicional da população do meio rural. Neste contexto, emerge a noção de que famílias do meio rural possuiriam orientação normativa e, conseqüentemente, preferências para famílias maiores, levando a um elevado nível de fertilidade (TROVATO; GRINDSTAFF, 1980).

3. Estratégia Empírica

De modo a examinar o diferencial de fecundidade total entre as mulheres residentes nos meios rural e urbano, inicialmente foi modelada a fertilidade total das mulheres brasileiras. Atentando-se às peculiaridades comumente observadas em dados de fertilidade, optou-se pela utilização de um modelo *hurdle* (MULLAHY, 1986). Este é um modelo em duas partes que relaxa a pressuposição de que os zeros e os valores positivos observados para a variável dependente sejam geradas pelo mesmo mecanismo (CAMERON; TRIVEDI, 2009). Em um segundo momento, no intuito de se analisar, de fato, os diferenciais de fecundidade total entre as mulheres residentes nos meios urbano e rural, aplica-se o método de decomposição de Blinder-Oaxaca (BLINDER, 1973; OAXACA, 1973).

3.1. Determinantes da Fecundidade Entre as Mulheres Residentes nos Meios Rural e Urbano

Uma das principais características dos dados de fertilidade é a não-negatividade da variável de interesse, haja visto que o número de filhos tidos por uma mulher é sempre igual ou maior do que zero. Atentos a este fato, diversos estudos empíricos se basearam em modelos de regressão para dados contáveis, com destaque para os modelos de Poisson e binomial negativo (WANG; FAMOYE, 1997). No contexto de modelagem da fertilidade, o modelo binomial negativo apresenta uma clara vantagem ao modelo padrão de Poisson, dada sua flexibilidade em cenários de ausência de equidispersão.

Outra característica inerente aos dados de fertilidade está na costeira observância de uma maior quantidade de zeros do que teoricamente suposto pelas distribuições de Poisson e binomial negativa (MELKERSSON; ROTH, 2000). Buscando contornar o problema da ausência de equidispersão em dados de contagem que Cragg (1971) introduziu o primeiro modelo *hurdle*, posteriormente popularizado por Mullahy (1986).

A tradução literal para *hurdle* é obstáculo, o que explica grande parte da lógica por trás desse modelo. Considerando-se o contexto da modelagem da fertilidade, o obstáculo reside na passagem dos zeros para valores positivos da variável de interesse. Parte-se então da noção de que os dados tenham origem em um processo que começa a gerar valores positivos apenas quando se cruza a barreira dos zeros, ou seja, até a barreira ser rompida apenas zeros são gerados no processo (HILBE; 2014).

O modelo *hurdle* é, em sua essência, um modelo de duas partes. Primeiro, um processo binário dá origem a valores positivos (1) e zeros (0). Depois, um segundo processo origina apenas valores positivos. Enquanto o primeiro processo é tipicamente modelo através de um modelo binário, o segundo processo, por sua vez, é costumeiramente estimado a partir de um modelo de contagem truncado em zero (HILBE, 2011).

Neste trabalho optou-se por utilizar um modelo *hurdle* cujo processo binário é estimado através da função log-log complementar. A escolha deste modelo, em detrimento das funções *logit* e *tobit*, é motivada pelo fato de o modelo log-log complementar ser assimétrico ao redor de zero, ao contrário dos dois modelos citados. Essa assimetria é capaz de melhor acomodar distribuições da variável binária que sejam distorcidas pelo fato de haver uma grande proporção de zeros ou uns (CAMERON; TRIVEDI, 2009). Já a segunda parte do modelo *hurdle* foi estimada através de um modelo binomial negativo truncado em zero. Este modelo considera, neste segundo estágio do modelo *hurdle*, a ausência de valores zero para a variável dependente, além de ser robusto à possível sobredispersão dos dados.

Os dois processos, log-log complementar e binomial negativo truncado em zero, são combinados através da função log-verossimilhança (HILBE, 2011) em (1):

$$\mathcal{L} = \ln(f(0)) + \{\ln[1 - f(0)] + \ln P(t)\}, \quad (1)$$

em que $f(0)$ representa a probabilidade da parte binária do modelo e $P(t)$ representa a probabilidade de se observar um valor positivo. Logo, a função log-verossimilhança do modelo *hurdle* é igual ao log da probabilidade de $y = 0$, somado ao log de $y = 1$ mais o log de y , dado que ele seja um valor positivo.

No caso do modelo log-log complementar, a probabilidade de se observar o valor zero é dada por (2)

$$f(0) = P(y = 0|x) = \exp(-\exp(x'_i \beta_b)) \quad (2)$$

e a probabilidade de $y = 1$ é expressa em (3)

$$1 - f(0) = P(y = 1|x) = 1 - \exp(-\exp(x'_i \beta_b)). \quad (3)$$

Por seu turno, a função de distribuição de probabilidades binomial negativa é dada por (4)[‡]

[‡] Para derivação completa da função de distribuição de probabilidade binomial negativa, consultar Hilbe (2011).

$$f(y|\mu, \alpha) = P(t) = \frac{\Gamma(y_i + 1/\alpha)}{\Gamma(y_i + 1)\Gamma(1/\alpha)} \left(\frac{1}{1 + \alpha\mu_i}\right)^{1/\alpha} \left(\frac{\alpha\mu_i}{1 + \alpha\mu_i}\right)^{y_i}, \quad (4)$$

em que $\Gamma(\cdot)$ representa uma distribuição gama; α representa o parâmetro de heterogeneidade binomial negativa; e μ é o parâmetro de média e variância.

Combinado (2), (3) e (4) em (1), tem-se a função log-verossimilhança do modelo hurdle log-log complementar binomial negativo, parametrizada em termos dos coeficientes do modelo (5),

se $y = 0$: $-\exp(x'_i\beta_b)$

$$\begin{aligned} \text{se } y = 1: & \ln\left(1 - \exp(-\exp(x'_i\beta_b))\right) + y_i \ln\left(\frac{\exp(x'_i\beta)}{1 + \exp(x'_i\beta)}\right) \\ & - \frac{1}{\alpha} \ln(1 + \exp(x'_i\beta)) + \ln \Gamma\left(y_i + \frac{1}{\alpha}\right) \\ & - \ln \Gamma(y_i + 1) - \ln \Gamma\left(\frac{1}{\alpha}\right) \\ & - \ln\left(1 - (1 + \exp(x'_i\beta))\right)^{-1/\alpha}, \end{aligned} \quad (5)$$

em que o subscrito b indica que $x'\beta$ é um componente do modelo binário e não do modelo binomial negativo.

A base de dados utilizada neste estudo adota uma amostragem complexa e a consideração do desenho amostral se torna imprescindível para obtenção de estimativas não-viesadas quanto ao erro-padrão e a variância, por exemplo (SILVA; PESSOA; LILA, 2002). A estimação por máxima verossimilhança, segundo Greene (2003), se faz necessária neste caso. Além disso, dada a natureza complexa da amostragem empregada na base de dados utilizada, deve-se definir corretamente o método de estimação das variâncias. Nesta pesquisa, optou-se pela utilização do método de linearização de Taylor. Este procedimento, trabalhando com uma Série de Taylor, fornece uma aproximação linear para o estimador não-linear de interesse, aplicando a ela as fórmulas usuais de estimadores lineares (ALVES; SILVA, 2007).

Ademais, destaca-se também a existência de métodos capazes de verificar o impacto da incorporação do desenho amostral sobre a estimação da variância, como aqueles desenvolvidos por Kish (1965) e Skinner (1989). O primeiro, conhecido como *design-effect* (DEFF), é definido como a razão entre a variância obtida ao se considerar o plano amostral e a variância obtida ao ignorá-lo. Já o segundo, denominado de *mispecification effect* (MEFF), compara a estimativa da variância do parâmetro obtida a partir da consideração do plano amostral com outra que desconsidera peso, conglomerado e estratificação. Ambos os métodos foram utilizados nesta pesquisa.

3.2. Diferenciais de Fecundidade Total Entre as Mulheres Residentes nos Meios Urbano e Rural

Dado que a função de verossimilhança é separável com respeito aos vetores de parâmetros β e β_b , a função log-verossimilhança pode ser escrita como a soma das funções log-verossimilhança dos dois modelos separados e, deste modo, o modelo hurdle pode ser maximizado através da maximização dos dois modelos separadamente (MCDOWELL, 2003). É com base nessa separabilidade que se torna possível a aplicação da decomposição de Blinder-

Oaxaca para o segundo estágio do modelo *hurdle*, que no presente artigo é estimado através de um modelo binomial negativo truncado em zero.

O método de decomposição desenvolvido por Blinder (1973) e Oaxaca (1973) permite decompor a variável de interesse entre dois grupos em uma parte explicada pelas diferenças em características observadas e outra parte atribuída às diferenças nos retornos à essas características. Posteriormente esse modelo foi estendido pela introdução de um terceiro termo de explicação das diferenças entre os grupos analisados, formado pela interação dos dois primeiros termos de modelo de decomposição original (DAYMONT; ANDRISANI, 1984).

Tanto o modelo de Blinder-Oaxaca quanto a extensão de Daymont e Andrisani (1984) foram extensamente utilizados no contexto de modelos de regressão linear. Todavia, o modelo binomial negativo truncado em zero possui natureza não-linear. Neste sentido, optou-se por utilizar a extensão da decomposição de Blinder-Oaxaca para modelos não-lineares proposta por Bauer e Sinning (2008).

A extensão de Daymont e Andrisani (1984), exposta em termos de expectativas condicionais e adaptada ao contexto do presente estudo, é dada por (6)

$$\Delta_U^{NL} = \{E_{\beta_U}(Y_{iR}|X_{iR}) - E_{\beta_U}(Y_{iU}|X_{iU})\} + \{E_{\beta_R}(Y_{iU}|X_{iU}) - E_{\beta_U}(Y_{iU}|X_{iU})\} \quad (6)$$

$$+ \{[E_{\beta_R}(Y_{iR}|X_{iR}) - E_{\beta_U}(Y_{iR}|X_{iR})]$$

$$+ [E_{\beta_R}(Y_{iU}|X_{iU}) - E_{\beta_U}(Y_{iU}|X_{iU})]\},$$

em que Δ_U^{NL} é a diferença média da fecundidade total; $E_{\beta_g}(Y_{ig}|X_{ig})$ se refere à expectativa condicional de Y_{ig} ; $E_{\beta_h}(Y_{ih}|X_{ih})$ se refere à expectativa condicional de Y_{ih} avaliada no vetor de parâmetros β_g , com $g, h = (U, R)$ e $g \neq h$; U se refere às mulheres residentes no meio urbano; e R se refere às mulheres residentes no meio rural.

No contexto de estimadores não-lineares, Bauer e Sinning (2008) sugerem a estimação dos componentes da equação 6 pelos análogos amostrais $S(\hat{\beta}_g|X_{ig})$ e $S(\hat{\beta}_h|X_{ih})$ das expectativas condicionais $E_{\beta_g}(Y_{ig}|X_{ig})$ e $E_{\beta_h}(Y_{ih}|X_{ih})$ para $g, h = (U, R)$ e $g \neq h$. O análogo amostral da média condicional para o modelo binomial negativo é dado por (7)

$$S(\hat{\beta}_{g,NB}, X_{ig}) = \bar{Y}_{g, \hat{\beta}_{g,NB}} = \frac{1}{N_g} \sum_{i=1}^{N_g} \exp(X_{ig} \hat{\beta}_{g,NB}). \quad (7)$$

em que NB se refere à regressão binomial negativa.

3.3. Natureza e fonte dos dados

A modelagem empírica da fertilidade total das mulheres brasileiras é feita com base em microdados da PNAD 2015, por esta ser a fonte mais recente de dados acerca da fertilidade e seus condicionantes. A variável independente utilizada no modelo *hurdle* e decomposta no procedimento de Blinder-Oaxaca é o número de filhos tidos, *Filhos*. Por se tratar da fertilidade total, a amostra utilizada abarca todas as mulheres que tenham, em teoria, ultrapassado seu período fértil. Seguindo Melkersson e Rooth (2000), foram consideradas mulheres com ao menos 45 anos de idade.

As variáveis independentes buscam abarcar uma gama de características socioeconômicas e demográficas teoricamente capazes de influenciar e explicar a decisão das mulheres brasileiras pela maternidade e o número de filhos tidos durante o período fértil, bem como o diferencial urbano-rural. Mais explicitamente, as covariadas utilizadas foram: *Educação*, anos de estudo da mulher; *Idade*, conjunto de *dummies* para *coortes* de idade [(45-54), (55-64), (65-74), (≥ 75)]; *Região*, conjunto de *dummies* para as macrorregiões do Brasil; *Ocupação*, *dummy* para indicação de participação no mercado de trabalho [1 = ocupada]; *Renda*, logaritmo natural da rendimento domiciliar *per capita*; *Cor*, *dummy* para indicação de cor da mulher [1 = não-

branca]; *Conjugal*, *dummy* para indicação de situação conjugal [1 = vive em companhia de cônjuge ou companheiro]. Além das covariadas apresentadas, utiliza-se também uma *dummy* de indicação da situação censitária, de modo a possibilitar a decomposição do diferencial da fertilidade total das mulheres residentes nos meios urbano e rural. A covariada *Rural*, deste modo, apresenta valor igual a 1 quando a mulher domicilia em zona rural.

Espera-se, a princípio, que os coeficientes estimados para as *dummies* de cor, situação conjugal e situação censitária apresentem coeficientes positivos em ambas as etapas do modelo *hurdle*, acreditando-se que as mulheres não-brancas, que vivem em companhia do companheiro e domiciliadas no meio rural possuam maior probabilidade de altos níveis de fertilidade total. Por outro lado, fica esperado que a *dummy* de ocupação e as variáveis de educação e renda apresentem sinais negativos, dado que se crê na menor probabilidade de uma alta fertilidade para mulheres ocupadas e de maiores níveis educacional e de renda.

4. Resultados e Discussão

Nesta sessão são apresentados os resultados empíricos do estudo. Primeiramente são expostas as estatísticas descritivas da amostra analisada. Posteriormente, são apresentados os resultados econométricos, divididos entre a análise dos determinantes socioeconômicos relacionados à fertilidade completa no Brasil e a análise da decomposição do diferencial de fertilidade completa para as mulheres domiciliadas nos meios urbano e rural do País.

4.1. Estatísticas descritivas

A Tabela 1 apresenta as estatísticas descritivas das variáveis socioeconômicas de interesse relacionadas à questão da fecundidade total no Brasil para o ano de 2015. Estas estatísticas são baseadas na amostra utilizada nesta pesquisa, abrangendo um total de 59.987 mulheres brasileiras respondentes da PNAD, que, considerando a complexidade amostral, correspondiam a uma população de mais de 35 milhões de pessoas. Além das estatísticas para o total da amostra, segregou-se também em mulheres residentes nos meios urbano e rural.

Tabela 1: Estatísticas descritivas das variáveis de interesse, Brasil 2015*

Variáveis	Brasil		Urbano		Rural	
	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.	Prop. ou média	Desv. Pad.
<i>Filhos</i>	2,976	2,379	2,792	2,223	4,186	2,948
<i>Educação</i>	6,501	4,939	6,957	4,923	3,509	3,886
<i>Renda</i>	6,807	0,866	6,873	0,857	6,374	0,802
<i>Ocupação</i>	0,388	0,487	0,379	0,485	0,451	0,498
<i>Cor</i>	0,495	0,500	0,479	0,500	0,598	0,490
<i>Conjugal</i>	0,555	0,497	0,533	0,499	0,696	0,460
<i>Idade</i>						
(45-54)	0,387	0,487	0,388	0,487	0,383	0,486
(55-64)	0,299	0,458	0,299	0,458	0,296	0,457
(65-74)	0,165	0,389	0,185	0,389	0,190	0,392
(≥ 75)	0,128	0,335	0,128	0,334	0,131	0,338
<i>Região</i>						
Norte	0,061	0,239	0,054	0,227	0,104	0,305
Nordeste	0,258	0,437	0,227	0,419	0,461	0,499
Sudeste	0,452	0,498	0,489	0,500	0,207	0,405
Sul	0,160	0,367	0,157	0,364	0,183	0,386
Centro-Oeste	0,069	0,254	0,073	0,260	0,046	0,209
<i>Rural</i>	0,132	0,339	-	-	-	-

Fonte: Elaboração do autor a partir dos microdados da PNAD 2015.

* “Prop.” significa proporção e “Desv. pad.” significa desvio-padrão.

Observa-se, de princípio, que o nível de fecundidade total da amostra analisada se encontrou em torno de 3 filhos por mulher, destacando-se que as mulheres rurais possuíam, em média, quase 1,5 filhos a mais do que as mulheres urbanas. Poder-se-ia pensar que essa diferença relativamente expressiva não confirme a tendência de uniformização da fecundidade das mulheres brasileiras, já destacada por Berquó e Cavenaghi (2005). Todavia, ressalta-se que esta pesquisa se baseia na fertilidade completa, trabalhando-se com mulheres com mais de 45 anos de idade. Quando analisada a diferença de fertilidade total por *coortes* de idade se observa uma clara tendência de diminuição do número de filhos tidos, havendo também uma acentuada aproximação dos níveis de fertilidade total entre as mulheres domiciliadas nos meios urbano e rural (Tabela 1A do Anexo).

Outra clara divergência entre mulheres residentes nos meios urbano e rural reside no nível de escolaridade. Enquanto as mulheres do meio urbano apresentaram, em média, aproximadamente 7 anos de estudo, as mulheres do meio rural possuíam a metade do nível de escolaridade. Estes números podem, a princípio, parecerem pequenos, mas, novamente, ressalta-se que esta pesquisa trabalha com mulheres que, em 2015, possuíam 45 anos de idade ou mais. Deste modo, essa é uma indicação do baixo nível de escolaridade média alcançada pelas mulheres das *coortes* analisadas, com as mulheres domiciliadas no meio rural se destacando negativamente.

Quanto à cor, quase metade do total de mulheres analisadas eram brancas, enquanto essa proporção foi de 52% para o meio urbano e 40% para o meio rural. Em relação à situação conjugal, 55% do total das mulheres residiam em conjunto com o cônjuge ou companheiro, ao passo que esse percentual era de 53 e 70% para as mulheres dos meios urbano e rural, respectivamente. No tangente ao nível de ocupação, a amostral geral e a subamostra urbana apresentaram níveis similares, com 39 e 38% das mulheres ocupadas, respectivamente. No meio rural, por outro lado, 45% das mulheres estavam participando no mercado de trabalho. Dado o fato de significativa parcela da amostra analisada possuir idade relativamente avançada, acredita-se que parte dessas mulheres tenham trabalhado quando mais jovens e estariam agora aposentadas. Ademais, o rendimento familiar *per capita*, medido em logaritmo natural, mostrou-se semelhante para as mulheres urbanas e rurais, com um valor ligeiramente superior para as mulheres do primeiro grupo.

Embora existam mulheres de todas as *coortes* de idade na amostra analisada, expressiva parte delas estão contidas no dois *coortes* mais novos, abarcando quase 70% do total da amostra. Essa proporção é similar para os dois subgrupos. Há predominância, na amostra, de mulheres das regiões Nordeste e, principalmente, Sudeste, com essa observação se mantendo também quando consideradas somente as mulheres do meio urbano. Todavia, para as mulheres do meio rural, embora essas também sejam as regiões geográficas de maior representatividade, o Nordeste contribui com a maior parte das mulheres deste grupo, 46%. Por fim, destaca-se que pouco mais de 13% das mulheres da amostra se encontravam domiciliadas no meio rural.

Finalizando a análise descritiva, conduziu-se um teste de médias, de modo analisar se os valores médios das variáveis de interesse seriam estatisticamente diferentes entre as subpopulações urbana e rural. Excetuando-se as variáveis de *coorte* de idade e de indicação geográfica, as diferenças entre as demais características populacionais foram estatisticamente significativas a 1% de significância.

4.2. Determinantes socioeconômicos da fertilidade completa no Brasil

A estimação de $P(y = 0|x)$ foi realizada através de um modelo log-log complementar, como especificado nas considerações metodológicas. Da amostral total analisada, aproximadamente 12% das mulheres não tiveram filhos durante a idade fértil. As mulheres do meio urbano seguiram esta tendência, dado que 12,5% delas eram mães, enquanto apenas 8% das mulheres domiciliadas no meio rural não possuíam filhos. O desbalanceamento entre a quantidade de mulheres com e sem filhos justifica a utilização do modelo log-log complementar, pois ele se ajusta melhor a este tipo de cenário do que os modelos *logit* e *probit*, por exemplo. A Tabela 2 mostra os resultados da estimação de $P(y = 0|x)$, considerando o modelo log-log complementar.

Tabela 2: Estimativas de $P(y = 0|x)$ para o modelo *hurdle* (log-log complementar).

Variáveis	BRASIL				URBANO				RURAL			
	Coef.	E. M. ¹	DEFF ²	MEFF ³	Coef.	E. M. ¹	DEFF ²	MEFF ³	Coef.	E. M. ¹	DEFF ²	MEFF ³
<i>Constante</i>	1,554***	-	1,227	1,356	1,532***	-	1,278	1,419	1,644***	-	1,048	1,179
<i>Educação</i>	-0,017***	-0,004***	1,282	1,363	-0,019***	-0,005***	1,307	1,392	0,003 ^{NS}	0,001 ^{NS}	1,055	1,148
<i>Renda</i>	-0,128***	-0,031***	1,291	1,315	-0,120***	-0,030***	1,332	1,353	-0,187***	-0,035***	1,091	1,189
<i>Ocupação</i>	0,071***	0,017***	1,160	1,147	0,058***	0,014***	1,233	1,222	0,130***	0,024***	1,148	1,142
<i>Cor</i>	-0,002 ^{NS}	-0,000 ^{NS}	1,297	1,300	-0,001 ^{NS}	-0,000 ^{NS}	1,293	1,295	-0,003 ^{NS}	-0,001 ^{NS}	1,114	1,110
<i>Conjugal</i>	0,463***	0,112***	1,225	1,300	0,448***	0,111***	1,280	1,353	0,580***	0,109***	1,185	1,239
<i>Idade</i>												
(55-64)	0,043***	0,010***	1,318	1,343	0,030*	0,008*	1,260	1,282	0,149***	0,028***	1,336	1,365
(65-74)	0,103***	0,025***	1,289	1,369	0,071***	0,018***	1,343	1,433	0,344***	0,065***	1,224	1,241
(≥ 75)	0,184***	0,044***	1,274	1,381	0,153***	0,038***	1,305	1,405	0,393***	0,074***	1,284	1,496
<i>Região</i>												
Nordeste	-0,136***	-0,033***	1,043	1,870	-0,126***	-0,031***	1,001	1,862	-0,164**	-0,031**	1,268	1,961
Sudeste	-0,130***	-0,031***	1,092	1,899	-0,121***	-0,030***	1,080	1,966	-0,185**	-0,035**	1,296	1,689
Sul	-0,062**	-0,015**	1,226	2,014	-0,069**	-0,017**	1,224	2,099	-0,007 ^{NS}	-0,001 ^{NS}	1,255	1,642
C. Oeste	-0,064*	-0,015*	1,278	2,235	-0,067*	-0,017*	1,261	2,267	-0,005 ^{NS}	-0,001 ^{NS}	1,327	2,129
<i>Rural</i>	-0,027 ^{NS}	-0,006 ^{NS}	1,260	1,199	-	-	-	-	-	-	-	-
Observações	59.987				52.503				7.484			
Pop. Considerada	35.127.751				30.481.620				4.646.131			
F (Prob.)	0,000				0,000				0,000			

Fonte: Elaboração do autor a partir dos microdados da PNAD 2015.

Notas: ¹ E. M. significa Efeito Marginal; ² DEFF significa *Design Effect*; ³ MEFF significa *Misspecification Effect*; *** Valor significativo a 1% de significância; ** Valor significativo a 5% de significância; * Valor significativo a 10% de significância; ^{NS} Valor não significativo a 10% de significância.

Os coeficientes e os efeitos marginais são apresentados juntamente com as estatísticas DEFF e MEFF, para a amostra geral e também para os subgrupos urbano e rural. Considerando-se os valores apontados pela estatística MEFF, é possível afirmar que, caso a natureza complexa do plano amostral da PNAD fosse desconsiderada, a variância de todas as variáveis seriam subestimadas, levando a obtenção de estimativas viesadas.

O modelo log-log complementar foi globalmente significativo (estatística F) para as três amostras – geral, urbano e rural – insinuando que tal modelo, considerando as variáveis independentes consideradas, descreve adequadamente os determinantes socioeconômicos da probabilidade da mulher ter ou não filhos. Resultados similares foram obtidos para a amostra geral e as duas subamostras analisadas, embora possam ser observadas ligeiras diferenças quanto à magnitude e significância de determinadas variáveis especialmente para a subamostra das mulheres domiciliadas no meio rural brasileiro. Por exemplo, a variável *Educação* foi estatisticamente igual a zero para as mulheres rurais, significando que a variação do nível educacional das mulheres desse grupo não influencia a probabilidade de a mulher ser mãe. Por outro lado, considerando uma mulher do meio urbano possuindo as características médias dessa subamostra, um ano adicional de estudo diminui a probabilidade de ser mãe em 0,005.

A variável *Renda* também apresentou impacto negativo sobre a probabilidade de a mulher ser mãe, sendo significativa a 1% de probabilidade para todas as amostras. O aumento de um desvio-padrão no logaritmo natural do rendimento familiar per capita, aproximadamente R\$ 14,00, diminui a probabilidade de a mulher ter filhos em 0,03. Para todas as amostras a variável *Cor* não foi estatisticamente, indicando que o fato de a raça da mulher não influenciar a probabilidade de ser mãe.

Diferentemente do esperado, se a mulher participa do mercado de trabalho, a probabilidade de ela possuir filhos é 0,017 maior do que uma mulher que não estivesse participando do mercado de trabalho, mantendo constante as demais variáveis. O impacto para as mulheres do meio urbano é ligeiramente menor (0,014), enquanto há um maior efeito sobre as mulheres do meio rural (0,024). A variável *Conjugal*, estatisticamente significativa a 1% de significância para as três amostras, apontou que, *ceteris paribus*, mulheres que viviam com o cônjuge ou companheiro tem probabilidade 0,11 maior de ser mãe do que sua contraparte, sendo o resultado similar para todas as amostras.

Os resultados observados para as variáveis de *coorte* de idade apontam que as mulheres situadas nas *coortes* mais velhas tendem a apresentar maior probabilidade de ter filhos em comparação às mulheres do *coorte* de idade mais novo, de 55 a 64 anos, sendo o efeito tão maior quanto mais elevada é a idade do *coorte*. As estimativas para as variáveis regionais demonstraram que, para a amostra geral e a subamostra urbana, as mulheres das demais regiões possuem menor probabilidade de serem mães do que as mulheres da região Norte, sendo o efeito mais expressivo para as mulheres das regiões Nordeste e Sudeste. Para a subamostra rural, as mulheres residentes nas regiões Nordeste e Sudeste também apresentaram impacto estatisticamente significativo sobre a probabilidade de se ter filhos, com magnitude similar à das demais amostras, embora os coeficientes relacionados às regiões Sul e Centro-Oeste tenham se mostrado estatisticamente iguais a zero.

Finalizando-se a análise do modelo binomial, destaca-se o fato de a variável *Rural* não ter se mostrado estatisticamente significativa. Ou seja, ficou observado que o fato de a mulher domiciliar no meio rural, mantendo constantes as demais variáveis, não apresenta efeito estatístico sobre a probabilidade de a mulher ter filhos. Apesar deste resultado inesperado, a continuidade da pesquisa não é afetada, tendo em vista que a comparação entre as mulheres dos

meios urbano e rural é feita, prioritariamente, em relação ao total de filhos tidos e não à probabilidade de se ter tido filhos.

Passando-se para os determinantes da fertilidade total das mães analisadas, expostos na Tabela 3, observou-se que a maior parte das covariadas foi significativa a 1% de significância, evidenciando a relevância desses regressores na explicação da variável dependente. Nas três regressões a variável *Ocupação* foi estatisticamente igual a zero, demonstrando que o fato de a mulher estar inserida no mercado de trabalho não possui efeitos estatísticos sobre o número de filhos. Para a amostra rural, a variável *Conjugal* também não se mostrou estatisticamente significativa, indicando que o fato de a mulher residir em companhia de seu cônjuge afeta apenas a probabilidade de ela ser mãe, mas não tem influência estatística sobre o número de filhos tidos. Diferentemente do observado para a subamostra rural, nas demais regressões a variável *Conjugal* foi estatisticamente significativa, impactando positivamente o número de filhos tidos.

As variáveis *Educação* e *Renda* apresentaram efeito negativo sobre o número total de filhos tidos pelas mães brasileiras, ou seja, a elevação do nível de escolaridade da mulher e o aumento do rendimento domiciliar *per capita* levam a uma diminuição do número de filhos tidos. Estes resultados vão ao encontro do observado por Madalozzo (2012) que, com base em dados de 2009, apontou que o aumento do nível educacional e da renda levam a uma diminuição do número de filhos das mulheres brasileiras. Deve-se ressaltar, inclusive, que o efeito dessas covariadas é mais acentuado para as mulheres domiciliadas no meio rural. Quanto à variável *Cor*, ficou observado que as mulheres não-brancas possuem, em média, maior número de filhos do que as mulheres brancas.

As covariadas de *coorte* de idade evidenciaram o anteriormente observado na análise das estatísticas descritivas. Em comparação ao *coorte* de idade mais baixa, as demais *coortes* apresentam expressivo efeito sobre o número de filhos tidos pelas mulheres brasileiras, com a magnitude se elevando a medida que se evolui para *coortes* de idade mais elevada. O fato de o efeito marginal estimado ter sido comparativamente mais elevado para as mulheres do meio rural, especialmente para as duas últimas *coortes*, corroborando a diminuição da divergência na fertilidade total entre as mulheres dos meios rural e urbano, conforme já apontado por Tejada *et al.* (2017).

A região Norte é aquela com maior média de fertilidade total no Brasil. Quando comparadas à ela, todas as variáveis regionais apresentaram expressivo efeito marginal, com destaque para as regiões Sudeste e Sul. Este resultado se mostrou válido tanto para o meio urbano quanto para o meio rural. Essas diferenças macrorregionais no padrão de fecundidade vêm há muito sendo comprovadas empiricamente para o Brasil (POTTER; SCHMERTMANN; CAVENAGHI, 2002; SCHMERTMANN; POTTER; CAVENAGHI, 2007). Por fim, mantendo as demais covariadas constantes, o fato de a mulher residir no meio rural faz com que ela tenham aproximadamente 50% mais filhos do que as mulheres domiciliadas no meio urbano. Este diferencial de fecundidade é há muito destacado pela literatura especializada no Brasil (BERQUÓ; CAVENAGHI, 2005).

Tabela 3: Estimativas de $P(y > 0|x)$ para o modelo *hurdle* (binomial negativo truncado em zero).

Variáveis	BRASIL				URBANO				RURAL			
	Coef.	E. M. ¹	DEFF ²	MEFF ³	Coef.	E. M. ¹	DEFF ²	MEFF ³	Coef.	E. M. ¹	DEFF ²	MEFF ³
<i>Constante</i>	2.063***	-	1.096	1.230	2.094***	-	1.061	1.225	2.178***	-	1.265	1.259
<i>Educação</i>	-0.039***	-0.114***	1.202	1.198	-0.039***	-0.108***	1.194	1.182	-0.035***	-0.148***	0.897	0.979
<i>Renda</i>	-0.119***	-0.344***	1.154	1.214	-0.121***	-0.331***	1.111	1.195	-0.113***	-0.474***	1.345	1.320
<i>Ocupação</i>	-0.005 ^{NS}	-0.014 ^{NS}	1.132	1.071	-0.011 ^{NS}	-0.031 ^{NS}	1.157	1.066	0.018 ^{NS}	0.075 ^{NS}	1.003	1.023
<i>Cor</i>	0.078***	0.226***	1.189	1.169	0.082***	0.223***	1.164	1.154	0.053***	0.224***	1.217	1.156
<i>Conjugal</i>	0.036***	0.104***	1.258	1.301	0.045***	0.123***	1.273	1.309	-0.014 ^{NS}	-0.058 ^{NS}	1.218	1.278
<i>Idade</i>												
(55-64)	0.199***	0.576***	1.170	1.048	0.182***	0.497***	1.219	1.067	0.269***	1.133***	1.160	1.131
(65-74)	0.425***	1.229***	1.198	1.177	0.412***	1.125***	1.246	1.223	0.474***	1.998***	1.231	1.197
(≥ 75)	0.566***	1.638***	1.208	1.284	0.568***	1.550***	1.292	1.388	0.553***	2.327***	1.077	1.074
<i>Região</i>												
Nordeste	-0.166***	-0.479***	1.030	1.741	-0.185***	-0.506***	0.908	1.614	-0.114***	-0.481***	1.083	1.576
Sudeste	-0.316***	-0.913***	1.032	1.547	-0.330***	-0.900***	0.904	1.445	-0.256***	-1.077***	1.502	1.713
Sul	-0.277***	-0.802***	1.042	1.480	-0.272***	-0.741***	0.914	1.376	-0.332***	-1.399***	1.276	1.495
C. Oeste	-0.190***	-0.550***	0.901	1.349	-0.200***	-0.545***	0.814	1.261	-0.178***	-0.751***	0.813	1.173
<i>Rural</i>	0.188***	0.544***	1.411	1.480	-	-	-	-	-	-	-	-
Observações	52.856				45.972				6.884			
Pop. Considerada	30.941.149				26.687.681				4.253.468			
F (Prob.)	0,000				0,000				0,000			

Fonte: Elaboração do autor a partir dos microdados da PNAD 2015.

Notas: ¹ E. M. significa Efeito Marginal; ² DEFF significa *Design Effect*; ³ MEFF significa *Misspecification Effect*; *** Valor significativo a 1% de significância; ** Valor significativo a 5% de significância; * Valor significativo a 10% de significância; ^{NS} Valor não significativo a 10% de significância.

4.3. Diferencial da fertilidade total para os meios urbano e rural

A decomposição do diferencial de fecundidade entre os meios urbano e rural no Brasil é apresentada na Tabela 4. O cálculo é feito com base apenas naquelas mulheres que possuíam filhos. Neste sentido, foram analisadas 7.077 mulheres residentes no meio rural brasileiro e 47.320 mulheres domiciliadas no meio urbano do Brasil. A diferença é apresentada em termos das mulheres domiciliadas no meio rural, de modo a se apresentar um diferencial positivo, facilitando a análise dos resultados. O método *bootstrap* foi aplicado, efetuando-se 50 repetições, de modo a gerar os erros-padrão dos componentes da equação de decomposição. A Tabela 4 aponta os resultados da decomposição, dividindo-os em três partes distintas. A primeira parte reflete a elevação média no nível de fecundidade das mulheres do meio urbano caso estas possuíssem as mesmas características das mulheres do meio rural. Estas características são captadas através dos valores das variáveis explicativas. Como observado na análise descritiva dos dados, são muitas as discrepâncias entre os valores médios entre os dois grupos analisados. Observa-se que, neste caso, haveria um aumento médio de 0,74 filhos entre as mulheres urbanas. Logo, este resultado fornece evidências para a sustentação da hipótese composicional, tendo em vista o potencial aumento na fecundidade das mulheres urbanas caso elas apresentassem as características das mulheres rurais.

Tabela 4: Diferencial urbano-rural de fecundidade total, Brasil 2015

Decomposição	Coefficientes	Erros-padrão
Características	0,7423***	0,0209
Coefficientes	0,7391***	0,0411
Interação	0,0307 ^{NS}	0,0317
Diferença bruta	1,5121***	0,0391

Fonte: Elaboração do autor a partir dos microdados da PNAD 2015.

Nota: *** Valor significativo a 1% de significância.

Por seu turno, a segunda parte da decomposição quantifica a mudança que seria observada no número de filhos tidos pelas mulheres do meio urbano quando aplicados os coeficientes estimados para as mulheres do meio rural sobre as características das mulheres urbanas. Em menor proporção muito similar daquela vista na parte anterior, a aplicação dos coeficientes do outro grupo seria capaz de aumentar em aproximadamente 0,74 o nível de fecundidade das mulheres residentes no meio urbano. Pode-se interpretar que esse resultado seja favorável ao reconhecimento da hipótese contextual. Tomando os coeficientes estimados como indicadores do impacto da variação das características socioeconômicas sobre a fertilidade total, é possível interpretá-los como indicadores do contexto ao qual as mulheres pertencem. Deste modo, quando transferidas ao contexto rural, as mulheres do meio urbano tenderiam a elevar seus níveis de fertilidade.

Já a terceira parte e última parte diz respeito a um termo de interação, medindo o efeito simultâneo das diferenças nas características e nos coeficientes dos dois grupos. Para interpretar este termo, deve-se pensar na aplicação da diferença nos coeficientes estimados na diferença dos dados (características) das duas amostras. Este termo pode então ser visto como o diferencial esperado pela soma das duas diferenças individuais (características e coeficientes) e o resultado propriamente dito. Neste sentido, o valor da interação de 0,03 aponta que a soma das duas diferenças foi ligeiramente mais elevada do que o diferencial propriamente dito. Todavia, esse valor não foi estatisticamente significativo, implicando que os efeitos das características e dos coeficientes afetam o diferencial apenas de modo separado, não sendo observada interação entre eles.

Conclusões

Em se tratando da escolha das mulheres pela maternidade, os resultados mostraram que o nível educacional não influencia a probabilidade de as mulheres do meio rural possuírem um ou mais filhos. Além disso, a raça também não pareceu ter influência sobre a maternidade para mulheres de ambos os meios analisados. Conforme inicialmente esperado, mulheres de *coorte* de idade mais avançada apresentaram maior probabilidade de serem mães. Contrariando as expectativas iniciais, o fato de a mulher domiciliar no meio rural não apresentou impacto sobre a probabilidade de a mulher possuir filhos.

No tangente à segunda equação do modelo *hurdle* estimado, que buscou medir o impacto de determinadas variáveis socioeconômicas sobre o número de filhos das mulheres brasileiras, destaca-se principalmente que o fato de a mulher residir no meio rural está associado a um aumento de aproximadamente 50% na fecundidade total em comparação com as mulheres domiciliadas no meio urbano. Ademais, ficou evidenciado que as variáveis de nível educacional e de rendimento familiar *per capita* afetam negativamente a fertilidade total, enquanto o fato de a mulher ser não-branca tem efeito positivo sobre o número de filhos tidos. As *coortes* de idade apresentaram efeitos positivos, com magnitude menor para *coortes* de menor idade, fornecendo evidências sobre uma possível convergência entre fertilidade total das mulheres dos meios rural e urbano.

Ao se considerar os resultados obtidos pela decomposição de Blinder-Oaxaca, destaca-se a estimação de que as mulheres domiciliadas no meio rural possuem, em média, 1,5 filhos a mais do que aquelas residentes no meio urbano. Esse diferencial é decomposto quase que igualmente entre os componentes de características e de coeficientes, com cada influenciando o diferencial de fertilidade em 0,75 filhos. Deste modo, conclui-se que tanto a hipótese contextual quanto a hipótese composicional podem ser sustentadas pelos resultados aqui encontrados. Deste modo, à medida que o diferencial entre os fatores socioeconômicos das mulheres dos meios urbano e rural brasileiros se estreitar, haverá uma tendência de o nível de fecundidade entre estes dois grupos se tornar próximos entre eles. Todavia, acredita-se que, com base nos efeitos de composição, ainda assim o nível de fertilidade total das mulheres do meio rural se manterá mais elevado do que aquele observado para o meio urbano.

Referências

- Alves, M. C. G. P., Silva, N. N. (2007). Métodos de estimação de variância em amostras provenientes de inquéritos domiciliares. *Revista de Saúde Pública*. 41(6). 938-946.
- Bauer, T. K., Sinning, M. (2008). An extension of the Blinder–Oaxaca decomposition to nonlinear models. *Advances in Statistical Analysis*. 92(2). 197–206.
- Becker, G. S. (1991). *A Treatise on the Family*. Cambridge: Harvard University Press.
- Becker, G. S., Lewis, H. G. (1973). On the Interaction between the Quantity and Quality of Children. *Journal of Political Economy*. 81(2). S279-S288.
- Berquó, E., Cavenaghi, S. M. (2005). Brazilian fertility regimes: profiles of women below and above replacement levels. Disponível em: <http://demoscope.ru/weekly/knigi/tours_2005/papers/iussp2005s51864.pdf>. Acesso em 04/05/2017.
- Berquó, E., Cavenaghi, S. M. (2006). Breve nota sobre a redução no número médio de filhos por mulher no Brasil. *Novos Estudos – CEBRAP*. (74). 11-15.
- Berquó, E., Cavenaghi, S. M. (2014). Notas sobre os diferenciais educacionais e econômicos da fecundidade no Brasil. *Revista Brasileira de Estudos Populacionais*. 31(2). 471-482.
- Blinder, A. S. (1973). Wage Discrimination: Reduced Form and Structural Estimates. *The Journal of Human Resources*. 8(4). 436–455.
- Cameron, A. C., Trivedi, P. K. (2009). *Microeconometrics Using Stata*. College Station: Stata Press.
- Cragg, J. G. (1971). Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods. *Econometrica*. 39(5). 829-844.
- Daymont, T. N., Andrisani, P. J. (1984). Job preferences, college major, and the gender gap in earnings. *Journal of Human Resources*. 19(3). 408–428.
- Greene, W. H. (2003). *Econometric Analysis*. Upper Saddle River: Prentice Hall.
- Hilbe, J. M. (2011). *Negative Binomial Regression*. Nova Iorque: Cambridge University Press.
- Hilbe, J. M. (2014). *Modeling Count Data*. Nova Iorque: Cambridge University Press.
- Hoem, J. M. (1987). Statistical analysis of a multiplicative model and its application to the standardization of vital rates: a review. *International Statistical Review*. 55(2). 119–152.
- Kish, L. (1965). *Survey sampling*. New York: Wiley.
- Kulu, H. (2013). Why Do Fertility Levels Vary between Urban and Rural Areas? *Regional Studies*. 47(6). 895-912.
- Madalozzo, R. (2012). Transitions in Fertility for Brazilian Women: An Analysis of Impact Factors. *PLoS ONE*. 7(7).
- McDowell, A. (2003). From the help desk: hurdle models. *The Stata Journal*. 3(2). 178–184.
- Melkersson, M., Rooth, D. (2000). Modeling female fertility using inflated count data models. *Journal of Population Economics*. 13(2). 189-203.
- Miro, C. A., Mertens, W., Davis, K. (1968). Influences Affecting Fertility in Urban and Rural Latin America. *The Milbank Memorial Fund Quarterly*. 46(3). 89-120.
- Mullahy, J. (1986). Specification and Testing in Some Modified Count Data Models. *Journal of Econometrics*. 33(3). 341-365.
- Oaxaca, R. (1973). Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review*. 14(3). 693–709.
- Potter, J. E., Schmertmann, C. P., Cavenaghi, S. M. (2002). Fertility and Development: Evidence from Brazil. *Demography*. 39(4). 739-761.
- Robinson, W. C. (1961). Urban-Rural Differences in Indian Fertility. *Population Studies*. 14(3). 218-234.



- Schmertmann, C. P., Potter, J. E., Cavenaghi, S. M. (2008). Exploratory Analysis of Spatial Patterns in Brazil's Fertility Transition. *Population Research and Policy Review*. 27(1). 1-15.
- Silva, P. L. N., Pessoa, D. G. C., Lila, M. F. (2002). Análise estatística de dados da PNAD: incorporando a estrutura do plano amostral. *Ciência & Saúde Coletiva*. 7(4). 659–670.
- Skinner, C. J. (1989). Domain means, regression and multivariate analysis. In: Skinner, C. J., Holt, D., Smith, T. M. F. (Eds.) *Analysis of Complex Surveys*. Chichester: Wiley. 59-87.
- Tejada, C. A. O., Triaca, L. M., Costa, F. K., Hellwig, F. (2017). The sociodemographic, behavioral, reproductive, and health factors associated with fertility in Brazil. *PLoS ONE*. 12(2).
- Trovato, F., Grindstaff, C. F. (1980). Decomposing the Urban-Rural Fertility Differential: Canada, 1971. *Rural Sociology*. 45(3). 448-468.
- Wang, W., Famoye, F. (1997). Modeling household fertility decisions with generalized Poisson regression. *Journal of Population Economics*. 10(3). 273-283.

Anexo

Tabela 1A – Fertilidade total por coortes de idade e situação censitária, Brasil 2015.

Situação censitária	Coorte de idade			
	(45-54)	(55-64)	(65-74)	(≥ 75)
Rural	3,314281	4,122637	5,166248	5,491869
Urbano	2,227819	2,639203	3,353319	3,953829
Diferença	1,086462	1,483434	1,812939	1,538040

Fonte: PNAD 2015.