

UNIVERSIDADE FEDERAL DE MINAS GERAIS  
FACULDADE DE CIÊNCIAS ECONÔMICAS  
CENTRO DE DESENVOLVIMENTO E PLANEJAMENTO REGIONAL

FABIANA DE FÁTIMA VIEIRA

AVALIAÇÃO DO ENSINO FUNDAMENTAL DE NOVE  
ANOS

BELO HORIZONTE  
2015

<p>Fabiana de Fátima Vieira</p>	<p>Avaliação do ensino fundamental de nove anos</p>		<p>UFMG – FACE CEDEPLAR 2015</p>
---	---	--	--

Fabiana de Fátima Vieira

# AVALIAÇÃO DO ENSINO FUNDAMENTAL DE NOVE ANOS

Dissertação apresentada ao Centro de Desenvolvimento e Planejamento Regional da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do Título de Mestre em Demografia.

Área de concentração: Demografia da Educação.

Orientador: Prof. Dr. Eduardo Luiz Gonçalves Rios-Neto

Belo Horizonte  
2015

*Após um longo período de questionamentos, após um longo período de busca e redescobrimento, após um longo período de dedicação intensa... A conclusão deste trabalho só foi possível pela paciência e apoio incondicional de papai, mamãe, minha irmã e meu marido. Não há palavras para expressar tamanha cumplicidade e gratidão, mas pelo menos posso dedicar todo esse trabalho a vocês...*

## **AGRADECIMENTOS**

Sem dúvida alguma a conclusão de um curso de mestrado é mérito de todo um corpo docente que dedica preciosas horas a formar profissionais e acima de tudo pessoas. Por isso agradeço a todos os professores do CEDEPLAR que fizeram parte da minha formação. Mas, devido a uma dedicação toda especial, uma enorme compreensão e muita solidariedade nos momentos de dúvidas e problemas pessoais devo agradecer de modo especial à Ana Maria Hermeto Camilo de Oliveira, Eduardo Luiz Gonçalves Rios-Neto, Paula Miranda-Ribeiro e Simone Wajnman. Não tenho a menor dúvida de que esta dissertação só foi concluída por causa destes professores. Obrigada pela paciência, pela fé que depositaram em mim e por fazerem tudo o que era possível (a até o que parecia impossível) para que eu pudesse ser uma demógrafa.

Agradeço também a alguns amigos que direta ou indiretamente foram cruciais para que esse trabalho fosse feito. À amiga Raquel Rangel de Meireles Guimarães agradeço de modo especial pelas contribuições nos exercícios empíricos. Aos amigos Ulisses Pereira dos Santos e Raphael Rocha Gouvea agradeço pelo apoio e pelo incentivo em todos os momentos. E a todos os amigos queridos que fizeram parte da coorte 2007 eu agradeço pela troca de conhecimento e pelas boas lembranças da convivência prazerosa.

Por fim, e acima de tudo, agradeço a Deus. Seu amor e misericórdia foram determinantes para que eu conseguisse construir esse trabalho e concluir meu mestrado.

*“Se a educação sozinha não pode transformar a sociedade, tampouco sem ela a sociedade muda.”*

*(Paulo Freire)*

## RESUMO

O ensino fundamental de nove anos foi uma política educacional adotada a nível nacional para dirimir as grandes diferenças no rendimento das crianças durante o ensino fundamental. Estas diferenças foram observadas entre aqueles estudantes que já haviam frequentado a escola ao ter acesso à educação infantil e aqueles que não haviam frequentado a escola até serem matriculados no ensino fundamental. A ideia era antecipar o ciclo de vida do estudante diminuindo em um ano a idade de matrícula obrigatória no ensino fundamental. Considerando que essa política poderia ter impactos sobre a matrícula e sobre a probabilidade de progressão por série, este trabalho analisou os dois indicadores usando dados da PNAD para verificar se estes impactos foram positivos e substanciais. O método usado neste trabalho foi o método da dupla diferença e o foco da análise foi o estado de Minas Gerais que foi comparado a duas outras unidades da federação: São Paulo e Bahia. Os resultados mostraram que inicialmente a antecipação do ciclo de vida do estudante não teve impactos substanciais sobre a matrícula e a progressão por série dos estudantes do ensino fundamental de Minas Gerais.

---

**Palavras-chave:** educação básica, ensino fundamental de nove anos, ciclo de vida do estudante.

## ABSTRACT

Elementary school of nine years was a educational policy adopted at national level to settle the large differences in the return (productivity) of children during elementary school. These differences were observed between those students who had already attended school, because they had access to early childhood education, and those who had not attended school until they are enrolled in elementary school. The idea was anticipating the life cycle of student decreasing in a year the age of compulsory registration in elementary school. Whereas this policy could have impacts on rate of enrollment and on probability of progression by series, this work analyzed these two indicators using data from the PNAD (National Household Sample Survey) to check if these impacts were positive and substantial. The method used in this work was the method of double difference and the focus of the analysis was the state of Minas Gerais that was compared to two other units of the federation: Sao Paulo and Bahia. The results showed that initially the anticipation of the life cycle of the student had no substantial impact on the registration and progression by series of elementary school students in Minas Gerais.

---

**Keywords:** Basic school, elementary school of nine years, school transition, life cycle of the student.



## LISTA DE ILUSTRAÇÕES

Quadro 1 – Nova organização do ensino fundamental.....	29
Quadro 2 - Estrutura simples do método da dupla diferença.....	32
Gráfico 1 - Representação gráfica dos dois grupos do método da dupla diferença.....	33
Gráfico 2 - Diferença obtida para o grupo de tratamento.....	34
Gráfico 3 - Diferença obtida entre o grupo de tratamento e o grupo de controle depois da intervenção.....	34
Gráfico 4 - Representação correta da estimação da dupla diferença.....	35
Quadro 3 - Desenho da amostra do exercício I.....	42
Quadro 4 - Desenho da amostra do exercício II.....	42
Quadro 5 - Desenho da amostra do exercício III.....	43

## LISTA DE TABELAS

Tabela 1: Matrículas no ensino fundamental em 2004 por UF.....	30
Tabela 2 - Número de matrículas de 1ª a 4ª série por unidade federativa – exercício I.....	45
Tabela 3: Número de matrículas de 1ª a 4ª série antes e depois da intervenção – exercício I.....	45
Tabela 4 - Número de matrículas de 1ª a 4ª série por unidade federativa – exercício II.....	47
Tabela 5 - Número de matrículas de 1ª a 4ª série antes e depois da intervenção – exercício II.....	47
Tabela 6 - Resumo dos resultados do Exercício I com o modelo de probabilidade linear – MG e SP.....	50
Tabela 7 - Resumo dos resultados do Exercício I com o modelo logit – MG e SP.....	50
Tabela 8 - Resumo dos resultados do Exercício I com o modelo de probabilidade linear com desenho complexo – MG e SP.....	51
Tabela 9 - Resumo dos resultados do Exercício I com o modelo logit com desenho complexo – MG e SP.....	51
Tabela 10 - Resumo dos resultados do Exercício II com o modelo de probabilidade linear – MG e SP.....	52
Tabela 11 - Resumo dos resultados do Exercício II com o modelo logit – MG e SP.....	53
Tabela 12 - – Resumo dos resultados do Exercício II com o modelo de probabilidade linear com desenho complexo – MG e SP.....	53

Tabela 13 – Resumo dos resultados do Exercício II com o modelo logit com desenho complexo – MG e SP.....	53
Tabela 14 - Resumo dos resultados do Exercício III com o modelo de probabilidade linear – MG e SP.....	54
Tabela 15 - Resumo dos resultados do Exercício III com o modelo de probabilidade linear – MG e SP.....	55
Tabela 16 - Resumo dos resultados do Exercício III com o modelo de probabilidade linear com desenho complexo – MG e SP.....	56
Tabela 17 - Resumo dos resultados do Exercício III com o modelo logit com desenho complexo – MG e SP.....	57
Tabela 18 - Resumo dos resultados do Exercício I com o modelo de probabilidade linear – MG e BA.....	58
Tabela 19 - Resumo dos resultados do Exercício I com o modelo logit – MG e BA.....	58
Tabela 20 - Resumo dos resultados do Exercício I com o modelo de probabilidade linear com desenho complexo – MG e BA.....	59
Tabela 21 - Resumo dos resultados do Exercício I com o modelo logit com desenho complexo – MG e BA.....	59
Tabela 22 - Resumo dos resultados do Exercício II com o modelo de probabilidade linear – MG e BA.....	60
Tabela 23 - Resumo dos resultados do Exercício II com o modelo logit – MG e BA.....	60

Tabela 24 - Resumo dos resultados do Exercício II com o modelo de probabilidade linear com desenho complexo – MG e BA.....	61
Tabela 25 - Resumo dos resultados do Exercício II com o modelo logit com desenho complexo – MG e BA.....	61
Tabela 26 - Resumo dos resultados do Exercício III com o modelo de probabilidade linear – MG e BA.....	62
Tabela 27 - Resumo dos resultados do Exercício III com o modelo logit – MG e BA.....	63
Tabela 28 - Resumo dos resultados do Exercício III com o modelo de probabilidade linear com desenho complexo – MG e BA.....	63
Tabela 29 - Resumo dos resultados do Exercício III com o modelo logit com desenho complexo – MG e BA.....	64

## **LISTA DE ABREVIATURAS E SIGLAS**

CEALE – Centro de Alfabetização, Leitura e Escrita da Faculdade de Educação da UFMG

INEP - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira

LDB – Lei de Diretrizes e Bases da Educação

MEC – Ministério da Educação

PNE – Plano Nacional de Educação

SAEB – Sistema de Avaliação da Educação Básica

SEB – Secretaria de Educação Básica

SEE – Secretaria Estadual de Educação de Minas Gerais

## SUMÁRIO

1 INTRODUÇÃO .....	13
2 A TRAJETÓRIA ESCOLAR E A EVOLUÇÃO DO ENSINO FUNDAMENTAL .....	15
2.1 A dinâmica do aprendizado e do desenvolvimento infantil .....	15
2.2 O ensino fundamental de nove anos .....	26
3 DADOS E METODOLOGIA .....	32
3.1 Modelos estatísticos utilizados .....	32
3.1.1 Método da dupla diferença (DD) ou <i>diferença-em-diferença</i> .....	32
3.1.2 Método de estimação logit .....	36
3.1.3 Método de estimação dos mínimos quadrados ordinários .....	38
3.1.4 Método de cálculo da Progressão por série .....	39
3.1.5 Desenho das amostras para os exercícios empíricos .....	40
3.1.5.1 Matrícula como variável de interesse .....	40
3.1.5.2 Probabilidade de progressão por série como variável de interesse .....	41
4 RESULTADOS .....	43
4.1 Descrição dos bancos de dados utilizados nos exercícios .....	43
4.1.1 Exercício I .....	43
4.1.2 Exercício II .....	45
4.2 Resultados dos modelos estatísticos .....	48
4.2.1 Minas Gerais como grupo de tratamento e São Paulo como grupo de controle .....	48
4.2.2 Minas Gerais como grupo de tratamento e Bahia como grupo de controle .....	56
5 CONSIDERAÇÕES FINAIS .....	64
REFERÊNCIAS BIBLIOGRÁFICAS .....	66

# 1 INTRODUÇÃO

Há muitos mecanismos de avaliação da educação, tanto em termos quantitativos quanto qualitativos. Nesta dissertação a taxa de matrícula e a probabilidade de progressão por série são os indicadores quantitativos escolhidos como *proxies* do grau de cobertura do sistema de ensino e do desempenho escolar do indivíduo para fornecer um pouco mais de informação sobre a educação no Brasil. Pretende-se com a análise destes indicadores contribuir para a avaliação do ensino fundamental com duração de nove anos e com essa avaliação iniciar a discussão a respeito da eficácia desta nova política educacional. Através do método estatístico *diferenças-em-diferenças* é feita a avaliação desta mudança no ensino fundamental brasileiro na tentativa de captar se houve algum avanço na trajetória escolar dos estudantes.

De acordo com as orientações do MEC, essa política tinha o objetivo de estreitar o hiato entre as crianças brasileiras no que diz respeito a anos de estudo completos e ao acesso à escola. O artigo 3º da LDB (lei 9.394 de 1996) estabelece como um dos princípios do ensino brasileiro que haja igualdade de condições para acesso e permanência na escola, mas a universalização do ensino fundamental não foi capaz de garantir o cumprimento deste princípio. A LDB já determinava que a duração do ensino fundamental era de no mínimo oito anos com a entrada aos sete anos de idade e que a entrada aos seis anos, aumentando a duração do ensino fundamental, era possível mas ficava a critério dos municípios implementá-la ou não. No entanto, a iniciação da vida escolar aos sete anos de idade foi se mostrando tardia quando os resultados obtidos com a educação infantil eram cada vez mais positivos e latentes. Os resultados obtidos nos anos iniciais do ensino fundamental pelas crianças que frequentaram a educação infantil eram, e ainda são, muito superiores aos das crianças que iniciavam o processo de educação formal já no ensino fundamental. No documento “Ensino Fundamental de nove anos – Orientações para a inclusão da criança de seis anos de idade”, o MEC ilustra isso com os resultados do SAEB de 2003 que apontam para o melhor desempenho do grupo que frequentou a educação infantil na competência de leitura. E é justamente o trabalho com a alfabetização e com o letramento o foco dos dois primeiros anos do ensino fundamental. As crianças que frequentam a educação infantil já entram no ensino fundamental com o processo de alfabetização e letramento bem avançados, e em alguns casos, já concluído. E aquelas que não tiveram essa oportunidade têm grandes dificuldades em se adaptar à rotina escolar e

desenvolver em pouco tempo a capacidade que as crianças que frequentaram a educação infantil desenvolveram com mais tempo.

Como a LDB assegura a educação infantil mas não a estabelece como prioridade, sua oferta estava muito restrita à iniciativa privada e conseqüentemente o acesso era muito limitado. Viu-se, então, no aumento da duração do ensino fundamental e na antecipação da entrada na escola a possibilidade de nivelar o grau de desenvolvimento das crianças que começavam o ensino fundamental e reduzir as diferenças de desempenho que tendem a aumentar à medida que a criança avança no sistema de ensino.

Essa política ainda é recente e é possível que seus resultados ainda sofram modificações ao longo do tempo, mas no período analisado aqui os resultados apontam para ineficácia do aumento da duração do ensino fundamental em Minas Gerais. O foco do trabalho é quantitativo e mesmo considerando-se as limitações deste tipo de estudo percebe-se que esta política educacional carece de ajustes para que se obtenha um resultado positivo.

Além desta introdução, o trabalho é formado por quatro capítulos. O capítulo dois trata da literatura que discute os determinantes do desempenho escolar e da literatura sobre o panorama do ensino fundamental brasileiro. O capítulo três trata das ferramentas estatísticas usadas nos exercícios empíricos. O capítulo quatro discute o resultado destes exercícios e o quinto capítulo estabelece um cotejo entre a literatura revisitada neste trabalho e os resultados obtidos.



## 2 A TRAJETÓRIA ESCOLAR E A EVOLUÇÃO DO ENSINO FUNDAMENTAL

### 2.1 A dinâmica do aprendizado e do desenvolvimento infantil

O estudo da trajetória escolar trata, na maioria dos casos, da análise da transição entre as séries ou anos-ciclo que compõem o sistema de ensino. Esta é uma forma muito simplista de analisar a educação e deixa de lado aspectos cruciais do desenvolvimento do indivíduo que têm grande influência sobre o desempenho escolar. Ignorar estes aspectos significa mitigar o retorno dos investimentos em educação e afetar negativamente a inserção destes indivíduos na economia e na sociedade. Neste capítulo são discutidos aspectos do desempenho escolar que vão além das notas e da aprovação nas séries ou nos ciclos.

A literatura nas áreas da economia e da demografia que contempla estudos sobre educação e capital humano normalmente se atém ao estudo de habilidades mensuráveis e captadas em testes e provas. Estas, denominadas habilidades cognitivas, estão ligadas à inteligência inata e podem ser analisadas através de testes de QI (Quociente de Inteligência) e de modo limitado também através de testes e provas que medem o grau de conhecimento formal adquirido em áreas específicas. No entanto, estas não são as únicas habilidades importantes para a formação do indivíduo e para sua inserção na sociedade e no mercado de trabalho. As chamadas habilidades não-cognitivas tais como paciência, auto controle, temperamento, motivação e perseverança são tão importantes quanto as primeiras. Segundo Heckman e Cunha (2006) “psicólogos especializados em desenvolvimento infantil defendem a importância de entender como as habilidades não cognitivas são formadas para analisar o impacto das políticas públicas voltadas para a educação infantil.” (p.1) Mas sua mensuração é extremamente difícil e controversa. Por esse motivo elas são negligenciadas e sua influência é tratada como algo secundário. Cunha *et all* (2005) analisaram a formação destas habilidades ao longo da vida dos indivíduos e identificaram o princípio da *self-productivity*. De acordo com esse princípio, algumas habilidades são fundamentais na formação das outras em um período posterior do

ciclo de vida individual, ou seja, uma ajuda a “criar” a outra. Dessa forma, como é possível negligenciar as habilidades não-cognitivas? Simplesmente, isso não é possível.

“Noncognitive abilities also matter for success both in the labor market and in schooling. Even when early childhood interventions do not boost IQ, they improve noncognitive skills, with substantial effects on schooling, labor market outcomes, and behavioral outcomes such as teenage pregnancy and participation in criminal activities. They raise achievement test scores, which can be influenced by schooling (and other inputs), even when they do not boost IQ. In light of this evidence, the neglect of noncognitive ability in evaluating human capital interventions, and in formalizing the skill formation process, is unwarranted.” (Cunha *et al*, 2005, p.16)

Além de se alimentarem<sup>1</sup>, as habilidades são determinadas pelo fator genético. Contudo, a maneira através da qual esse componente genético das habilidades cognitivas e das não-cognitivas se manifesta e se desenvolve é influenciado pelo ambiente familiar e pelo *background* familiar. Estes podem ser chamados de “fatores de risco” do desenvolvimento das habilidades. Os principais fatores de risco são fazer parte de família de baixa renda, fazer parte de família mono parental, fazer parte de alguma etnia que é minoria na localidade onde a criança mora, ter pais com baixos níveis de educação formal e de habilidades, receber pouco estímulo cognitivo e emocional dos familiares e ter pais com habilidades laborais e intelectuais particularmente baixas.

“The quantitatively important constraints facing disadvantaged children are the ones determining their early environment’s parental background, and the like. The empirically important market failure in the life cycle of child skill formation is the inability of children to buy their parents or the lifetime resources that parents provide, and not the inability of families to secure loans for a child’s education when the child is an adolescent. Our analysis has major implications for the way policies should be designed in order to help low income and disadvantaged populations. Evidence from disadvantaged populations demonstrates that enriched early interventions can raise measured ability and other skills.” (Cunha *et al*, 2005, p.6)

“(…) a forma como os pais lidam com os filhos, do ponto de vista social, emocional e cognitivo é determinante de seu sucesso escolar. E, por sua vez, é fortemente condicionada pelo nível sociocultural dos mesmos e das práticas educativas consideradas como culturalmente válidas nos vários segmentos da sociedade.” (Oliveira & Silva, 2006, p.6)

Todos estes fatores de risco, além de impactarem na formação das habilidades, aumentam a probabilidade da criança deixar a escola, de ser reprovada, de ter notas baixas ou de ter

---

<sup>1</sup> Heckman e Cunha (2006) afirmam que as habilidades não cognitivas influenciam o acúmulo das habilidades cognitivas em períodos posteriores mas a recíproca não é verdadeira. Uma influencia a outra, mas este efeito não é observado em qualquer período do tempo.

problemas de comportamento na escola. Além disso, altas taxas de repetência nas primeiras séries aumentam as chances de que a criança tenha baixo rendimento no futuro, pois geralmente o estudante fica traumatizado e estigmatizado. Uma possível maneira de evitar que isso aconteça é investir diretamente nas famílias com ambiente adverso. Poderia ser muito proveitoso criar mecanismos e políticas públicas para diminuir o impacto do ambiente familiar adverso no qual muitas crianças estão inseridas. Essas políticas poderiam ter impactos futuros significativos já que as habilidades consolidadas nos primeiros anos de vida influenciam na probabilidade de praticar crimes, de ter filhos durante a adolescência, de conseguir empregos com boa remuneração e influenciam também no retorno obtido por cada ano de estudo completo. Esta ideia é um complemento às políticas que investem somente na construção de mais escolas, na formação de professores e na compra de materiais didáticos. Embora em muitos casos este tipo de política seja extremamente necessário devido a grande insuficiência de oferta de vagas na educação e a baixa qualidade do ensino ofertado, não é suficiente para sanar o problema de baixo desempenho nas primeiras séries do ensino formal.

"Evidências substanciais de estudos econômicos, sociológicos e de políticas públicas sugerem que crianças de famílias com certas desvantagens têm mais chance de cometer crimes, ter filhos fora do casamento e de deixar a escola. Intervenções feitas mais cedo e que parcialmente remediam os efeitos dos ambientes adversos na infância podem reverter alguns danos causados pelas famílias com certas desvantagens e ter um retorno econômico alto relativamente a outras políticas. Eles beneficiarão não só as próprias crianças, mas também os filhos das mesmas, assim como a sociedade de modo geral." (Heckman & Masterov, 2005, p.4)

Na grande maioria dos casos, crianças que moram com ambos os pais têm melhor desempenho escolar do que aquelas que moram com um só deles porque o tempo que um só dispõe para ajudar o filho nas tarefas escolares e para acompanhar seu desenvolvimento social é insuficiente (esta deficiência seria suprida pelo outro) e a renda de um só normalmente é menor do que a renda de um casal o que diminui substancialmente o investimento na educação dos filhos. Mas vale ressaltar que, famílias com restrição orçamentária não só terão dificuldade de custear a faculdade dos filhos como também apresentam déficit nos investimentos em educação ao longo de toda a vida deles. Não é adequado usar o argumento simplista de que o indivíduo cuja família é de baixa renda não consegue alto nível de escolaridade e tampouco bons empregos porque a família não pode custear o nível superior. Este indivíduo está em desvantagem desde a infância quando não teve suas habilidades e capacidades estimuladas. Mas isso só fica claro quando ele não consegue sucesso profissional

e por isso a análise deste problema normalmente se limita a vida adulta. É importante olhar mais para trás e perceber que a falta de investimento começou cedo e não é só a questão financeira, mas também a falta de estímulo das habilidades e capacidades.

“Children whose parents have higher incomes have access to better-quality primary and secondary schools. Children’s tastes for education and their expectations about their life chances are shaped by those of their parents. Educated parents are better able to develop scholastic aptitude in their children by assisting and directing their studies.” (Cunha *et al*, 2005, p.28)

Com os dados obtidos com o teste PIAT nos EUA para as crianças de 4 a 12 anos, Cunha *et al* perceberam um significativo hiato entre o desempenho escolar de indivíduos de diferentes grupos de renda e também grandes hiatos por grupos de raça (brancos, hispânicos e negros). Mas esses hiatos são drasticamente reduzidos quando os dados são controlados por escolaridade e habilidade cognitiva da mãe e pela estrutura familiar. Nota-se que os hiatos que surgem entre as crianças de diferentes grupos de renda aparecem cedo, mas estes hiatos podem ser eliminados se forem oferecidos meios de melhorar o ambiente adverso no qual algumas crianças estão inseridas.

“Differences in levels of cognitive and noncognitive skills by family income and family background emerge early and persist. If anything, schooling widens these early differences, but the main gaps in these skills that are found in adulthood emerge before schooling begins.” (Cunha *et al*, 2005, p.16)

Heckman e Cunha (2006) esmiúçam a influência do ambiente e do *background* familiar sobre o desempenho futuro dos indivíduos. Em seu trabalho eles contemplam tanto as habilidades cognitivas quanto as não-cognitivas dos indivíduos já que estudos que relacionam notas de testes, como *proxies* de habilidades cognitivas, com variáveis de *background* familiar e outras variáveis socioeconômicas apresentam muitos problemas em seus resultados devido à endogeneidade das variáveis independentes usadas para explicar as habilidades cognitivas. Eles perceberam que as características socioeconômicas, ou *inputs*, dos pais afetam as habilidades cognitivas de modo mais significativo nas idades iniciais (nas idades iniciais da infância este tipo de habilidade normalmente se desenvolve mais) e mais tarde estes *inputs* afetam mais as habilidades não cognitivas (nas idades posteriores as habilidades não cognitivas normalmente se desenvolvem mais). Além desta diferença de impacto em termos temporais, a magnitude do impacto do ambiente familiar sobre os dois tipos de habilidade é diferente: as habilidades não cognitivas são mais afetadas pelas variáveis que captam o

investimento dos pais nas crianças do que as habilidades cognitivas. Mas o mais importante é que a influência dos pais perdura ao longo da vida escolar da criança independente da magnitude deste impacto e do tipo de habilidade que é influenciada.

Para entender melhor esta diferença do impacto do ambiente e do *background* familiar sobre o desenvolvimento das habilidades dos indivíduos é importante incorporar na discussão o conceito de período crítico de aprendizado. Costuma-se definir, na literatura sobre a trajetória escolar, um período crítico de aprendizado na vida da criança que vai de três a sete anos de idade. Neste período a formação cognitiva e social das crianças está no seu momento mais importante e todo tipo de influência é extremamente significativo. Nos primeiros anos de vida esse desenvolvimento cognitivo se dá com muita rapidez e o indivíduo apresenta os mais elevados níveis de aproveitamento daquilo que é aprendido. A memória é muito mais duradoura e a absorção do conhecimento é muito mais rápida e eficaz. A discussão sobre este período crítico é vasta, mas a idade de seis a sete anos é vista como crucial.

“(…) Por volta de 6 anos todos os mecanismos básicos para aprender já estão bem estabelecidos. Também por volta do 6º aniversário começam a se consolidar as áreas de associação terciárias na córtex pré-frontal. Esse cabeamento neuronal permite à criança maior autocontrole e menor dependência do córtex visual.” (Oliveira & Silva, 2006, p.12)

Cunha *et all* (2005) apontam para o fato de que a infância, mais do que um único período crítico, possui mais de um estágio e isso é de extrema importância na construção de políticas públicas educacionais. Independentemente de qual seja o período crítico do aprendizado na infância, esta não pode ser vista como um período único e de características imutáveis no que diz respeito ao aprendizado da criança. Estes estágios são identificados por Heckman e Cunha (2006): “(…) with our estimated technology, it is possible to define and measure critical and sensitive periods in the life cycle of child development, and to determine which inputs at which ages most affect the evolution of skills.” (p.3) Através da função que estima e formaliza a produção tecnológica das habilidades cognitivas os autores percebem que a idade crítica na qual o investimento dos pais nas crianças tem mais influência sobre as habilidades cognitivas é entre seis e sete anos ou entre oito e nove anos. E a idade crítica na qual o investimento dos pais nas crianças tem mais influência sobre as habilidades não cognitivas é entre oito e nove anos ou entre dez e onze anos. As habilidades cognitivas, medidas através do coeficiente de inteligência, não mudam quase nada depois dos dez anos de idade. De acordo

com os autores, o QI de um indivíduo com dez anos de idade é muito próximo do QI observados nas idades posteriores. Mas as habilidades não cognitivas ainda podem ser estimuladas depois desta idade obtendo-se um bom retorno do estímulo.

“(...) Quem é favorecido nesse período tem aumentadas as chances de colher mais e mais benefícios ao longo da vida escolar. Quem fracassa e/ou experimenta dificuldades, começa a se envergonhar perante seus pares, começa a duvidar de suas próprias capacidades cognitivas, passa por uma situação de queda da auto-estima, tende a evitar livros, se desinteressa de leitura, inclina-se a se afastar do mundo dos valores da cultura letrada, atrasa-se no desenvolvimento vocabular, lê menos, (...) aprende menos. Desse modo, entra numa espiral de fracassos que pavimenta uma larga estrada para o insucesso pessoal, social e econômico. (...) sem o efetivo desenvolvimento e sem a posse e o incremento de tais habilidades na fase inicial de escolarização, os esforços para dominar o instrumento cultural do manejo da língua escrita e a agregação de valor educacional ao longo da vida escolar que esse domínio propicia tornam-se crescentemente penosos, dispendiosos e/ou infrutíferos. Ao contrário, o domínio e o incremento desse conjunto de competências na fase infantil, em especial da consciência fonológica e da consciência fonêmica, favorecem e facilitam a agregação posterior de valores educacionais. E diminuem os custos e as penas dessa agregação.” (Oliveira & Silva, 2006, p.2)

Mesmo reconhecendo a pluralidade dos determinantes do desempenho escolar e a grande influência do ambiente e do *background* familiar sabe-se que dirimir os diferenciais socioeconômicos e as diferenças deste *background* é algo que beira o impossível atualmente. Por isso outros aspectos também devem ser considerados no estudo dos determinantes do desempenho escolar para maximizar o retorno das políticas públicas de melhoria deste desempenho. Um dos aspectos é a influência do ambiente escolar. No caso do Brasil, os recursos físicos escassos das escolas e a forma de promover ou reprovar a criança (recuperação, retenção, formação de classes de acordo com notas e índices de aproveitamento dos estudantes) ainda afetam muito as crianças por toda vida escolar. A educação infantil é uma maneira interessante de preparar a criança para esta etapa. Ela antecipa o contato com letras e números e favorece o desenvolvimento cognitivo além de fazer com que as crianças se adaptem mais cedo à rotina escolar e suas regras. O objetivo não é antecipar a educação formal ou a aquisição de conhecimento formal. A ideia é fazer com que a criança desenvolva as suas habilidades cognitivas, adquira disciplina e conduta compatíveis com a convivência em sociedade e mais especificamente com o ambiente escolar. Na educação infantil a criança cria identidade própria, desenvolve sua independência e começa a perceber suas capacidades. Ela começa a desenvolver a linguagem e a comunicação, sua capacidade motora, sua percepção de espaço, as relações sociais e as regras que norteiam as mesmas. O indivíduo privado disso muito provavelmente terá desempenho muito aquém do desejado na escola

quando comparado ao indivíduo que teve acesso a educação infantil. A educação infantil é uma ponte entre o ambiente familiar e a educação formal. O ideal é evitar choques vividos pelas crianças durante períodos de transição dentro da sua vida escolar, e o maior deles é quando a criança deixa o ambiente familiar para se “aventurar” no ambiente escolar. A iniciação na educação formal é muito mais tranquila para a criança que frequentou a educação infantil. O conceito de educação básica englobando educação infantil e ensino fundamental deveria ser mais difundido e aplicado.

(...) A Educação infantil pode beneficiar as crianças por duas razões principais. A primeira, derivada da perspectiva cognitiva, é que frequentar instituições de educação infantil, pode ensinar às crianças habilidades cognitivas ou potencializar seu desenvolvimento cognitivo. Por exemplo, crianças que aprendem letras e sons nas instituições de educação infantil aprenderiam a ler mais facilmente na primeira série. A segunda perspectiva se desenha na abordagem do curso da vida. Ela teoriza que, ao proporcionar uma socialização mais antecipada, as instituições de educação infantil poderiam ajudar as crianças a se ajustar ao papel de estudantes. Para os estudantes de ensino fundamental que frequentaram instituições de educação infantil, a primeira série não é uma experiência ‘nova’ dado que eles já estão de algum modo acostumados a estar separados de casa, a trabalhar em grupo e a seguir instruções. (Doris & Alexander, 1998, p.358-359)

Além disso, nas primeiras séries a criança cria um laço afetivo com a escola e com seus integrantes facilitando o aprendizado e estimulando a frequência. A figura do professor que estimula e elogia ajuda a criança a se esforçar cada vez mais para se igualar ao colega e ter o seu esforço reconhecido pelos demais. Respeitar regras, desenvolver habilidades de estímulo cognitivo, receber novos conhecimentos e conviver com outras crianças estimula cada criança a crescer e a desenvolver habilidades que facilitam sua inserção e atuação em todas as esferas da sociedade. Por tudo isso, esse período de iniciação na vida escolar é tão crucial.

“(…), crianças nas primeiras séries são otimistas, interessadas em aprender e tendem a formar laços afetivos com seus professores. Todos esses fatores intensificam sua receptividade à influência escolar e estão notavelmente ausentes nas séries mais elevadas.” (Doris & Alexander, 1993, p.409)

À medida que a criança avança em idade e no sistema de ensino ela perde seu entusiasmo pela escola – que cai progressivamente ano após ano – e ela tende a se auto-avaliar de forma cada vez menos positiva. Por isso, conseguir um bom desempenho nas primeiras séries facilita a progressão nas séries subsequentes. Neste caso, a criança estaria mais preparada para o aprofundamento e para o mais elevado grau de dificuldade do conteúdo das séries posteriores.

Ademais, as diferenças de desempenho que surgem entre os estudantes logo nas primeiras séries, em termos de nota e avaliação, crescem ao longo dos anos e podem culminar em repetência ou rendimento muito abaixo daquele considerado razoável.

“More motivated children are more likely to stay in school and have higher achievement tests. Our analysis suggests that social policy should be more active in attempting to alter non cognitive traits, including values, especially for children from disadvantaged environments who receive poor discipline and little encouragement at home. This more active social policy approach would include mentoring programs and stricter enforcement of discipline in the schools. Although these programs are controversial, they are likely to be effective and to produce substantial saving to society from reduced pathological behavior.” (Cunha *et al*, 2005, p.22)

Retomando a discussão sobre a influência do ambiente e do *background familiar*, de acordo com Oliveira e Silva (2006), no caso do processo de alfabetização, que ocorre normalmente quando as crianças são iniciadas na vida escolar, as características da criança e as características da família são os principais determinantes do desempenho escolar. Nesse caso o QI da criança não é o mais importante e sim o trabalho com a linguística. De acordo com Oliveira e Silva (2006), crianças de nível socioeconômico mais baixo tendem a apresentar mais dificuldade na alfabetização pela falta de estímulo de conversação e de enriquecimento de vocabulário e pela falta de acesso a livros e jogos educativos que estimulam essa parte linguística. A qualidade do diálogo que os pais estabelecem com as crianças é fundamental para o processo de alfabetização. Os pais de nível socioeconômico mais elevado utilizam frases mais complexas e estimulam mais que a criança estabeleça diálogos e conversações mais longos e mais complexos em termos de vocabulário e de construção gramatical. A parte fonológica também é crucial nessa fase e os problemas relacionados a ela podem ser detectados com muita facilidade quando a criança não responde aos estímulos citados acima. Mas quando a criança sequer tem esse tipo de estímulo, seja em casa ou na escola, a detecção precoce deste tipo de problema não é possível. A educação infantil conseguiria abrandar o impacto negativo de ambientes familiares desfavoráveis e diminuir o hiato entre as crianças que aparece antes mesmo de iniciarem a vida escolar.

Ngaruiya (2006) aborda um conceito muito interessante ao destacar que a criança deve estar pronta para a escola e a escola deve estar pronta para a criança. No caso da instituição, ele está falando de assegurar infraestrutura adequada, profissionais qualificados e bem treinados e projetos pedagógicos sólidos e robustos que garantam a qualidade do ensino ofertado. No



caso da criança, a educação infantil cumpre esta tarefa de preparar a criança para a educação formal. É importante que a concepção do currículo da educação infantil esteja relacionada com o currículo do ensino fundamental.

Os diagnósticos precoces aqui apontados não são os únicos benefícios da educação infantil. Barros e Mendonça (1999) fizeram um estudo sobre a educação pré-escolar no Brasil que discorre sobre o estado nutricional das crianças, sobre o desempenho escolar das mesmas na educação formal e sobre o desempenho no mercado de trabalho. Os autores utilizaram a Pesquisa sobre Padrões de Vida do IBGE (PPV) de 1996 e 1997 e mostraram que o fato de frequentar a pré-escola diminui a probabilidade de que a criança tenha baixa estatura para a sua idade e diminui a probabilidade de que o seu peso esteja abaixo daquele considerado ideal para sua altura. Eles perceberam também que cada ano adicional de pré-escola aumenta entre 0,4 e 0,6 anos a escolaridade do indivíduo, aumenta a probabilidade de completar qualquer série e reduz a repetência entre três e cinco pontos percentuais em média.

Observa-se, portanto, que investir na educação das crianças desde cedo pode ser muito mais barato e eficaz do que investir mais tarde quando já há problemas sociais e educacionais tais como repetência, notas baixas, pouca frequência, etc. entre os jovens. Para comprovar isso, Cunha *et al* (2005) analisam um estudo feito no Reino Unido por Michael Rutter, alguns colegas e o *English and Romanian Adoptees Study Team*. Eles acompanharam o desenvolvimento físico e intelectual de um grupo de 165 crianças romenas adotadas por britânicos e 52 crianças britânicas adotadas também por britânicos. As crianças romenas que foram alvo do estudo nasceram em uma época difícil no seu país, pois uma política pró-natalidade mal sucedida superlotou orfanatos com crianças abandonadas que sofreram por falta de cuidados médicos, falta de alimentação adequada e falta de carinho e afeto. A análise foi feita quando as crianças alcançaram 4 e 6 anos. Em ambas as idades foi observado que as crianças romenas adotadas que conseguiram alcançar o mesmo nível de desenvolvimento físico e intelectual das crianças britânicas adotadas foram aquelas cuja adoção se deu antes dos seis meses de idade. As que foram adotadas depois desta idade não alcançaram o mesmo nível de desenvolvimento das britânicas. Isso mostra que intervenções feitas mais cedo são muito mais eficazes.

“The main finding from the recent literature is that we should prioritize, and shift our priorities, in a marginal fashion by redirecting a given total sum of expenditure

on skill investment to earlier ages relative to how it is currently allocated toward disadvantaged populations that do not provide enriched environments for their children.” (Cunha *et al*, 2005, p.87)

Outro exemplo de que a intervenção precoce é bem sucedida é o *Perry Preschool Program*. Nesse programa o público alvo eram crianças norte-americanas negras, com idade entre três e quatro anos, cuja família apresentava vulnerabilidade social devido a pobreza e que tinham grande probabilidade de apresentar desenvolvimento escolar insatisfatório. Estas crianças frequentaram aulas cinco vezes por semana em classes de 20 a 25 indivíduos com quatro professoras desenvolvendo atividades com elas (um professor para cada oito crianças). Uma vez por semana a família era visitada por uma professora do programa para intensificar o trabalho feito nas aulas e envolver os pais nas atividades e no desenvolvimento da criança. O programa foi executado de 1962 a 1967 sendo a mensuração e coleta de dados foi feita anualmente quando as crianças tinham de 3 a 11 anos e ao completarem as seguintes idades: 14, 15, 19, 27 e 40 anos. Cento e vinte e três crianças foram selecionadas para fazer parte do estudo e, aleatoriamente, elas foram divididas em dois grupos: controle e tratamento (58 crianças ficaram neste grupo). O grupo de tratamento apresentou melhor preparação para o início da educação formal; apresentou melhor desempenho escolar, a maioria dos integrantes deste grupo tinha mais responsabilidade e comprometimento com a escola ao completar 14 anos de idade, apresentou maior percentual de conclusão do ensino médio, auferiu maiores rendimentos quando chegaram a idade de 40 anos, apresentou maior índice de empregabilidade aos 27 e aos 40 anos e o percentual de indivíduos presos mais de cinco vezes até os 40 anos foi bem inferior se comparado ao grupo de controle. Além disso, até os pais dos indivíduos do grupo de tratamento apresentaram mais comprometimento com a escolaridade dos seus filhos quando comparados aos pais do outro grupo. Com relação à qualidade de vida, no que diz respeito ao poder de compra, o grupo de tratamento tinha maior percentual de indivíduos que possuíam casa própria e carro. Em suma, o programa apresentou grande retorno social ao longo do tempo melhorando a inserção social e econômica e o comportamento dos indivíduos do grupo de tratamento trouxe benefícios não só para eles mesmos mas também para a sociedade de modo geral. Heckman e Cunha (2006) afirmam que o *Perry School Program* não elevou o QI das crianças, mas suas chances de se tornarem bem sucedidas econômica e socialmente foram muito maiores do que as das crianças que não foram contempladas pelo programa. De forma bem concisa eles pautam que “(...) Skill begets

skill. Later remediation of early skill deficits can be costly. (...) returns to investment are higher for the young and the disadvantaged". (p.62).

Para suprir as carências de um ambiente familiar desfavorável e conseguir retornos semelhantes aos do *Perry School Program*, a educação infantil deve ser de alta qualidade. Não basta frequentar qualquer tipo de pré-escola. E a entrada deve se dar o mais cedo possível para que diagnósticos precoces de todos os tipos de deficiências reduzam ao máximo ou até mesmo eliminem os hiatos existentes entre as crianças de famílias de situação socioeconômica desfavorável e as crianças de nível socioeconômico elevado como já foi discutido anteriormente. Os retornos dos investimentos feitos na infância são altos e os custos para remediar, mais tarde, a falta destes investimentos na infância são altíssimos. No entanto, vale salientar que investimentos em educação jamais devem ser feitos pontualmente. Esse é o princípio da *complementarity* de Cunha *et all* (2005). Investimento nas idades iniciais deve ser seguido de mais investimento nas idades posteriores, ou seja, o investimento na primeira infância é crucial, mas não é suficiente, ele aumenta muito o retorno dos investimentos futuros, mas sozinho não consegue garantir que os indivíduos estejam preparados para a vida profissional. Se os investimentos futuros não acontecerem os efeitos dos investimentos empregados durante a pré-escola desaparecerão.

Na próxima seção é apresentada a nova política educacional que está diretamente relacionada à educação infantil e ao aproveitamento do melhor período para o desenvolvimento cognitivo das crianças. O aumento da duração do ensino fundamental é uma antecipação da entrada na escola já que não há oferta de educação infantil para todas as crianças. É uma tentativa de diminuir a diferença entre aqueles que entram no ensino fundamental depois de frequentar a educação infantil e aqueles que tem o primeiro contato com ambiente escolar só no ensino fundamental. Este trabalho estimulará a discussão da eficácia desta política.

## 2.2 O ensino fundamental de nove anos

Em 1872 foi realizado o primeiro censo brasileiro e no mesmo já havia dados educacionais. Neste ano, o percentual de pessoas de cinco anos e mais alfabetizadas era de 17,7%. Em 1920 este percentual subiu para 28,8%, em 1940 subiu para 38,9%, mas caiu para 32,9% em 1950. Nos censos de 1960 a 2000 a proporção de alfabetizados superou a de analfabetos alcançando os seguintes percentuais da população de cinco anos e mais: 53,3%, 61,3%, 68,1%, 75,8% e 83,3%. Um dos aspectos responsáveis pela queda no número de analfabetos foi a expansão da cobertura escolar, especialmente no ensino fundamental.

O resultado da expansão do ensino fundamental de modo pouco criterioso como aconteceu no Brasil foi a queda na qualidade do ensino e o grande impacto negativo no desempenho dos estudantes. Dados do SAEB (Sistema de Avaliação da Educação Básica) revelam que, em 2001, somente 4,48% dos estudantes da quarta série do ensino fundamental possuíam nível de leitura compatível com a série na qual estavam matriculados, 36,2% possuíam nível intermediário de leitura, 37% estavam no nível crítico e 22% estavam no nível muito crítico. Estas classificações foram definidas pelo SAEB, sendo que a situação crítica é caracterizada pela leitura de frases pequenas e simples e a situação muito crítica é caracterizada pela ausência de habilidades de leitura, estágio no qual as crianças não conseguem sequer responder às questões do questionário de Avaliação do SAEB (CEALE, 2004).

Isto mostra que o processo de alfabetização no Brasil ainda apresenta falhas e que o processo de letramento está ainda mais distante de contemplar 100% da população de cinco anos e mais. A evolução da cultura e do mercado de trabalho fez com que a alfabetização se tornasse um conceito insuficiente e, então, surgiu o conceito de letramento que vai além da alfabetização e melhor se adequa à evolução da educação. O processo de letramento engloba a questão do conhecimento, da cultura e da capacidade de utilizar a língua escrita em vários cenários sociais, tais como comunicação, trabalho acesso à informação, etc.

“A *alfabetização* é a aquisição do código da escrita e da leitura, a codificação através da escrita e a decodificação através da leitura; em suma, em seu sentido mais restrito, alfabetizar-se é aprender a ler e a escrever.

O termo *letramento* é o uso das práticas sociais da leitura e da escrita e difundiu-se rapidamente no meio acadêmico; porém, anteriormente, transitou pela mídia e nas escolas na tentativa de produzir algum sentido para além do termo alfabetização, que já não era suficiente para explicar o processo de aquisição do código escrito.

(...)

As atividades de alfabetização são aquelas de descoberta e automação do sistema alfabético de escrita, relacionando a automatização das relações grafemas-fonemas, ou seja, relacionadas à aquisição da base alfabética. E as atividades de letramento são aquelas onde existe uma interação com o material escrito, lendo e escrevendo diferentes gêneros em variados suportes, para diferentes interlocutores, isto é, uma preocupação com os usos sociais da leitura e da escrita.” (Filgueiras, 2004, p. 4-5)

Dois fatores são usualmente apontados como causas do fracasso no processo de alfabetização: a implantação do sistema de ciclos junto com a adoção do sistema de progressão continuada e a aplicação de métodos ultrapassados e inadequados de alfabetização (CEALE, 2004). Dados do SAEB mostram que boa parte dos estudantes não está matriculada sob o sistema de ciclos e sim sob o sistema de séries o que diminui a responsabilidade da implantação deste sistema no fracasso do processo de alfabetização. Ainda assim, a implantação dos mesmos desvirtua a comunicação entre professores, as propostas curriculares e os planos pedagógicos dificultando a continuidade e a fluidez das atividades escolares. Mas o sistema de ciclos pode ser bem sucedido na fase de alfabetização já que muitas vezes o período de um ano é insuficiente para alfabetizar uma criança que acabou de entrar na escola e a retenção não seria benéfica para ela. Na década de 1980, por exemplo, “(...) de cada mil crianças que iniciavam a 1ª série, menos da metade chegava à 2ª; menos de um terço conseguia concluir este nível de ensino. Quer dizer: não havia analfabetos na escola ou sua presença não era percebida porque, em função justamente da repetência, eles se concentravam na 1ª série, ou porque, em razão da evasão, eles abandonavam a escola (ou, dependendo do ponto de vista, dela eram expulsos).” (CEALE, 2004, p.18) Neste caso a promoção criteriosa deveria ser feita com o acompanhamento desta criança que não alcançou o nível de aproveitamento desejado de modo que ela alcance o nível dos colegas. No entanto, não é isso que acontece.

De modo geral, o processo de alfabetização e letramento começa mais tarde para as crianças de camadas sociais menos favorecidas devido ao restrito acesso à educação infantil. Iniciar o ensino fundamental aos seis anos pode ser uma forma de fazer com que as crianças que não frequentam nem a creche nem a pré-escola tenham acesso a pelo menos um ano de estudo dedicado ao processo de alfabetização.

As atuais exigências de democratização do acesso à escola pública de qualidade se acompanham de demandas bastante complexas: a permanência das crianças de camadas populares nessa escola e ampliação de suas oportunidades de acesso à cultura escrita, pois tais oportunidades já são precocemente vivenciadas por camadas sociais mais favorecidas. Isso implica o direito daquelas crianças à alfabetização e ao letramento, em processos de aprendizagem que assegurem progressivas capacidades e habilidades (...). (CEALE, 2004, p.21)

“(...) No Brasil, as crianças de classe média e as crianças que frequentam escolas privadas concluem seu processo de alfabetização ao final de 6 anos, antes de ingressar na 1ª. série do Ensino Fundamental. Proposta apresentada ao país pelo grupo intitulado Todos pela Educação preconiza como meta que a alfabetização esteja concluída ao final dos 8 anos (3ª série do E.F. de 9 anos). Esse tipo de proposta possivelmente reflete uma concepção de alfabetização diferente da que é adotada em outros países. Além disso, ao tomar como aceitável ou desejável essa meta, estabelece como aceitável uma inequidade de acesso. Dada a cumulatividade da aprendizagem, especialmente a aprendizagem de vocabulário por intermédio de leitura, o retardo para concluir a etapa da alfabetização também implica uma inequidade de progresso.”(Oliveira & Silva, 2006, p.19)

Os problemas no processo de letramento levaram ao aumento do número de anos de estudo obrigatórios para concluir o ensino fundamental. Esta ampliação já estava prevista na lei nº 9.394 de 1996, a LDB, e é uma das metas do PNE (Plano Nacional de Educação): garantir que todas as crianças de seis anos, independente da classe social, estejam matriculadas na escola. O objetivo desta ampliação é aumentar a qualidade da educação aumentando a taxa de matrícula das crianças de seis anos, diminuir a vulnerabilidade destas crianças a situações de risco, diminuir o abandono escolar, aumentar o sucesso no processo de aprendizado e aumentar a escolaridade média. Na lei que estabelece o PNE, lei nº 10.172 de 9 de janeiro de 2001, é apresentada a proposta de “ampliar para nove anos a duração do Ensino Fundamental obrigatório, com início aos seis anos de idade, à medida que for sendo universalizado o atendimento na faixa etária de 7 a 14 anos” com o objetivo de “oferecer maiores oportunidades de aprendizagem no período de escolarização obrigatória e assegurar que, ingressando mais cedo no sistema de ensino, as crianças prossigam nos estudos, alcançando maior nível de escolaridade.”

O novo ensino fundamental é dividido em duas etapas, os anos iniciais e os anos finais. A primeira etapa é composta de dois ciclos: o Ciclo Inicial de Alfabetização (CIA) e o Ciclo Complementar de Alfabetização (CCA). O primeiro atende as crianças de seis a oito anos e o segundo atende as crianças de nove e dez anos. A segunda etapa é composta de um ciclo de duração de quatro anos que atende as crianças de 11 a 14 anos.

Em agosto 2003, o governo de Minas Gerais instituiu o ensino fundamental de nove anos de duração nas escolas da rede estadual de ensino do estado, com matrícula a partir dos seis anos de idade, através do Decreto nº 43.506. Neste mesmo decreto o governo estadual incumbiu a SEE (Secretaria Estadual de Educação de Minas Gerais) de definir e divulgar as normas para a organização do novo ensino fundamental. Em dezembro do mesmo ano, a SEE publicou a Resolução nº 469 cumprindo a determinação do governo estadual. Com base nesta resolução, o ensino fundamental passa a ser organizado de acordo com o quadro 1.

Etapa de ensino	Faixa etária prevista	Duração
Educação infantil	Até 5 anos de idade	
Creche	Até 3 anos de idade	
Pré-escola	De 4 a 5 anos de idade	
Ensino Fundamental:	Até 14 anos de idade	9 anos
*Anos Iniciais:	De 6 a 10 anos de idade	5 anos
Ciclo Inicial de Alfabetização	De seis a 8 anos de idade	3 anos
Ciclo Complementar de Alfabetização	De nove a dez anos de idade	2 anos
*Anos Finais	De 11 a 14 anos de idade	4 anos

**Quadro 1: Nova organização do ensino fundamental**

Inicialmente, no ano de implantação do ensino fundamental de nove anos em Minas Gerais – 2004 – ,o CIA foi composto por três fases: a fase introdutória era aquela destinada aos estudantes que ingressassem no ensino fundamental aos seis anos, completos até 30 de abril de 2004 e aos estudantes que completassem sete anos de idade no período de 1º de agosto a 31 de dezembro de 2004; a Fase I era destinada aos estudantes provenientes da Fase Introdutória e a Fase II era destinada aos estudantes provenientes da Fase I. Os estudantes que completassem sete anos no período compreendido entre 1º de agosto e 31 de dezembro de 2004 podiam ser matriculados na Fase I, desde que evidenciassem domínio dos objetivos da Fase Introdutória, em função de sua escolaridade ou experiências anteriores. Já o CCA, que compreende dois anos, dá seguimento ao Ciclo Inicial e tem como objetivo a consolidação, ampliação e aprofundamento dos conhecimentos e capacidades consideradas essenciais ao processo de alfabetização e letramento dos estudantes, compreendendo duas fases: a Fase III, para os concluintes do Ciclo Inicial e a Fase IV, destinada aos estudantes que alcançarem os objetivos da Fase III. Os estudantes que já haviam ingressado no ensino fundamental antes da mudança no mesmo e os estudantes que completassem sete anos até 31 de julho de 2004 podiam concluí-lo em oito anos.

Em fevereiro de 2006, o governo federal sancionou a Lei nº 11.274 alterando a LDB e dispondo sobre a duração de nove anos para o ensino fundamental com matrícula obrigatória a partir dos seis anos de idade em todo o território nacional. Com base nesta lei os municípios, os estados e o distrito federal tiveram prazo até 2010 para implementar o ensino fundamental de nove anos.

**TABELA 1: Matrículas no ensino fundamental em 2004 por UF**

UF	Matrículas no ensino fundamental	Matrículas no ensino fundamental aos seis anos de idade	Matrículas no ensino fundamental de 9 anos
Acre	151.535	2.931	2.913
Alagoas	729.780	20.393	2.920
Amapá	135.778	906	-
Amazonas	771.257	22.077	28.963
Bahia	3.079.584	78.703	126.141
Ceará	1.787.660	71.733	658.492
Distrito Federal	370.219	925	3.493
Espírito Santo	565.489	3.964	4.018
Goiás	1.059.068	28.464	143.203
Maranhão	1.562.374	38.402	75.771
Mato Grosso	611.230	21.048	240.741
Mato Grosso do Sul	442.544	5.318	-
Minas Gerais	3.482.994	107.952	3.055.451
Pará	1.614.942	27.239	57.163
Paraíba	843.607	15.187	41.071
Paraná	1.683.914	33.878	130.495
Pernambuco	1.752.632	49.339	148.001
Piauí	715.891	31.486	38.392
Rio de Janeiro	2.474.150	77.495	2.078.574
Rio Grande do Norte	608.651	19.091	272.935
Rio Grande do Sul	1.669.790	24.708	128.694
Rondônia	318.801	6.522	27.938
Roraima	78.062	648	-
Santa Catarina	952.887	14.844	85.777
São Paulo	5.862.955	9.019	37.876
Sergipe	412.449	4.369	-
Tocantins	274.191	5.957	9.106

**Fonte: Dados do INEP – Síntese do Censo Escolar de 2004 e dados divulgados pelo informativo “Ampliação do ensino fundamental para nove anos – Relatório do Programa”.**

Apesar de se tornar obrigatória somente em 2006, a implementação da política teve início em meados de 2003 em várias unidades da federação, pois neste ano já havia 3.900.000 estudantes matriculados no ensino fundamental de nove anos e em 2004 este número já era de



7.398.128 estudantes<sup>2</sup>. A TAB.1 apresenta os dados de matrícula no ano de 2004 e estes dados mostram que neste ano somente dois estados já estavam avançados nessa mudança: Minas Gerais com 87,72% dos estudantes matriculados no ensino fundamental de nove anos e Rio de Janeiro com 84,01%.

Uma vez que a ampliação do ensino fundamental é uma realidade, é interessante acompanhar esta ampliação e verificar se essa mudança é ou não eficaz em termos de melhoras nos indicadores educacionais. Entende-se como eficaz, neste trabalho, o incremento na progressão por série ou na taxa de matrícula como resultado do aumento do tempo de aprendizado do estudante que começa o seu ciclo de vida estudantil mais cedo entrando no ensino fundamental um ano antes dos outros indivíduos que frequentaram o ensino fundamental antes da implementação da política. Vale ressaltar que essa é uma alternativa a escassez de vagas na educação infantil. Com a análise dos exercícios empíricos será possível inferir se o aumento na duração do ensino fundamental é tão benéfico em termos de aprendizado esperado quanto frequentar a educação infantil.

---

<sup>2</sup> Dados divulgados pelo informativo “Ampliação do ensino fundamental para nove anos – Relatório do Programa”.

## 3 DADOS E METODOLOGIA

### 3.1 Modelos estatísticos utilizados

#### 3.1.1 Método da dupla diferença (DD) ou *diferença-em-diferença*

O método de dupla diferença consiste basicamente em analisar o impacto causado por determinada intervenção em certa variável, que é denominada variável de impacto (ou variável de interesse). Este método, que estima relações causais, é um método bastante simples e consegue contornar muitos problemas de endogeneidade que em geral surgem quando são realizadas comparações entre indivíduos heterogêneos.

Um grupo de controle, ou contra-factual, é eleito para servir de parâmetro de comparação com o grupo que passou pela intervenção. Além da comparação entre os grupos há também a comparação entre diferentes períodos para captar o efeito do tempo sobre a variável de impacto. O grupo de comparação também é útil para evitar que o efeito captado da intervenção seja espúrio. No caso deste trabalho especificamente, a dupla diferença se assemelha a um experimento natural porque a política pública será analisada através de dois grupos distintos: um que adotou a política e outro que não adotou.

É possível calcular a dupla diferença com e sem a utilização de regressões. Sem a regressão, a dupla diferença é obtida através de médias (valores agregados) da variável de intervenção ou de certo indicador de interesse.

	Grupo de tratamento	Grupo de controle
Antes	$T_A$	$C_A$
Depois	$T_D$	$C_D$

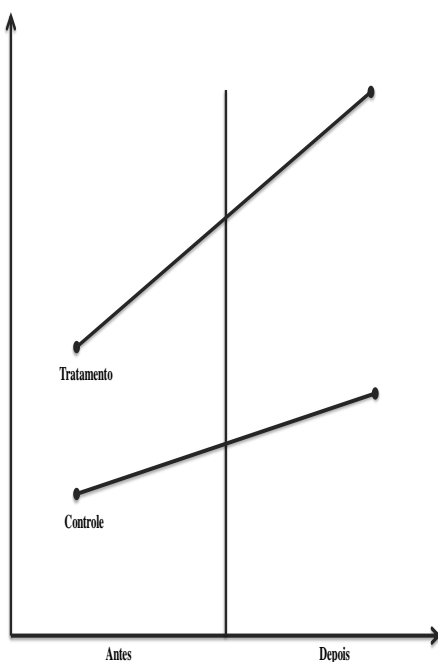
Quadro 2: Estrutura simples do método da dupla diferença

O estimador da dupla diferença é a variação média ocorrida no grupo de tratamento menos a variação média ocorrida no grupo de controle. Como não é possível que todo indivíduo da amostra passe pela intervenção e ao mesmo tempo não passe pela intervenção, o cálculo deste efeito não é factível. A saída é calcular a diferença média com grupos de controle que simulem este mesmo indivíduo. Neste caso, o estimador é obtido de maneira simples:

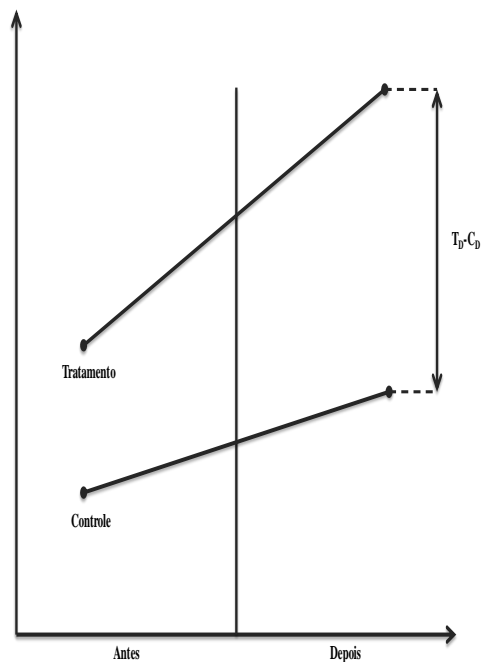
$$\text{Efeito da intervenção: } (T_D - T_A) - (C_D - C_A)$$

Essa maneira é muito simples e não consegue captar outros fatores que influenciam no efeito da intervenção. A dupla diferença estimada via regressão permite que seja captada a influência do tempo sobre o efeito da intervenção assim como a influência de outros fatores que servem como controle para a regressão. Os GRAF. 1 a 4 representam a ideia da dupla diferença.

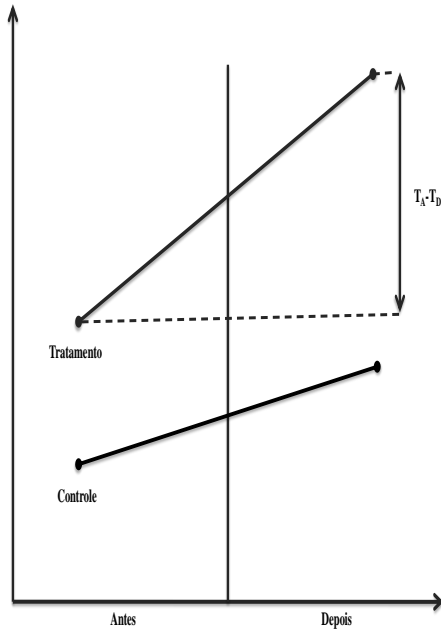
**GRÁFICO 1 – Representação gráfica dos dois grupos do método da dupla diferença**



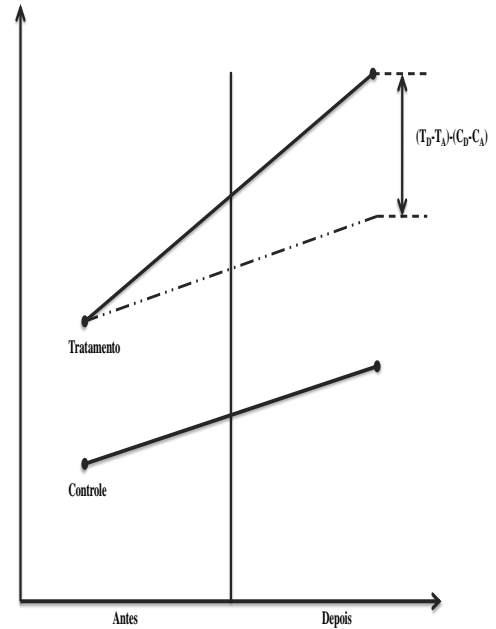
**GRÁFICO 2 – Diferença obtida para o grupo de tratamento**



**GRÁFICO 3 – Diferença obtida entre o grupo de tratamento e o grupo de controle depois da intervenção**



**GRÁFICO 4 – Representação correta da estimação da dupla diferença**



O modelo básico da regressão pode ser descrito na seguinte forma genérica:

$$y_i = \beta_0 + \beta_1 \text{tratamento} + \beta_2 \text{tempo} + \beta_3 \text{interação} + \beta_i \text{variáveis de controle} + e_i \quad (1)$$

sendo que a interação = tratamento x tempo.

Os três efeitos captados pela regressão que usa o método da dupla diferença são:

- efeito tratamento ( $\beta_1$ ): capta a diferença inerente aos grupos de tratamento e de controle. Esta diferença independe da intervenção ou política aplicada.
- efeito tempo ( $\beta_2$ ): capta a tendência da variável ao longo do tempo.
- efeito global ou efeito interação ( $\beta_3$ ): é o coeficiente de interesse que capta o impacto puro da intervenção que é dado pelo efeito do tratamento ao longo do tempo.

O estimador da dupla diferença é o  $\beta_3$  que é obtido através de um modelo paramétrico linear. Também chamado de termo de interação, ele apenas fornece informações sobre o grupo de tratamento e nada pode ser inferido a respeito do grupo de controle. Há alguns problemas que podem tornar o estimador de dupla diferença viesado. Entre eles está o tamanho da série temporal (o número de períodos observados) utilizada e a correlação entre as variáveis de controle.

Corrigido o problema da correlação entre as variáveis de controle, para estimar a variável de impacto é de extrema importância o uso das variáveis de controle na regressão. Elas ajudam a explicar a variação nos valores observados da variável de impacto de cada indivíduo dos dois grupos e a controlar as diferenças entre os dois grupos que não são causadas pela intervenção. Isso impede que a correlação entre o erro e as variáveis que seriam omitidas torne o estimador viesado.

Como todo modelo econométrico, a dupla diferença precisa de alguns pressupostos para ser utilizada. O primeiro deles é o pressuposto da aleatoriedade. As intervenções ou políticas avaliadas pelo método deveriam ser aleatórias com relação aos indivíduos tratados. A variável de impacto passa por uma mudança que deveria ser exógena e o grupo de tratamento e o grupo de controle deveriam ser observados pelo menos em dois períodos diferentes – antes e depois na intervenção. Isso quase não ocorre nos experimentos das ciências sociais aplicadas. Conforme mencionado anteriormente, a comparação entre duas UF, uma que adotou e outra que não adotou a referida política, pode gerar este experimento quase natural.

(...) Comparações simples de variáveis de interesse pré-tratamento e pós-tratamento para aqueles indivíduos expostos a alguma intervenção têm grande chance de estarem contaminadas por tendências temporais da variável de impacto ou pelo efeito de outros eventos. No entanto, quando somente uma fração da população é exposta à intervenção, um grupo de comparação que não sofreu intervenção pode ser usado para identificar a variação temporal na variável de impacto que não se deve à exposição ao tratamento. (Abadie, 2003, p.1)

Com relação ao tipo de dados que pode ser utilizado no método da dupla diferença, uma vez que nesse caso o estimador é calculado em termos da média da amostra, é indicado o uso de dados em painel ou *cross-section*. (Blundell & MaCurdy, 1999). Este trabalho utiliza uma base de dados *cross-section*, a PNAD (Pesquisa Nacional por Amostra de Domicílios) realizada anualmente pelo IBGE desde 1971. Esta pesquisa abrange grandes temas de

interesse da Demografia tais como migração, educação, trabalho, fecundidade, família, características físicas dos domicílios, rendimento, dentre outros. Ela é representativa para o Brasil e para as unidades da federação, sendo que até 2004 ela não cobria as áreas rurais de Rondônia, Acre, Amazonas, Roraima, Pará e Amapá.

O modelo de dupla diferença será estimado através de dois métodos de regressão – *logit* e mínimos quadrados ordinários – que serão explicados nas duas próximas sub-seções.

### 3.1.2 Método de estimação logit

A regressão logística é um tipo de regressão desenvolvida especialmente para variáveis binárias, aquelas para as quais só há duas respostas possíveis. Esse tipo de variável pode ser classificado da seguinte forma:

$\begin{cases} 0 = \text{sucesso} \\ 1 = \text{fracasso} \end{cases}$ . Estas variáveis possuem distribuição de Bernoulli com as seguintes probabilidades:

$\begin{cases} P(Y = 1) = \pi \\ P(Y = 0) = 1 - \pi \end{cases}$ . Sendo que  $\pi = E(Y)$ .

Quando  $Y_i$  tem distribuição de Bernoulli com parâmetro  $\pi_i$ , a função de probabilidade de massa é:

$$f(y_i, \pi_i) = \pi_i^{y_i} (1 - \pi_i)^{1-y_i} = (1 - \pi_i) \left[ \frac{\pi_i}{(1 - \pi_i)} \right]^{y_i} = (1 - \pi_i) \exp \left[ y_i \log \left( \frac{\pi_i}{1 - \pi_i} \right) \right] \quad (2)$$

para  $y_i=0$  e 1. Esta distribuição está na família exponencial natural. O parâmetro natural  $Q(\pi)=\log \left[ \frac{\pi}{1-\pi} \right]$ , que é o logaritmo da chance de resposta 1, é chamado de *logit* de  $\pi$ . “Modelos lineares generalizados que usam a ligação (link) logit são chamados modelos logit.” (Agresti, 1990).

Ainda de acordo com Agresti (1990), se denotarmos os dois possíveis resultados por 0 e 1 obteremos a variável aleatória de Bernoulli com média

$$E(Y)=1*P(Y=1)+0*P(Y=0)=P(Y=1). \quad (3)$$

Esta probabilidade pode ser chamada de  $\pi(x)$ , refletindo sua dependência dos valores das variáveis explanatórias  $\mathbf{X}=(X_1,\dots,X_k)$ . E, se

$$E(Y^2)=1^2*\pi(\mathbf{x})+0^2*[1-\pi(\mathbf{x})]=\pi(\mathbf{x}), \quad (4)$$

a variância de  $Y_{is}$  será dada por

$$V(Y)=E(Y^2)-[E(Y)]^2=\pi(\mathbf{x})*[1-\pi(\mathbf{x})]. \quad (5)$$

Se o modelo de variável resposta binária fosse um modelo de probabilidade linear (relação linear entre  $x$  e  $\pi(\mathbf{x})$ ) ele seria denotado da seguinte maneira:

$$E(Y)=\pi(\mathbf{x})=\alpha+\beta x. \quad (6)$$

Mas como a variância não é constante, depende de  $\mathbf{x}$   $\{V(Y)=\pi(\mathbf{x})*[1-\pi(\mathbf{x})]\}$ , os estimadores não seriam de mínima variância dentre os não-viesados. Por isso, não podemos usar um modelo linear. É preciso usar o logit.

Assim, o modelo de regressão logística pode ser representado pela seguinte função:

$$\pi(x) = \frac{\exp(\alpha+\beta x)}{1+\exp(\alpha+\beta x)} \quad (7)$$

Neste modelo, a chance de se obter a resposta ( $Y=1$ ) é

$$\frac{\pi(x)}{1-\pi(x)} = \exp(\alpha + \beta x) = e^\alpha (e^\beta)^x. \quad (8)$$

Para linearizar esta função, calcula-se o logaritmo da mesma.

$$\log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \alpha + \beta x. \quad (9)$$

O efeito de  $x_i$  sobre  $Y$  é obtido calculando-se o  $\exp(\beta)$ . A partir daí é possível saber como a variação de  $x_i$  afeta  $Y$ . O modelo logit é adequado para a variável dependente binária, mas o coeficiente estimado não é linear e a interpretação não corresponde a um efeito marginal linear.

### 3.1.3 Método de estimação dos mínimos quadrados ordinários

Em geral, ao estimar uma regressão, o intuito é inferir para toda uma população aquilo que observamos na amostra. A regressão é uma função matemática que identifica as relações entre as variáveis independentes e a variável dependente. Essa função pode ser expressa de forma simples

$$Y_i = \beta_0 + \beta_i X_i + u_i \quad (10)$$

Mas como não é possível estimar a função para toda a população porque normalmente não há dados de todos os indivíduos da população, estima-se a função da amostra

$$Y_i = \hat{\beta}_0 + \hat{\beta}_i X_i + \hat{u}_i \Rightarrow Y_i = \hat{Y}_i + \hat{u}_i \quad (11)$$

Podemos calcular o erro amostral manipulando a função da seguinte maneira

$$\hat{u}_i = Y_i - \hat{Y}_i \quad (12)$$

Para obter uma estimativa que seja o mais próxima possível da função da população Gauss sugeriu a minimização da soma dos quadrados dos erros. Neste caso, a soma dos erros é a menor possível, e é isso que dá a esse método de estimação o nome de mínimos quadrados ordinários (MQO). Neste método são escolhidos os coeficientes estimados  $\hat{\beta}_i$  que minimizam a  $\sum \hat{u}_i^2$ .



Embora a estimativa do modelo de MQO para uma variável dependente binária seja inadequada, a interpretação do coeficiente beta para a interação oferece uma ideia assintótica do efeito marginal linear.

### 3.1.4 Método de cálculo da Progressão por série

A probabilidade de progressão por série é um indicador que expressa a probabilidade de que um indivíduo que cursou a série  $i+1$  seja aprovado nesta série dado que ele foi aprovado na série  $i$ .

$$e_i = \frac{P_{i+1}}{P_i} \quad (13)$$

Onde:  $e_i$  = PPS = probabilidade de progressão da série  $i$  para a série  $i+1$ ;

$P_{i+1}$  = pessoas no grupo etário que concluíram pelo menos a série  $i+1$ ;

$P_i$  = pessoas no grupo etário que concluíram pelo menos a série  $i$ .

Este indicador foi criado com base no método demográfico da Razão de Progressão por Parturição. Rios-Neto (2004) argumenta que a utilidade deste indicador consiste em fornecer um bom panorama das diversas transições dentro do sistema de ensino, identificar o padrão de evasão definitiva (abandono por parte do aluno) para cada série escolar, possibilitar o cálculo dos anos médios de estudo e também permitir previsões dos anos médios de estudo.

### **3.1.5 Desenho das amostras para os exercícios empíricos**

Para estimar o impacto da política educacional que aumenta em um ano a duração do ensino fundamental foram feitos três tipos de exercícios com a metodologia da dupla diferença. Em dois deles a variável de interesse foi estar matriculado e em outro foi a probabilidade de progressão por série. Nestes três exercícios foram testados os estudantes de Minas Gerais, de modo que eles formaram o grupo de tratamento. Como grupos de controle foram escolhidos os estudantes de São Paulo e os estudantes da Bahia, sendo que foram feitas regressões separadas com cada grupo de controle.

#### **3.1.5.1 Matrícula como variável de interesse**

Basicamente, o que diferencia os dois tipos de exercício com essa variável de resposta é a definição do tempo na análise – os períodos antes e depois – do método da dupla diferença. No primeiro exercício as amostras foram selecionadas agrupando dados da PNAD de vários anos (entre 2001 e 2007). As observações foram coletadas em anos diferentes da pesquisa e agrupadas em um mesmo banco de dados, sendo que a PNAD de determinado ano forneceu os dados do período “antes” e a PNAD de outro ano posterior forneceu os dados do período “depois”. A definição de qual PNAD usar foi feita com base no ano de implementação da política em Minas Gerais (2003) e também com base nas idades de interesse, pois o foco é o início do ensino fundamental. No segundo exercício as observações foram coletadas no mesmo ano, ou seja, na mesma PNAD. Neste caso, o período “antes” foi determinado por um grupo de indivíduos com idade  $x$  e o período “depois” foi determinado como um grupo de indivíduos de idade  $y$ . A definição de qual PNAD usar foi feita com base nos mesmos critérios usados no primeiro exercício.

Para captar de forma mais efetiva o impacto da política, a análise é feita não só para os estudantes que haviam acabado de ingressar no ensino fundamental no momento da intervenção, mas também para os estudantes que cursavam a segunda e a terceira séries. Desse modo é possível perceber se a política fez efeito em outras séries do ensino

fundamental e não só na primeira, ou seja, se o efeito da mesma seria cumulativo ou ainda se o impacto da mesma poderia ser mais tardio sendo observado somente em séries posteriores.

Os quadros 3 e 4 ilustram a maneira através da qual foram desenhadas as amostras para os dois tipos de exercícios empíricos realizados para a variável de interesse matrícula (estar matriculado). No lado esquerdo estão as PNAD que foram utilizadas para formar o banco de dados e no lado direito estão as faixas etárias que compõem as amostras.

2001 - 2005	Crianças de sete anos - estavam na primeira série antes ou depois da intervenção
2001+2002 - 2005+2006	Crianças de sete e oito anos - estavam na primeira ou na segunda série antes ou depois da intervenção
2001+2002+2003 - 2005+2006+2007	Crianças de sete, oito e nove anos - estavam na primeira, segunda ou terceira séries antes ou depois da intervenção

**Quadro 3: Desenho da amostra do exercício I.**

2005	Crianças de sete anos - estavam na primeira série naquele ano posterior à intervenção Crianças de oito anos - estavam na primeira série no ano anterior à intervenção
2006	Crianças de sete e oito anos - estavam na primeira ou na segunda série nos dois primeiros anos posteriores à intervenção Crianças de nove e dez anos - estavam na primeira ou na segunda série nos dois primeiros anos anteriores à intervenção
2007	Crianças de sete, oito e nove anos - estavam na primeira, ou na segunda ou na terceira série nos três primeiros anos posteriores à intervenção Crianças de onze, doze e treze anos - estavam na primeira, ou na segunda ou na terceira série nos três primeiros anos anteriores à intervenção

**Quadro 4: Desenho da amostra do exercício II.**

### 3.1.5.2 Probabilidade de progressão por série como variável de interesse

Neste exercício as amostras foram compostas por indivíduos de doze e treze anos e os dados foram retirados das PNAD 2007 e 2011. As observações de 2007 são o período “antes” do

método da dupla diferença e as observações de 2011 são o período “depois” do método da dupla diferença. O objetivo é identificar o impacto da política na probabilidade de progressão por série, nas séries iniciais do ensino fundamental. Teoricamente, adiantar a entrada do indivíduo no sistema de ensino possibilitaria mais tempo para a adaptação do estudante e para trabalhar com alfabetização o que teria um impacto positivo e mais imediato nas séries iniciais. Em especial, a probabilidade de progressão do quinto para o sexto ano do novo ensino fundamental é analisada aqui por ser um gargalo no ensino brasileiro (da antiga quarta para a quinta série).

$e_0$	Probabilidade de concluir o primeiro ano do Ensino Fundamental dado que tinha 0 anos de estudo completos
$e_1$	Probabilidade de concluir o segundo ano do Ensino Fundamental dado que concluiu o primeiro
$e_2$	Probabilidade de concluir o terceiro ano do Ensino Fundamental dado que concluiu o segundo
$e_3$	Probabilidade de concluir o quarto ano do Ensino Fundamental dado que concluiu o terceiro
$e_4$	Probabilidade de concluir o quinto ano do Ensino Fundamental dado que concluiu o quarto

**Quadro 5: Desenho da amostra do exercício III.**

## 4 RESULTADOS

### 4.1 Descrição dos bancos de dados utilizados nos exercícios

#### 4.1.1 Exercício I

Conforme já foi explicitado na seção anterior, o exercício 1 foi feito com base na seguinte amostra:

2001 - 2005	Crianças de sete anos - estavam na primeira série antes ou depois da intervenção
2001+2002 - 2005+2006	Crianças de sete e oito anos - estavam na primeira ou na segunda série antes ou depois da intervenção
2001+2002+2003 - 2005+2006+2007	Crianças de sete, oito e nove anos - estavam na primeira, segunda ou terceira séries antes ou depois da intervenção

**Quadro 3: Desenho da amostra do exercício I**

Para fornecer uma ideia da situação da matrícula nas primeiras quatro séries do ensino fundamental a TAB. 2 mostra o número absoluto de matrículas em cada unidade da federação, independentemente da idade do indivíduo matriculado em cada um dos três períodos de tempo contemplados pelo exercício I. Neste caso, o aumento expressivo da matrícula é atribuído ao fato de que em cada cenário subsequente a amostra é aumentada pelas observações de outros anos e por incorporar indivíduos de outras idades além dos sete anos, de modo que o aumento da matrícula não deve ser interpretado entre os anos e a comparação só pode ser feita entre as unidades da federação para o mesmo período. O ponto é fornecer informações gerais sobre os bancos de dados que foram obtidos seguindo as definições apresentadas no quadro 1. Este número de matrículas é uma aproximação obtida a partir dos dados de uma pesquisa abrangente e não podem ser tomados como substitutos dos dados de matrícula do Censo Escolar do INEP. Observa-se que os estados de Minas Gerais, São Paulo e Bahia por terem população mais expressiva apresentam os mais elevados números de matrículas.

**TABELA 2: Número de matrículas de 1ª a 4ª série por unidade federativa – exercício I**

Estado	2001-2005	2001-2006	2001-2007
Rondônia	268.836	535.863	804.809
Acre	148.929	289.195	436.946
Amazonas	713.043	1.459.747	2.224.449
Roraima	70.017	143.055	222.150
Pará	1.546.414	3.078.235	4.637.401
Amapá	102.971	235.961	367.520
Tocantins	309.617	627.368	909.633
Maranhão	1.652.081	3.262.649	4.882.796
Piauí	840.130	1.654.771	2.418.722
Ceará	1.849.956	3.706.381	5.468.711
Rio Grande do Norte	661.456	1.328.597	2.001.926
Paraíba	930.124	1.830.196	2.640.850
Pernambuco	1.939.002	3.826.244	5.682.898
Alagoas	877.285	1.717.889	2.503.699
Sergipe	473.488	959.454	1.406.515
<b>Bahia</b>	<b>3.524.215</b>	<b>6.877.010</b>	<b>10.212.491</b>
<b>Minas Gerais</b>	<b>3.276.095</b>	<b>6.488.162</b>	<b>9.853.229</b>
Espírito Santo	577.514	1.181.433	1.791.598
Rio de Janeiro	2.421.146	4.742.899	7.086.754
<b>São Paulo</b>	<b>5.830.835</b>	<b>11.834.704</b>	<b>17.905.794</b>
Paraná	1.759.065	3.471.059	5.224.681
Santa Catarina	908.565	1.866.246	2.821.220
Rio Grande do Sul	1.677.396	3.365.535	5.066.089
Mato Grosso do Sul	419.690	832.112	1.223.239
Mato Grosso	565.620	1.125.548	1.687.608
Goiás	1.059.454	2.070.295	3.063.293
Distrito Federal	372.188	760.852	1.143.668

Fonte dos dados básicos: PNAD - IBGE.

**TABELA 3: Número de matrículas de 1ª a 4ª série antes e depois da intervenção – exercício I**

	Antes	Depois
2001-2005	17.501.065	17.274.067
2001-2006	34.660.411	34.611.049
2001-2007	51.664.618	52.024.071

Fonte dos dados básicos: PNAD - IBGE.

A TAB. 3 apresenta o número de matrículas em todo o país para os dois períodos do método da dupla diferença para o exercício I. Nos períodos analisados – 2001-2005; 2001-2006 e 2001-2007 – o número de matrículas antes e depois da intervenção apresenta valores bem

aproximados, mas a política ainda não havia sido implementada em todas as unidades da federação.

#### 4.1.2 Exercício II

Assim como a TAB. 2, a TAB. 4 desagrega as matrículas para os dois períodos do método da dupla diferença. Nos anos analisados –2005; 2006 e 2007 – o número de matrículas antes e depois da intervenção apresenta valores bem díspares com um número muito maior de observações no período posterior à política. Isso é facilmente explicado observando que nas amostras retiradas das PNAD de 2006 e de 2007 foram considerados também os estudantes de outros anos do ensino fundamental. Assim como as tabelas descritivas apresentadas para o exercício I, o intuito é simplesmente fornecer informações gerais sobre os bancos de dados formados de acordo com a descrição do quadro 4.

2005	Crianças de sete anos - estavam na primeira série naquele ano posterior à intervenção Crianças de oito anos - estavam na primeira série no ano anterior à intervenção
2006	Crianças de sete e oito anos - estavam na primeira ou na segunda série nos dois primeiros anos posteriores à intervenção Crianças de nove e dez anos - estavam na primeira ou na segunda série nos dois primeiros anos anteriores à intervenção
2007	Crianças de sete, oito e nove anos - estavam na primeira, ou na segunda ou na terceira série nos três primeiros anos posteriores à intervenção Crianças de onze, doze e treze anos - estavam na primeira, ou na segunda ou na terceira série nos três primeiros anos anteriores à intervenção

**Quadro 4: Desenho da amostra do exercício II.**

**TABELA 4: Número de matrículas de 1ª a 4ª série por unidade federativa – exercício II**

Estado	2005	2006	2007
Rondônia	165.424	163.390	160.256
Acre	95.845	90.853	95.592
Amazonas	431.324	436.691	459.568
Roraima	40.157	47.938	49.951
Pará	942.087	941.140	944.967
Amapá	72.003	69.684	75.280
Tocantins	150.465	153.084	138.682
Maranhão	756.569	773.492	803.877
Piauí	377.301	376.830	344.063
Ceará	896.102	887.094	857.534
Rio Grande do Norte	324.968	304.998	313.794
Paraíba	409.034	404.194	390.162
Pernambuco	926.510	917.810	895.538
Alagoas	426.785	396.370	370.231
Sergipe	219.010	231.528	216.819
Bahia	1.591.157	1.555.954	1.567.272
Minas Gerais	1.602.101	1.636.149	1.696.544
Espírito Santo	273.885	312.632	311.896
Rio de Janeiro	1.224.825	1.168.844	1.184.864
São Paulo	2.980.624	3.075.740	3.145.279
Paraná	863.038	844.645	898.890
Santa Catarina	451.503	480.012	479.643
Rio Grande do Sul	841.610	853.728	873.428
Mato Grosso do Sul	206.718	209.071	181.505
Mato Grosso	278.783	281.360	279.726
Goiás	531.253	514.755	476.169
Distrito Federal	194.986	208.996	201.492

Fonte dos dados básicos: PNAD - IBGE.

A TAB. 4 apresenta o número de matrículas de primeira à quarta série em cada ano. Entre 2005 e 2006 o número de matrículas na Bahia caiu 2,2% e em Minas Gerais e em São Paulo cresceu 2,1% e 3,1% respectivamente. Entre 2006 e 2007 houve crescimento de 0,73% na Bahia, de 3,7% em Minas Gerais e de 2,3% em São Paulo.

**TABELA 5: Número de matrículas de 1ª a 4ª série antes e depois da intervenção – exercício II**

	Antes	Depois
2005	3.330.697	5.698.753
2006	5.853.160	9.224.688
2007	3.164.730	12.336.578

Fonte dos dados básicos: PNAD - IBGE.



As tabelas que compõem os apêndices B, C, D e E vão além dos dados de matrícula e apresentam a distribuição das variáveis dependentes utilizadas em todas as regressões. Estas tabelas fornecem informações sucintas, porém importantes, para descrever as características socioeconômicas dos indivíduos que compõem as amostras. Ao analisá-las percebe-se que em todos os períodos contemplados, nas amostras dos exercícios I, II e III, as características identificadas na amostra de cada estado foram praticamente as mesmas. Mudanças nos percentuais de distribuição de cada variável foram observadas, mas nada modificou significativamente as características das amostras.

Ao considerar as características individuais, nota-se que nos três estados a maioria da amostra é composta por indivíduos do sexo masculino. No que diz respeito a raça, há diferenças, pois em MG e na BA a maioria da amostra é composta por indivíduos pardos ou negros e em SP a maioria é de brancos ou amarelos. O tipo de família mais comum nos três estados é casal com mais de um filho, mas vale ressaltar que a amostra da BA possui mais indivíduos cujo tipo de família é mãe com filhos (sem o pai), quando comparada a MG e SP. Comparando a renda domiciliar per capita (em salários mínimos) há grande diferença entre os três estados sendo que a renda dos indivíduos da BA não chega à metade da renda dos indivíduos de SP, e MG possui um valor acima da BA e abaixo de SP.

Com relação às características dos chefes do domicílio no qual os indivíduos vivem, a maioria das amostras dos três estados é composta por chefes ocupados e do sexo masculino. Em MG e na BA a maioria deles é pardo ou negro e em SP a maioria é branco ou amarelo. Na BA a maioria dos chefes de domicílio possui em média zero anos de estudo enquanto em MG e SP a maioria possui em média quatro anos de estudo<sup>3</sup>.

Ao analisar as características do domicílio identifica-se que as amostras dos três estados são compostas, em sua maioria, de indivíduos cujos domicílios ficam na zona urbana, em região não-metropolitana, com água encanada e com rede elétrica. Em MG e em SP na maioria dos domicílios há coleta de lixo e esgotamento sanitário, mas na BA a maioria não possui coleta de lixo e tampouco esgotamento sanitário<sup>4</sup>.

---

<sup>3</sup> Nas amostras do exercício II, a maioria dos chefes de domicílio de SP possuía 11 anos de estudo.

<sup>4</sup> Na amostra do estado da BA para o exercício II, na maioria dos domicílios há coleta de lixo.

Este foi um breve panorama dos bancos de dados usados nas estimações das regressões que pode auxiliar na interpretação dos resultados das mesmas que é apresentado na próxima seção.

## **4.2 Resultados dos modelos estatísticos**

O objetivo principal destes exercícios é verificar se a intervenção feita na duração do ensino fundamental teve impacto positivo na matrícula ou na progressão por série. Conforme explicado na seção 3.1.1, esta informação é dada pela razão de probabilidade da variável *dummy* interação. Deste modo a análise feita nesta seção será focada nesta variável que faz parte da metodologia da dupla diferença.

A seguir serão apresentados os resultados resumidos de cada exercício empírico realizado e as tabelas com estes resultados completos estão no final do trabalho no Apêndice A.

### **4.2.1 Minas Gerais como grupo de tratamento e São Paulo como grupo de controle**

#### **A – Exercício I**

O exercício I foi feito com mais de uma PNAD filtrando de cada uma os indivíduos que possuíam determinada idade de interesse definida de acordo com a idade que seria necessária para cursar certos anos do ensino fundamental (ver quadro 3).

Os resultados da estimação pelo método de probabilidade linear (TAB. 6) mostram que a variável chave do modelo usado neste trabalho, a variável *dummy* interação, é estatisticamente

significativa<sup>5</sup> nos dois últimos períodos de análise: 2001-2006 e 2001-2007 e o sinal do coeficiente em ambos períodos é negativo. Com relação às variáveis que captam a situação socioeconômica e de *background* familiar, poucas foram estatisticamente significativas<sup>6</sup>. Para os estudantes de MG e SP estas variáveis são ainda menos importantes na explicação do que pode influenciar a matrícula do que para os estudantes de MG e BA.

**Tabela 6 – Resumo dos resultados do Exercício I com o modelo de probabilidade linear – MG e SP**

Variável	2001-2005			2001-2006			2001-2007		
	Coeficiente	P-valor	Desvio padrão	Coeficiente	P-valor	Desvio padrão	Coeficiente	P-valor	Desvio padrão
Ano	0,120	0,00	0,012	0,064	0,00	0,006	0,048	0,00	0,004
Tratamento	-0,014	0,23	0,012	-0,012	0,10	0,007	-0,006	0,29	0,005
<b>Interação</b>	<b>0,008</b>	<b>0,67</b>	<b>0,020</b>	<b>-0,023</b>	<b>0,03</b>	<b>0,011</b>	<b>-0,031</b>	<b>0,00</b>	<b>0,007</b>

Fonte dos dados básicos: PNAD - IBGE.

No modelo logit (TAB.7) os resultados são os mesmos para todos os períodos analisados<sup>7</sup> em termos de significância quando o comparamos com o modelo de probabilidade linear. Nos dois períodos nos quais a variável interação foi estatisticamente significativa a razão de chance foi inferior a um mostrando que a probabilidade de um indivíduo do grupo de tratamento estar matriculado é inferior à probabilidade de um indivíduo do grupo de controle estar matriculado.

**Tabela 7 – Resumo dos resultados do Exercício I com o modelo logit – MG e SP**

Variável	2001-2005			2001-2006			2001-2007		
	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão
Ano	11,802	0,00	3,888	2,639	0,00	0,272	2,267	0,00	0,178
Tratamento	0,885	0,30	0,104	0,896	0,17	0,072	0,943	0,37	0,062
<b>Interação</b>	<b>0,760</b>	<b>0,54</b>	<b>0,337</b>	<b>0,594</b>	<b>0,00</b>	<b>0,081</b>	<b>0,541</b>	<b>0,00</b>	<b>0,057</b>

Fonte dos dados básicos: PNAD - IBGE.

<sup>5</sup> Neste trabalho é considerada como estatisticamente significativa a variável cujo coeficiente ou razão de chance tem p-valor de até 10%.

<sup>6</sup> No modelo de probabilidade linear somente a variável se o indivíduo reside em um domicílio que fica na zona urbana foi estatisticamente significativa no período 2001-2005. No período 2001-2006, as variáveis tipo de família na qual o indivíduo está inserido, condição de ocupação do chefe do domicílio, se o indivíduo reside em um domicílio que fica na região metropolitana, se o indivíduo reside em um domicílio que fica na zona urbana e se o indivíduo reside em um domicílio onde há água encanada foram estatisticamente significativas. No último período de análise somente três variáveis foram estatisticamente significativas: tipo de família na qual o indivíduo está inserido, sexo do indivíduo e renda domiciliar per capita.

<sup>7</sup> Com uma única exceção: no período 2001-2007 a variável se no domicílio há água encanada foi estatisticamente significativa no modelo logit, mas não foi no modelo de probabilidade linear.

Como os resultados para a variável interação contrariaram a expectativa inicial, as regressões foram estimadas novamente, dessa vez usando a amostragem de desenho complexo cujo objetivo é extirpar da amostra observações que atrapalhem a estimação dos coeficientes e das razões de chance das regressões. Os resultados destas novas estimações (TAB. 8 e TAB. 9) não trouxeram melhora alguma para os modelos com relação ao p-valor e tampouco com relação ao desvio padrão<sup>8</sup>. Assim, o exercício I aponta para a ineficácia da política de antecipação da entrada do ensino fundamental em Minas Gerais em termos de impacto sobre a matrícula.

**Tabela 8 – Resumo dos resultados do Exercício I com o modelo de probabilidade linear com desenho complexo – MG e SP**

Variável	2001-2005			2001-2006			2001-2007		
	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão
Ano	0,120	0,00	0,010	0,065	0,00	0,007	0,048	-	-
Tratamento	-0,014	0,33	0,015	-0,012	0,15	0,009	-0,006	-	-
<b>Interação</b>	<b>0,007</b>	<b>0,65</b>	<b>0,016</b>	<b>-0,023</b>	<b>0,03</b>	<b>0,011</b>	<b>-0,031</b>	-	-

Fonte dos dados básicos: PNAD - IBGE.

**Tabela 9 – Resumo dos resultados do Exercício I com o modelo logit com desenho complexo – MG e SP**

Variável	2001-2005			2001-2006			2001-2007		
	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão
Ano	11,795	0,00	3,869	2,658	0,00	0,294	2,262	-	-
Tratamento	0,888	0,33	0,109	0,894	0,19	0,076	0,942	-	-
<b>Interação</b>	<b>0,743</b>	<b>0,49</b>	<b>0,319</b>	<b>0,590</b>	<b>0,00</b>	<b>0,087</b>	<b>0,541</b>	-	-

Fonte dos dados básicos: PNAD - IBGE.

## B – Exercício II

<sup>8</sup> A única diferença tanto no modelo de probabilidade linear quanto no modelo logit foi observada no período 2001-2006 no qual a variável tipo de família na qual o indivíduo está inserido deixa de ser significativa na regressão com amostra de desenho complexo. Vale ressaltar que no período 2001-2007 não foi possível estimar de forma satisfatória as regressões com amostragem de desenho complexo porque nesta amostra há um ou mais estratos com apenas uma observação.

O exercício II foi feito com somente uma PNAD para cada período analisado. A amostra foi obtida filtrando do mesmo banco de dados os indivíduos de todas as idades de interesse (ver quadro 4).

Neste exercício, assim como no anterior, nos resultados obtidos através do modelo de probabilidade linear (TAB. 10), a variável interação é estatisticamente significativa somente para os dois últimos períodos de análise (2006 e 2007) e os coeficientes tem sinal negativo. Além disso, poucas variáveis que captam a situação socioeconômica e de *background* familiar foram estatisticamente significativas<sup>9</sup>.

**Tabela 10 – Resumo dos resultados do Exercício II com o modelo de probabilidade linear – MG e SP**

Variável	2005			2006			2007		
	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão
Ano	-0,095	0,00	0,011	-0,049	0,00	0,005	-0,009	0,06	0,005
Tratamento	0,013	0,42	0,015	-0,002	0,75	0,006	0,005	0,34	0,006
<b>Interação</b>	<b>-0,023</b>	<b>0,21</b>	<b>0,018</b>	<b>-0,040</b>	<b>0,00</b>	<b>0,009</b>	<b>-0,043</b>	<b>0,00</b>	<b>0,008</b>

**Fonte dos dados básicos: PNAD - IBGE.**

No modelo logit (TAB. 11) a variável interação só foi estatisticamente significativa para o período 2007. Assim como no exercício anterior o valor da razão de chance desta variável foi inferior a um. Mais uma vez a intervenção na duração do ensino fundamental se mostrou ineficaz quando comparamos estudantes de MG com estudantes de SP no que diz respeito a matrícula. Com relação às demais variáveis, os resultados foram quase os mesmos obtidos no modelo de probabilidade linear<sup>10</sup>.

<sup>9</sup> No modelo de probabilidade linear, no período 2005 somente as variáveis raça do chefe do domicílio e se o domicílio fica na região metropolitana foram estatisticamente significativas; no período 2006 as variáveis condição de ocupação do chefe do domicílio, raça do chefe do domicílio, se o domicílio fica na região metropolitana, se no domicílio há água encanada e se no domicílio há eletricidade foram estatisticamente significativas; no período 2007 as variáveis sexo do indivíduo, tipo de família na qual o indivíduo está inserido, escolaridade do chefe do domicílio, raça do chefe do domicílio, renda domiciliar per capita e se no domicílio há eletricidade foram estatisticamente significativas.

<sup>10</sup> Comparando os resultados do modelo logit com os resultados do modelo de probabilidade linear com relação às variáveis de condições socioeconômicas e *background* familiar é possível identificar as seguintes diferenças: no período 2005 a variável sexo do chefe de domicílio não foi estatisticamente significativa no modelo de probabilidade linear, no período 2006 a variável renda domiciliar per capita não foi estatisticamente significativa no modelo de probabilidade linear e no período 2007 a variável raça do chefe do domicílio não foi estatisticamente significativa no logit.

**Tabela 11 – Resumo dos resultados do Exercício II com o modelo logit – MG e SP**

Variável	2005			2006			2007		
	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão
Ano	0,096	0,00	0,033	0,103	0,00	0,029	0,767	0,07	0,110
Tratamento	1,338	0,59	0,716	0,521	0,05	0,176	1,158	0,38	0,192
<b>Interação</b>	<b>0,708</b>	<b>0,53</b>	<b>0,387</b>	<b>1,140</b>	<b>0,71</b>	<b>0,407</b>	<b>0,435</b>	<b>0,00</b>	<b>0,088</b>

**Fonte dos dados básicos: PNAD - IBGE.**

Para verificar os resultados obtidos em ambos os modelos, a estimação por amostragem de desenho complexo foi feita também neste exercício (TAB. 12 e TAB. 13). Não houve mudança alguma nos resultados para a variável interação em nenhum dos dois modelos. Só foram observadas mudanças nas variáveis de condições socioeconômicas e *background* familiar<sup>11</sup> o que somente corrobora que essa política educacional não teve impacto sobre a matrícula.

**Tabela 12 – Resumo dos resultados do Exercício II com o modelo de probabilidade linear com desenho complexo – MG e SP**

Variável	2005			2006			2007		
	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão
Ano	-0,096	0,00	0,010	-0,049	0,00	0,005	-0,009	0,07	0,005
Tratamento	0,012	0,07	0,007	-0,002	0,65	0,004	0,006	0,25	0,005
<b>Interação</b>	<b>-0,020</b>	<b>0,15</b>	<b>0,014</b>	<b>-0,040</b>	<b>0,00</b>	<b>0,009</b>	<b>-0,043</b>	<b>0,00</b>	<b>0,009</b>

**Fonte dos dados básicos: PNAD - IBGE.**

**Tabela 13 – Resumo dos resultados do Exercício II com o modelo logit com desenho complexo – MG e SP**

Variável	2005			2006			2007		
	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão
Ano	0,097	0,00	0,034	0,105	0,00	0,031	0,764	0,06	0,111
Tratamento	1,337	0,59	0,718	0,543	0,08	0,189	1,166	0,36	0,197
<b>Interação</b>	<b>0,728</b>	<b>0,56</b>	<b>0,391</b>	<b>1,085</b>	<b>0,82</b>	<b>0,398</b>	<b>0,430</b>	<b>0,00</b>	<b>0,092</b>

**Fonte dos dados básicos: PNAD - IBGE.**

<sup>11</sup> No modelo de probabilidade linear foram observadas as seguintes mudanças: no período 2005 a variável sexo do chefe do domicílio deixou de ser estatisticamente significativa; no período 2006 as variáveis raça do chefe do domicílio e se no domicílio há eletricidade deixam de ser estatisticamente significativas e a renda domiciliar per capita passa a ser; no período 2007 as variáveis tipo de família na qual o indivíduo está inserido e se no domicílio há eletricidade deixam de ser estatisticamente significativas. No modelo logit foram observadas as seguintes mudanças: no período 2006 a variável raça do chefe do domicílio deixa de ser estatisticamente significativa; no período 2007 a variável tipo de família na qual o indivíduo está inserido deixa de ser estatisticamente significativa.

### C – Exercício III

No exercício cuja variável de resposta é a progressão por série (PPS), no modelo de probabilidade linear (TAB. 14), a variável interação é estatisticamente significativa somente para  $e_3$  e  $e_4$  e o sinal do coeficiente é positivo. Poucas das variáveis de condição socioeconômica e de *background* familiar são estatisticamente significativas para  $e_0$  e  $e_1$ , mas esse número cresce para  $e_2$ ,  $e_3$  e  $e_4$ <sup>12</sup>.

**Tabela 14 – Resumo dos resultados do Exercício III com o modelo de probabilidade linear – MG e SP**

PPS	Variável	Coefficiente	P-valor	Desvio padrão
$e_0$	Ano	-0,024	0,00	0,006
	Tratamento	0,008	0,28	0,008
	<b>Interação</b>	<b>-0,016</b>	<b>0,14</b>	<b>0,011</b>
$e_1$	Ano	0,002	0,41	0,002
	Tratamento	-0,002	0,38	0,003
	<b>Interação</b>	<b>0,003</b>	<b>0,46</b>	<b>0,004</b>
$e_2$	Ano	0,005	0,23	0,004
	Tratamento	-0,018	0,00	0,005
	<b>Interação</b>	<b>0,009</b>	<b>0,20</b>	<b>0,007</b>
$e_3$	Ano	0,017	0,01	0,007
	Tratamento	-0,034	0,00	0,008
	<b>Interação</b>	<b>0,038</b>	<b>0,00</b>	<b>0,012</b>
$e_4$	Ano	0,042	0,00	0,012
	Tratamento	-0,080	0,00	0,015
	<b>Interação</b>	<b>0,046</b>	<b>0,03</b>	<b>0,021</b>

Fonte dos dados básicos: PNAD - IBGE.

No modelo logit (TAB. 15) a interação só é significativa para  $e_0$ . Esse resultado, além de ser diferente do resultado obtido com o modelo de probabilidade linear, é diferente do esperado pois a razão de chance desta variável é mais uma vez inferior a um mostrando que os indivíduos de MG tem menor chance de concluir a primeira série do ensino fundamental quando comparados aos indivíduos de São Paulo. Nada podemos inferir sobre as outras PPS

<sup>12</sup> No modelo de probabilidade linear, para  $e_0$  somente a escolaridade do chefe do domicílio é estatisticamente significativa; para  $e_1$  a raça do chefe do domicílio e se no domicílio há eletricidade são estatisticamente significativas; para  $e_2$  o sexo do indivíduo, a escolaridade do chefe do domicílio, se o domicílio fica na zona urbana, se no domicílio há água encanada e se no domicílio há eletricidade são estatisticamente significativas; para  $e_3$  a raça do indivíduo, o sexo do indivíduo, a renda domiciliar per capita, a escolaridade do chefe do domicílio e se o domicílio fica na zona urbana são estatisticamente significativas; para  $e_4$  o sexo do indivíduo, a renda domiciliar per capita, a escolaridade do chefe do domicílio, a raça do chefe do domicílio, se no domicílio há água encanada e se o domicílio fica na zona urbana são estatisticamente significativas.

pois nenhuma das suas razões de chance foi estatisticamente significativa impossibilitando conclusões a respeito do maior ponto de estrangulamento do ensino fundamental ( $e_4$ ). Os resultados obtidos para as variáveis de condições socioeconômicas e de *background* familiar são quase os mesmos dos obtidos no modelo de probabilidade linear<sup>13</sup>. Novamente é observado que o número destas variáveis que são estatisticamente significativas no modelo cresce de uma PPS para outra sugerindo que nas séries iniciais características sociais, situação econômica e características dos pais tem pouca ou quase nenhuma importância mas continuam sendo relevantes nas séries subsequentes.

**Tabela 15 – Resumo dos resultados do Exercício III com o modelo logit – MG e SP**

PPS	Variável	Razão de chance	P-valor	Desvio padrão
$e_0$	Ano	0,467	0,00	0,106
	Tratamento	1,611	0,13	0,501
	<b>Interação</b>	<b>0,531</b>	<b>0,08</b>	<b>0,190</b>
$e_1$	Ano	1,764	0,42	1,252
	Tratamento	0,688	0,48	0,367
	<b>Interação</b>	<b>1,581</b>	<b>0,67</b>	<b>1,711</b>
$e_2$	Ano	1,892	0,21	0,953
	Tratamento	0,344	0,00	0,117
	<b>Interação</b>	<b>0,990</b>	<b>0,99</b>	<b>0,582</b>
$e_3$	Ano	1,968	0,01	0,526
	Tratamento	0,614	0,01	0,113
	<b>Interação</b>	<b>1,583</b>	<b>0,20</b>	<b>0,571</b>
$e_4$	Ano	1,596	0,00	0,224
	Tratamento	0,584	0,00	0,067
	<b>Interação</b>	<b>1,160</b>	<b>0,43</b>	<b>0,218</b>

**Fonte dos dados básicos: PNAD - IBGE.**

Para corroborar os resultados encontrados foi feita a estimação por amostragem de desenho complexo também para os modelos estatísticos cuja variável resposta é PPS (TAB. 16 e TAB. 17). Nesse exercício a única mudança importante observada foi no modelo logit pois nesta

<sup>13</sup> No modelo de logit, para  $e_0$  somente a escolaridade do chefe do domicílio é estatisticamente significativa; para  $e_1$  a raça do chefe do domicílio, se no domicílio há coleta de lixo e se no domicílio há eletricidade são estatisticamente significativas; para  $e_2$  o sexo do indivíduo, tipo de família na qual o indivíduo está inserido, a escolaridade do chefe do domicílio, se o domicílio fica na zona urbana e se no domicílio há água encanada são estatisticamente significativas; para  $e_3$  o sexo do indivíduo, a renda domiciliar per capita, a escolaridade do chefe do domicílio, se o domicílio fica na região metropolitana e se o domicílio fica na zona urbana são estatisticamente significativas; para  $e_4$  os resultados são exatamente iguais aos do modelo de probabilidade linear.



nova estimação a razão de chance da variável interação não foi estatisticamente significativa para nenhuma PPS<sup>14</sup>.

Com os resultados dos três exercícios empíricos – que analisaram duas variáveis de resposta diferentes: matrícula e PPS – na tentativa de captar algum impacto da política de aumento da duração do ensino fundamental concomitantemente com a antecipação do ciclo de vida do estudante percebe-se que Minas Gerais não apresentou retorno positivo desta política educacional quando comparamos essa unidade da federação com São Paulo.

**Tabela 16 – Resumo dos resultados do Exercício III com o modelo de probabilidade linear com desenho complexo – MG e SP**

PPS	Variável	Coefficiente	P-valor	Desvio padrão
e <sub>0</sub>	Ano	-0,026	0,00	0,007
	Tratamento	0,008	0,15	0,006
	<b>Interação</b>	<b>-0,014</b>	<b>0,21</b>	<b>0,011</b>
e <sub>1</sub>	Ano	0,002	0,33	0,002
	Tratamento	-0,003	0,44	0,003
	<b>Interação</b>	<b>0,003</b>	<b>0,50</b>	<b>0,004</b>
e <sub>2</sub>	Ano	0,004	0,18	0,003
	Tratamento	-0,018	0,00	0,005
	<b>Interação</b>	<b>0,009</b>	<b>0,28</b>	<b>0,008</b>
e <sub>3</sub>	Ano	0,018	0,01	0,007
	Tratamento	-0,034	0,00	0,010
	<b>Interação</b>	<b>0,037</b>	<b>0,00</b>	<b>0,012</b>
e <sub>4</sub>	Ano	0,041	0,00	0,014
	Tratamento	-0,077	0,00	0,016
	<b>Interação</b>	<b>0,045</b>	<b>0,05</b>	<b>0,023</b>

**Fonte dos dados básicos: PNAD - IBGE.**

<sup>14</sup> Algumas mudanças foram observadas nas demais variáveis em ambos os modelos: no modelo de probabilidade linear para e<sub>0</sub> se no domicílio há eletricidade passa a ser a única variável estatisticamente significativa do modelo, para e<sub>1</sub> se no domicílio há coleta de lixo passa a ser a única variável estatisticamente significativa do modelo, para e<sub>2</sub> se no domicílio há eletricidade e se no domicílio há água encanada deixam de ser estatisticamente significativas, para e<sub>3</sub> se o domicílio está na região metropolitana passa a ser estatisticamente significativa e para e<sub>4</sub> se o domicílio está na zona urbana deixa de ser estatisticamente significativa; no modelo logit para e<sub>0</sub> nenhuma variável foi estatisticamente significativa, para e<sub>1</sub> se no domicílio há coleta de lixo deixa de ser estatisticamente significativa e renda domiciliar per capita passa a ser, para e<sub>2</sub> tipo de família na qual o indivíduo está inserido deixa de ser estatisticamente significativa, para e<sub>3</sub> se o domicílio está na região metropolitana deixa de ser estatisticamente significativa e para e<sub>4</sub> se o a escolaridade do chefe do domicílio deixa de ser estatisticamente significativa e o tipo de família na qual o indivíduo está inserido passa a ser. Para e<sub>2</sub> tipo de família na qual o indivíduo está inserido deixa de ser estatisticamente significativa.

Tabela 17 – Resumo dos resultados do Exercício III com o modelo logit com desenho complexo – MG e SP

PPS	Variável	Razão de chance	P-valor	Desvio padrão
e <sub>0</sub>	Ano	0,433	0,00	0,096
	Tratamento	1,582	0,16	0,515
	<b>Interação</b>	<b>0,569</b>	<b>0,13</b>	<b>0,210</b>
e <sub>1</sub>	Ano	1,845	0,37	1,257
	Tratamento	0,683	0,47	0,361
	<b>Interação</b>	<b>1,596</b>	<b>0,67</b>	<b>1,740</b>
e <sub>2</sub>	Ano	1,801	0,24	0,902
	Tratamento	0,356	0,00	0,123
	<b>Interação</b>	<b>1,015</b>	<b>0,98</b>	<b>0,623</b>
e <sub>3</sub>	Ano	2,106	0,01	0,600
	Tratamento	0,612	0,02	0,130
	<b>Interação</b>	<b>1,500</b>	<b>0,27</b>	<b>0,546</b>
e <sub>4</sub>	Ano	1,576	0,00	0,248
	Tratamento	0,595	0,00	0,072
	<b>Interação</b>	<b>1,174</b>	<b>0,45</b>	<b>0,247</b>

Fonte dos dados básicos: PNAD - IBGE.

#### 4.2.2 Minas Gerais como grupo de tratamento e Bahia como grupo de controle

##### A – Exercício I

Na análise na qual o grupo de controle é o Estado da Bahia, no modelo de probabilidade linear (TAB. 18), a variável de maior importância no modelo de diferenças em diferenças, a interação, não foi estatisticamente significativa para o primeiro período (2001-2005) mas foi estatisticamente significativa e apresentou sinal negativo para os demais períodos (2001-2006 e 2001-2007). Com relação às demais variáveis usadas para melhorar o ajuste do modelo e reduzir os fatores não explicados, poucas foram significativas no primeiro período de análise.

Nos demais períodos, que abarcam observações de indivíduos de séries posteriores o número de variáveis deste grupo que foram estatisticamente significativas aumentou.<sup>15</sup>

**Tabela 18 – Resumo dos resultados do Exercício I com o modelo de probabilidade linear – MG e BA**

Variável	2001-2005			2001-2006			2001-2007		
	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão
Ano	0,138	0,00	0,018	0,076	0,00	0,009	0,051	0,00	0,006
Tratamento	0,103	0,00	0,015	0,085	0,00	0,009	0,069	0,00	0,006
<b>Interação</b>	<b>-0,010</b>	<b>0,67</b>	<b>0,023</b>	<b>-0,036</b>	<b>0,00</b>	<b>0,012</b>	<b>-0,035</b>	<b>0,00</b>	<b>0,008</b>

**Fonte dos dados básicos: PNAD - IBGE.**

No modelo logit o resultado para a variável interação foi diferente (TAB. 19). Ela foi estatisticamente significativa no primeiro e no terceiro períodos analisados. No período 2001-2005 a razão de chance foi superior a um mostrando que os indivíduos do grupo de tratamento tinham maior probabilidade de estarem matriculados do que os indivíduos do grupo de controle depois que a intervenção foi feita. No entanto, no terceiro período a razão de chance da variável interação foi inferior a um mostrando que, nesse caso, a intervenção não melhorou o cenário das matrículas em séries posteriores no estado de Minas Gerais. O impacto da política aconteceu na primeira série mas não foi observado nas posteriores.

**Tabela 19 – Resumo dos resultados do Exercício I com o modelo logit – MG e BA**

Variável	2001-2005			2001-2006			2001-2007		
	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão
Ano	2,533	0,00	0,348	1,688	0,00	0,122	1,495	0,00	0,083
Tratamento	1,880	0,00	0,207	1,838	0,00	0,133	1,798	0,00	0,104
<b>Interação</b>	<b>3,552</b>	<b>0,00</b>	<b>1,157</b>	<b>0,925</b>	<b>0,50</b>	<b>0,106</b>	<b>0,823</b>	<b>0,03</b>	<b>0,072</b>

**Fonte dos dados básicos: PNAD - IBGE.**

<sup>15</sup> No modelo de probabilidade linear no período 2001-2005 as variáveis de características individuais, socioeconômicas e de *background* familiar que foram estatisticamente significativas foram tipo de família na qual o indivíduo está inserido, sexo do indivíduo, se no domicílio há esgotamento sanitário e se no domicílio há eletricidade. No período 2001-2006 mais duas variáveis foram significativas além daquelas do período anterior: raça do chefe do domicílio e se no domicílio há água encanada. No período 2001-2007 mais quatro variáveis foram significativas além daquelas do período 2001-2005: condição de ocupação do chefe do domicílio, escolaridade do chefe do domicílio, se o domicílio se localiza em região metropolitana e se no domicílio há água encanada. No modelo logit os resultados foram quase iguais aos obtidos com a estimação através do modelo de probabilidade linear. A única diferença é que no último período de análise as variáveis condição de ocupação do chefe do domicílio e escolaridade do chefe do domicílio não foram estatisticamente significativas no logit.

Na tentativa de melhorar a qualidade das estimações e corroborar os resultados foi feito um exercício com amostragem complexa. Mas mesmo com este controle não foi obtida nenhuma melhora significativa no desvio-padrão dos coeficientes nem no p-valor (TAB. 20 e TAB. 21). Não houve nenhuma mudança importante nos resultados obtidos através de estimação com amostra simples nem no modelo de probabilidade linear nem no modelo logit. Os resultados para a variável interação foram iguais e em cada período estimado uma variável de característica socioeconômica ou de *background* familiar deixou de ser significativa.<sup>16</sup> Este exercício mostrou um pequeno impacto da política, quando comparamos crianças de Minas Gerais com crianças da Bahia, logo no início da sua implementação na primeira série que não se estendeu às séries posteriores.

**Tabela 20 – Resumo dos resultados do Exercício I com o modelo de probabilidade linear com desenho complexo – MG e BA**

Variável	2001-2005			2001-2006			2001-2007		
	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão
Ano	0,137	0,00	0,025	0,076	0,00	0,014	0,051	0,00	0,010
Tratamento	0,103	0,00	0,019	0,085	0,00	0,012	0,069	0,00	0,009
<b>Interação</b>	<b>-0,010</b>	<b>0,73</b>	<b>0,028</b>	<b>-0,036</b>	<b>0,03</b>	<b>0,016</b>	<b>-0,035</b>	<b>0,00</b>	<b>0,012</b>

Fonte dos dados básicos: PNAD - IBGE.

**Tabela 21 – Resumo dos resultados do Exercício I com o modelo logit – MG e BA**

Variável	2001-2005			2001-2006			2001-2007		
	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão
Ano	2,514	0,00	0,503	1,690	0,00	0,168	1,496	0,00	0,120
Tratamento	1,876	0,00	0,234	1,844	0,00	0,157	1,798	0,00	0,134
<b>Interação</b>	<b>3,515</b>	<b>0,00</b>	<b>1,206</b>	<b>0,926</b>	<b>0,58</b>	<b>0,128</b>	<b>0,822</b>	<b>0,08</b>	<b>0,093</b>

Fonte dos dados básicos: PNAD - IBGE.

## B – Exercício II

<sup>16</sup> Nos dois primeiros períodos analisados no modelo de probabilidade linear a variável relativa ao esgotamento sanitário deixou de ser estatisticamente significativa, e no último período a variável condição de ocupação do chefe do domicílio deixou de ser estatisticamente significativa. Na estimação com o logit a variável tipo de família na qual o indivíduo está inserido deixou de ser significativa em todos os períodos analisados. E no período 2001-2006 a variável raça do chefe do domicílio também deixou de ser estatisticamente significativa.

No segundo exercício, usando o modelo de probabilidade linear (TAB. 22), é possível observar que a variável interação é estatisticamente significativa em dois períodos: 2005 e 2006 e em ambos o coeficiente tem sinal positivo. Poucas das variáveis de característica socioeconômica ou de *background* familiar foram estatisticamente significativas<sup>17</sup>.

**Tabela 22 – Resumo dos resultados do Exercício II com o modelo de probabilidade linear – MG e BA**

Variável	2005			2006			2007		
	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão
Ano	-0,174	0,00	0,016	-0,122	0,00	0,008	-0,066	0,00	0,007
Tratamento	0,040	0,03	0,018	0,000	0,97	0,008	0,008	0,27	0,007
<b>Interação</b>	<b>0,058</b>	<b>0,01</b>	<b>0,022</b>	<b>0,032</b>	<b>0,00</b>	<b>0,011</b>	<b>0,015</b>	<b>0,12</b>	<b>0,009</b>

**Fonte dos dados básicos: PNAD - IBGE.**

No modelo logit (TAB. 23) a variável interação só foi estatisticamente significativa no primeiro período analisado (2005). A razão de chance da variável interação neste período foi inferior a um mostrando que, neste caso, os indivíduos do grupo de tratamento tinham menor chance de estar matriculados do que os indivíduos do grupo de controle. Observa-se que este resultado é diferente do obtido no Exercício I e aponta para a ineficácia da política.

**Tabela 23 – Resumo dos resultados do Exercício II com o modelo logit – MG e BA**

Variável	2005			2006			2007		
	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão
Ano	0,221	0,00	0,037	0,128	0,00	0,022	0,359	0,00	0,043
Tratamento	6,011	0,00	2,651	1,418	0,19	0,373	1,349	0,06	0,217
<b>Interação</b>	<b>0,318</b>	<b>0,01</b>	<b>0,145</b>	<b>0,909</b>	<b>0,73</b>	<b>0,249</b>	<b>0,932</b>	<b>0,70</b>	<b>0,171</b>

**Fonte dos dados básicos: PNAD - IBGE.**

<sup>17</sup> Na estimação usando o modelo de probabilidade linear no período 2005 as variáveis de característica socioeconômica ou de *background* familiar estatisticamente significativas foram tipo de família na qual o indivíduo está inserido, sexo do indivíduo, escolaridade do chefe de domicílio, raça do chefe de domicílio e se no domicílio há eletricidade. No período 2006 só duas variáveis deste grupo foram significativas, se o domicílio fica localizado em zona urbana e se no domicílio há eletricidade. No período 2007 quatro variáveis deste grupo foram estatisticamente significativas: tipo de família na qual o indivíduo está inserido, sexo do indivíduo, raça do chefe de domicílio e renda domiciliar per capita. Na estimação usando o logit, com relação ao grupo de variáveis de característica socioeconômica ou de *background* familiar, foram obtidos quase os mesmos resultados do modelo de probabilidade linear. No período 2005 a variável condição de ocupação do chefe de domicílio também foi estatisticamente significativa, no período 2006 só a variável se no domicílio há eletricidade foi estatisticamente significativa e no período 2007 a variável tipo de família na qual o indivíduo está inserido não foi estatisticamente significativa. Assim como no exercício I, os resultados obtidos para este grupo de variáveis nos modelos de probabilidade linear e logit foram muito semelhantes.

No exercício II também foram feitas estimações dos dois modelos por amostragem de desenho complexo (TAB 24 e TAB. 25). Assim como no exercício I, os resultados obtidos para a variável interação não mudaram. Há poucas mudanças, pontuais, nas variáveis de característica socioeconômica ou de *background* familiar que não melhoram os resultados das estimações<sup>18</sup>. Desse modo, o Exercício II mostra que Minas Gerais não teve resultados melhores em sua matrícula ao antecipar a entrada das crianças no ensino fundamental quando comparada ao estado da Bahia.

**Tabela 24 – Resumo dos resultados do Exercício II com o modelo de probabilidade linear com desenho complexo – MG e BA**

Variável	2005			2006			2007		
	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão
Ano	-0,175	0,00	0,012	-0,122	0,00	0,010	-0,067	0,00	0,010
Tratamento	0,040	0,01	0,015	0,001	0,87	0,006	0,008	0,22	0,007
<b>Interação</b>	<b>0,062</b>	<b>0,00</b>	<b>0,015</b>	<b>0,032</b>	<b>0,01</b>	<b>0,013</b>	<b>0,014</b>	<b>0,25</b>	<b>0,013</b>

Fonte dos dados básicos: PNAD - IBGE.

**Tabela 25 – Resumo dos resultados do Exercício II com o modelo logit com desenho complexo – MG e BA**

Variável	2005			2006			2007		
	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão
Ano	0,219	0,00	0,049	0,128	0,00	0,024	0,358	0,00	0,056
Tratamento	5,908	0,00	2,925	1,470	0,17	0,412	1,369	0,09	0,251
<b>Interação</b>	<b>0,331</b>	<b>0,02</b>	<b>0,155</b>	<b>0,878</b>	<b>0,65</b>	<b>0,252</b>	<b>0,921</b>	<b>0,71</b>	<b>0,203</b>

Fonte dos dados básicos: PNAD - IBGE.

### C – Exercício III: PPS

A análise da estimação com a probabilidade de progressão por série no modelo de probabilidade linear (TAB. 26) mostra que houve impacto da política em  $e_0$ ,  $e_1$  e  $e_2$ . Nestas

<sup>18</sup> Para a estimação no modelo de probabilidade linear foram observadas as seguintes mudanças: no período 2005 a variável sexo do indivíduo deixa de ser estatisticamente significativa, no período 2006 só a variável sexo do indivíduo é estatisticamente significativa e no período 2007 a variável tipo de família no qual o indivíduo está inserido deixa de ser estatisticamente significativa. Para as estimação no modelo logit no período 2005 a variável condição de ocupação do chefe do domicílio deixa de ser estatisticamente significativa, no período 2006 só a variável sexo do indivíduo é estatisticamente significativa e no período 2007 a variável condição de ocupação do chefe do domicílio passa a ser estatisticamente significativa.

três PPS a variável interação foi estatisticamente significativa e os coeficientes apresentaram sinal negativo. Dentre as demais variáveis, de característica socioeconômica ou de *background* familiar, poucas foram estatisticamente significativas em  $e_0$ , mas o número destas variáveis que foram estatisticamente significativas subiu consideravelmente para cada PPS<sup>19</sup> mostrando que aspectos sociais, econômicos e familiares perderam sua influência nas séries iniciais mas continuam tendo importância nas séries subsequentes. Isso também pode ser observado nos exercícios anteriores cuja variável de resposta é a matrícula.

**Tabela 26 – Resumo dos resultados do Exercício III com o modelo de probabilidade linear – MG e BA**

PPS	Variável	Coefficiente	P-valor	Desvio padrão
$e_0$	Ano	-0,018	0,03	0,008
	Tratamento	0,014	0,07	0,008
	<b>Interação</b>	<b>-0,021</b>	<b>0,05</b>	<b>0,011</b>
$e_1$	Ano	0,021	0,00	0,006
	Tratamento	0,036	0,00	0,006
	<b>Interação</b>	<b>-0,017</b>	<b>0,04</b>	<b>0,008</b>
$e_2$	Ano	0,043	0,00	0,010
	Tratamento	0,047	0,00	0,009
	<b>Interação</b>	<b>-0,029</b>	<b>0,02</b>	<b>0,013</b>
$e_3$	Ano	0,042	0,00	0,014
	Tratamento	0,054	0,00	0,013
	<b>Interação</b>	<b>0,011</b>	<b>0,54</b>	<b>0,018</b>
$e_4$	Ano	0,098	0,00	0,021
	Tratamento	0,069	0,00	0,020
	<b>Interação</b>	<b>-0,013</b>	<b>0,64</b>	<b>0,027</b>

Fonte dos dados básicos: PNAD - IBGE.

No modelo logit (TAB. 27) a variável interação foi estatisticamente significativa para somente para  $e_0$  e  $e_3$  sendo que os resultados para as demais variáveis independentes foram quase

<sup>19</sup> No modelo de probabilidade linear, escolaridade do chefe do domicílio e raça do chefe do domicílio foram estatisticamente significativas para  $e_0$ ; sexo do indivíduo, escolaridade do chefe do domicílio, condição de ocupação do chefe do domicílio e se no domicílio há eletricidade foram estatisticamente significativas para  $e_1$ ; sexo do indivíduo, escolaridade do chefe do domicílio, renda domiciliar per capita, se no domicílio há esgotamento sanitário e se no domicílio há eletricidade foram estatisticamente significativas para  $e_2$ ; sexo do indivíduo, tipo de família na qual o indivíduo está inserido, escolaridade do chefe do domicílio, sexo do chefe do domicílio, raça do chefe do domicílio, renda domiciliar per capita, se o domicílio fica localizado em região metropolitana e se no domicílio há eletricidade foram estatisticamente significativas para  $e_3$ ; sexo do indivíduo, escolaridade do chefe do domicílio, raça do chefe do domicílio, renda domiciliar per capita, se no domicílio há água encanada, se no domicílio há esgotamento sanitário e se no domicílio há eletricidade foram estatisticamente significativas para  $e_4$ . No modelo logit os resultados foram quase iguais, somente algumas variáveis deixaram de ser estatisticamente significativas: para  $e_0$  a escolaridade do chefe do domicílio; para  $e_1$  se no domicílio há eletricidade e sexo do indivíduo; para  $e_3$  a renda domiciliar per capita; e para  $e_4$  se no domicílio há água encanada.

iguais àqueles obtidos pela estimação através do modelo de probabilidade linear. A razão de chance de  $e_0$  foi inferior a um apontando para ineficácia da política com relação a probabilidade de concluir a primeira série. Já a razão de chance de  $e_3$  foi superior a um sinalizando que a política pode ter ajudado na progressão escolar da quarta para a quinta série. Mas no que diz respeito ao principal gargalo do ensino fundamental, não há sinais de eficácia da intervenção.

**Tabela 27 – Resumo dos resultados do Exercício III com o modelo logit – MG e BA**

PPS	Variável	Razão de chance	P-valor	Desvio padrão
$e_0$	Ano	0,601	0,03	0,144
	Tratamento	2,055	0,02	0,622
	<b>Interação</b>	<b>0,421</b>	<b>0,02</b>	<b>0,153</b>
$e_1$	Ano	1,734	0,04	0,471
	Tratamento	4,860	0,00	1,808
	<b>Interação</b>	<b>1,753</b>	<b>0,50</b>	<b>1,452</b>
$e_2$	Ano	1,825	0,00	0,351
	Tratamento	2,111	0,00	0,434
	<b>Interação</b>	<b>1,015</b>	<b>0,97</b>	<b>0,368</b>
$e_3$	Ano	1,451	0,01	0,216
	Tratamento	1,555	0,00	0,234
	<b>Interação</b>	<b>2,099</b>	<b>0,01</b>	<b>0,590</b>
$e_4$	Ano	1,664	0,00	0,199
	Tratamento	1,350	0,01	0,145
	<b>Interação</b>	<b>1,115</b>	<b>0,53</b>	<b>0,194</b>

Fonte dos dados básicos: PNAD - IBGE.

**Tabela 28 – Resumo dos resultados do Exercício III com o modelo de probabilidade linear com desenho complexo – MG e BA**

PPS	Variável	Coefficiente	P-valor	Desvio padrão
$e_0$	Ano	-0,019	0,05	0,010
	Tratamento	0,016	0,01	0,006
	<b>Interação</b>	<b>-0,021</b>	<b>0,10</b>	<b>0,013</b>
$e_1$	Ano	0,022	0,02	0,010
	Tratamento	0,036	0,00	0,007
	<b>Interação</b>	<b>-0,017</b>	<b>0,10</b>	<b>0,010</b>
$e_2$	Ano	0,042	0,00	0,013
	Tratamento	0,050	0,00	0,011
	<b>Interação</b>	<b>-0,030</b>	<b>0,04</b>	<b>0,015</b>
$e_3$	Ano	0,043	0,03	0,020
	Tratamento	0,052	0,00	0,016
	<b>Interação</b>	<b>0,010</b>	<b>0,64</b>	<b>0,022</b>
$e_4$	Ano	0,100	0,00	0,026
	Tratamento	0,068	0,00	0,023
	<b>Interação</b>	<b>-0,015</b>	<b>0,63</b>	<b>0,032</b>

Fonte dos dados básicos: PNAD - IBGE.



Ao fazer as mesmas estimações para amostras de desenho complexo o resultado não muda para a variável de maior interesse (TAB. 28 e TAB. 29). Somente há mudanças pontuais nas variáveis de característica socioeconômica ou de *background* familiar que foram estatisticamente significativas<sup>20</sup>, corroborando o impacto restrito da política sobre apenas uma progressão por série:  $e_3$ .

**Tabela 29 – Resumo dos resultados do Exercício III com o modelo logit com desenho complexo – MG e BA**

PPS	Variável	Razão de chance	P-valor	Desvio padrão
$e_0$	Ano	0,588	0,04	0,150
	Tratamento	2,149	0,02	0,676
	<b>Interação</b>	<b>0,427</b>	<b>0,03</b>	<b>0,166</b>
$e_1$	Ano	1,798	0,04	0,524
	Tratamento	4,562	0,00	1,994
	<b>Interação</b>	<b>1,728</b>	<b>0,52</b>	<b>1,478</b>
$e_2$	Ano	1,821	0,00	0,350
	Tratamento	2,136	0,00	0,442
	<b>Interação</b>	<b>0,997</b>	<b>1,00</b>	<b>0,399</b>
$e_3$	Ano	1,491	0,02	0,258
	Tratamento	1,461	0,02	0,236
	<b>Interação</b>	<b>2,076</b>	<b>0,01</b>	<b>0,587</b>
$e_4$	Ano	1,679	0,00	0,229
	Tratamento	1,345	0,01	0,159
	<b>Interação</b>	<b>1,097</b>	<b>0,64</b>	<b>0,217</b>

**Fonte dos dados básicos: PNAD - IBGE.**

Os exercícios que compararam Minas Gerais a Bahia mostraram impactos positivos da política tanto na matrícula quanto na progressão por série (somente  $e_3$ ). Mas esses impactos foram muito limitados indicando que a eficácia da política educacional também é questionável quando comparamos estes dois estados.

<sup>20</sup> Na estimação usando o modelo de probabilidade linear com amostragem de desenho complexo foram observadas as seguintes diferenças com relação às estimações com amostragem simples: em  $e_0$  a escolaridade do chefe deixou de ser estatisticamente significativa; em  $e_1$  se no domicílio há eletricidade deixa de ser estatisticamente significativa e a renda domiciliar per capita passa a ser; em  $e_2$  se no domicílio há esgotamento sanitário e a renda domiciliar per capita deixam de ser estatisticamente significativas; em  $e_3$  tipo de família no qual o indivíduo está inserido deixa de ser estatisticamente significativa; e em  $e_4$  se no domicílio há água encanada deixa de ser estatisticamente significativa. Ao fazer a mesma comparação entre os modelos de regressão logit observa-se que em  $e_0$  o sexo do indivíduo deixa de ser estatisticamente significativo e em  $e_3$  o tipo de família no qual o indivíduo está inserido e no domicílio há eletricidade deixam de ser estatisticamente significativas.

## 5 CONSIDERAÇÕES FINAIS

Na tentativa de melhorar o desempenho dos estudantes no ensino fundamental foi instituído um programa de ampliação da sua duração antecipando a entrada do estudante neste nível de ensino. Se antes a matrícula era obrigatória aos sete anos de idade ela passa a ser obrigatória aos seis anos para iniciar mais cedo o ciclo de vida escolar das crianças. Essa antecipação foi analisada neste trabalho através da matrícula e da probabilidade de progressão por série, comparando o estado de Minas Gerais com os estados de São Paulo e Bahia. Na comparação com São Paulo, Minas Gerais não apresentou nenhum resultado positivo em nenhuma das duas variáveis analisadas. Pelo contrário, em alguns exercícios verificou-se que o impacto da política foi negativo sobre a matrícula e sobre a progressão por série. E na comparação entre Minas Gerais e Bahia só houve resultados positivos sobre a matrícula na primeira série e sobre a probabilidade de concluir o quarto ano dado que concluiu o terceiro. As pequenas diferenças entre os resultados obtidos quando comparamos MG com SP e os resultados quando comparamos MG com BA sugerem que essa política educacional não foi eficaz nos períodos analisados.

Esgotar a análise do ensino fundamental de nove anos não é o objetivo deste trabalho. O intuito é começar a exploração deste assunto e estimular mais exercícios e debates. É possível que ainda seja cedo para tentar mensurar os efeitos da intervenção e o retorno positivo pode vir depois de mais tempo de consolidação da mesma. Também é possível que este impacto esteja consolidado em outro tipo de habilidade que não é captada pela matrícula ou pela probabilidade de progressão por série. Mas sem dúvida, os resultados aqui obtidos apontam para a necessidade de realizar mais estudos que os corroborem ou os refutem, e ao mesmo tempo, apontam para a necessidade de verificar a possibilidade de que essa política não seja uma alternativa eficaz à oferta de educação infantil universalizada.

Os resultados deste trabalho não criticam a antecipação do ciclo de vida do estudante e sim questionam a eficiência desta antecipação, através da antecipação da entrada no ensino fundamental. A literatura revisada aqui deixa bem claro que quanto mais cedo a criança tem acesso à educação melhor será seu desempenho escolar e melhor será o desenvolvimento de todas as suas habilidades. A questão é: se esta melhora no desempenho não é conseguida

colocando a criança mais cedo no ensino fundamental há grande probabilidade de que a saída seja a educação infantil. Um ano a mais no ensino fundamental não é capaz de suprir a falta de acesso à educação infantil. Percebe-se que investir na educação infantil pode ser muito mais eficaz até mesmo no retorno do investimento que é feito no próprio ensino fundamental. Falta direcionar melhor as políticas educacionais voltando-as para o aumento da oferta de vagas na educação infantil sem diminuir a qualidade e aumentar também a qualidade do ensino fundamental que nem carecia de ampliação e sim de aprimoramento e reformulação.

## REFERÊNCIAS BIBLIOGRÁFICAS

ATHEY, Susan; IMBENS, Guido. **Identification and Inference in Nonlinear Difference-in-Differences Models**. Econometric Society, vol. 74(2), pages 431-497.

ABADIE, Alberto. **Semiparametric Difference-in-Differences Estimators**. *Review of Economic Studies* (2005) nº 72, p. 1–19.

AGRESTI, Alan. **Categorical data analysis**. John Wiley & Sons, 1990. Chapter 4: Models for binary response variables.

BARROS, Ricardo Paes de; MENDONÇA, Rosane; SANTOS, Daniel Domingues dos; QUINTAES, Giovani. Determinantes do Desempenho Educacional do Brasil. **Pesquisa e Planejamento Econômico**. Rio de Janeiro, v. 31, n. 1, p.1-42, abril. 2001.

BERTRAND ,Marianne; DUFLO ,Esther; MULLAINATHAN, Sendhil. 2004. How Much Should We Trust Differences-in-Differences Estimates?, **The Quarterly Journal of Economics**, MIT Press, vol. 119(1), pages 249-275, February.

BLUNDELL, Richard; MaCURDY, Thomas; Labor Supply: A Review of alternative approaches. Cap. 5. p 1607-1617. **Handbook of Labor Economics**, ASHENFELTER, O.;CARD, D. v. 3, 1999, Elsevier Science B.V.

BRASIL. Ministério da Educação. **Lei nº 9.394**, de 20 de dezembro de 1996.

BRASIL. Ministério da Educação. **Lei nº 11.274**, de 06 de fevereiro de 2006.

BRASIL. Ministério da Educação . Secretaria de Educação Básica. **Ensino Fundamental de nove anos – orientação para a inclusão da criança de seis anos de idade**. 2 ed. Brasília, 2007.

CEALE (Centro de Alfabetização, Leitura e Escrita), FAE - UFMG. **Orientações para a organização do ciclo inicial de alfabetização – Caderno 1**. Belo Horizonte: Secretaria de Estado da Educação de Minas Gerais, 2004.

CUNHA, Flavio; HECKMAN, James J.; **Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation**. Chicago, 2006.

CUNHA, Flavio; HECKMAN, James J.; The technology of skill formation. **Seminário sobre Educação de Primeira Infância**. Rio de Janeiro, novembro, 2005.

CUNHA, Flavio; HECKMAN, James J.; LOCHNER, Lance; MASTEROV Dimitriy V. **Interperiting the Evidence on Life Cycle Skill Formation**. NBER Working Paper No. 11331, Cambridge, 2005.

DOLTON, Peter; **Treatment effects Models – Strategies for identification: Diff-in-diff, LATE, regression discontinuity design lecture**. Março de 2007.

DORIS R. Entwisle; ALEXANDER Karl L. Alexander; Entry Into School: The Beginning School Transition and Educational Stratification in the United States. **Annual Review of Sociology**, Vol. 19. (1993), pp. 401-423.

DORIS R. Entwisle; ALEXANDER Karl L.; Facilitating the Transition to First Grade: The Nature of Transition and Research on Factors Affecting It. **The Elementary School Journal**, Vol. 98, No. 4, Special Issue: Transitions. (Mar., 1998), pp.351-364.

FILGUEIRAS, K. F. . (Con)fusões entre alfabetização e letramento: as dificuldades de aprendizagem no processo de alfabetização numa escola pública e numa escola privada.. In: **VI Encontro de Pesquisa em Educação da Região Sudeste**, 2004, Rio de Janeiro.

GUJARATI, Dadomar N. **Econometria básica**. Cap. 3, p. 42-58. 3 ed., São Paulo: Pearson Makron Books, 2000.

HECKMAN, James J.; MASTEROV, Dimitry V.; The Productivity Argument for Investing in Young Children. **Working Paper 5, Invest in Kids Working Group - Committee for Economic Development**. October, 2004

IMBENS, Guido; WOOLDRIDGE, Jeffrey. **What's new in econometrics?** Lecture 10 from Summer Institute 2007.

MENDONÇA, Rosane S. P. de ; BARROS, Ricardo Paes de . Uma avaliação dos custos e benefícios da educação pré-escolar no Brasil. **A Primeira Infância no Desenvolvimento Sustentável**, 1999, Porto Alegre.

MINAS GERAIS. Governo do Estado. **Decreto nº 43.506**, de 06 de agosto de 2003.

MINAS GERAIS. Governo do Estado. **Resolução nº 469**, de 23 de dezembro de 2003.

NGARUIYA, Samuel. **Pre-School Education and School Readiness: the Kenya Experience**. Ade Biennale on Education in Africa, Gabon, 2006

OLIVEIRA, João Batista Araújo e; SILVA, Luiz Carlos Faria da. **O impacto das séries iniciais: educação infantil, analfabetismo funcional e equidade**. Anais do Seminário sobre Educação e Equidade, IETS, 2006.

POWERS, Daniel; XIE, Yu. **Statistical Methods for categorical data analysis**. Academic Press, 2000. Cap. 1 -3.

RIOS-NETO, Eduardo Luiz Gonçalves. O método Probabilidade de Progressão por Série. **Introdução à demografia da educação**. Associação Brasileira de Estudos Populacionais – ABEP, Campinas, 2004, p. 145-158.

SILVA, Ceris S. R.; CASTANHEIRA, Maria L.; Ações para a melhoria da qualidade da alfabetização na rede estadual e ensino de Minas Gerais. **Anais do VII Encontro de extensão da Universidade Federal de Minas Gerais**. Belo Horizonte, 2004.

**APÊNDICE A – Tabelas com os resultados completos das regressões**

**Tabela A 1 – Exercício I – Matrícula - Modelo de probabilidade linear – MG e SP**

Variável	2001-2005			2001-2006			2001-2007		
	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão
Ano	0,120	0,00	0,012	0,064	0,00	0,006	0,048	0,00	0,004
Tratamento	-0,014	0,23	0,012	-0,012	0,10	0,007	-0,006	0,29	0,005
Interação	0,008	0,67	0,020	-0,023	0,03	0,011	-0,031	0,00	0,007
Casal com mais de um filho	0,066	0,20	0,051	0,049	0,07	0,027	0,040	0,02	0,017
Mãe com filhos	0,069	0,19	0,052	0,045	0,10	0,027	0,039	0,03	0,018
Outros tipos de família	0,022	0,70	0,058	0,036	0,25	0,031	0,033	0,10	0,020
Raça	-0,002	0,82	0,010	-0,003	0,57	0,006	0,000	0,96	0,004
Sexo	0,008	0,35	0,009	0,002	0,66	0,005	0,007	0,04	0,003
Condição de ocupação do chefe do domicílio	0,004	0,57	0,006	0,009	0,01	0,003	-0,002	0,45	0,002
Escolaridade do chefe do domicílio	0,000	0,88	0,001	0,001	0,08	0,001	0,001	0,24	0,000
Raça do chefe do domicílio	0,001	0,94	0,010	-0,003	0,54	0,005	0,000	0,96	0,004
Sexo do chefe do domicílio	0,014	0,24	0,012	-0,003	0,65	0,006	0,002	0,62	0,004
Renda domiciliar per capita	0,001	0,83	0,003	0,001	0,65	0,002	0,003	0,03	0,001
Domicílio fica em área metropolitana	0,009	0,37	0,010	0,018	0,00	0,005	0,006	0,14	0,004
Domicílio fica na zona urbana	-0,050	0,02	0,022	-0,027	0,02	0,012	-0,010	0,25	0,009
No domicílio há água encanada	0,024	0,30	0,023	0,030	0,02	0,013	0,025	0,02	0,011
No domicílio há coleta de lixo	0,007	0,68	0,017	-0,011	0,24	0,009	-0,007	0,32	0,007
No domicílio há esgotamento sanitário	-0,016	0,22	0,013	-0,010	0,18	0,007	-0,007	0,17	0,005
No domicílio há eletricidade	0,049	0,24	0,041	0,027	0,26	0,024	0,003	0,94	0,034
R2		0,044			0,158			0,009	
R2 ajustado		0,040			0,144			0,009	
Nº de observações		4.303			12.918			23.158	

Fonte dos dados básicos: PNAD - IBGE.

**TABELA A 2 – Exercício I – Matrícula - Modelo logit – MG e SP**

Variável	2001-2005			2001-2006			2001-2007		
	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão
Ano	11,802	0,00	3,888	2,639	0,00	0,272	2,267	0,00	0,178
Tratamento	0,885	0,30	0,104	0,896	0,17	0,072	0,943	0,37	0,062
Interação	0,760	0,54	0,337	0,594	0,00	0,081	0,541	0,00	0,057
Casal com mais de um filho	1,773	0,27	0,918	1,600	0,10	0,459	1,605	0,04	0,360
Mãe com filhos	1,878	0,24	0,997	1,517	0,16	0,446	1,573	0,05	0,363
Outros tipos de família	1,119	0,85	0,655	1,362	0,36	0,459	1,451	0,16	0,383
Raça	0,965	0,78	0,121	0,958	0,55	0,069	0,993	0,90	0,055
Sexo	1,099	0,39	0,119	1,024	0,71	0,065	1,108	0,05	0,057
Condição de ocupação do chefe do domicílio	1,046	0,58	0,085	1,119	0,02	0,052	0,975	0,43	0,032
Escolaridade do chefe do domicílio	1,003	0,85	0,014	1,014	0,10	0,008	1,007	0,27	0,006
Raça do chefe do domicílio	1,012	0,92	0,123	0,956	0,53	0,069	0,994	0,91	0,055
Sexo do chefe do domicílio	1,188	0,26	0,181	0,967	0,66	0,074	1,031	0,61	0,063
Renda domiciliar per capita	1,009	0,83	0,041	1,013	0,61	0,026	1,062	0,02	0,028
Domicílio fica em área metropolitana	1,110	0,39	0,136	1,269	0,00	0,091	1,087	0,13	0,060
Domicílio fica na zona urbana	0,496	0,03	0,160	0,696	0,03	0,115	0,857	0,29	0,124
No domicílio há água encanada	1,383	0,28	0,414	1,464	0,02	0,228	1,427	0,02	0,213
No domicílio há coleta de lixo	1,092	0,64	0,202	0,877	0,25	0,099	0,905	0,30	0,087
No domicílio há esgotamento sanitário	0,810	0,18	0,127	0,875	0,17	0,084	0,895	0,15	0,069
No domicílio há eletricidade	1,640	0,21	0,646	1,321	0,26	0,328	1,029	0,95	0,427
Pseudo R2		0,088			0,029			0,019	
Log pseudolikelihood		-1240,79			-3733,45			-5999,79	
Nº de observações		4.303			12.918			23.158	

Fonte dos dados básicos: PNAD - IBGE.



**TABELA A 3 – Exercício I – Matrícula - Modelo de probabilidade linear com desenho complexo – MG e SP**

Variável	2001-2005			2001-2006			2001-2007		
	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão
Ano	0,120	0,00	0,010	0,065	0,00	0,007	0,048	-	-
Tratamento	-0,014	0,33	0,015	-0,012	0,15	0,009	-0,006	-	-
Interação	0,007	0,65	0,016	-0,023	0,03	0,011	-0,031	-	-
Tipo de família	-0,005	0,68	0,011	-0,003	0,60	0,006	0,000	-	-
Raça	-0,002	0,83	0,010	-0,003	0,62	0,006	0,000	-	-
Sexo	0,009	0,35	0,009	0,002	0,65	0,005	0,007	-	-
Condição de ocupação do chefe do domicílio	0,003	0,64	0,007	0,009	0,01	0,004	-0,002	-	-
Escolaridade do chefe do domicílio	0,000	0,81	0,001	0,001	0,05	0,001	0,001	-	-
Raça do chefe do domicílio	0,001	0,94	0,010	-0,004	0,52	0,006	0,000	-	-
Sexo do chefe do domicílio	0,017	0,14	0,011	-0,003	0,59	0,006	0,002	-	-
Renda domiciliar per capita	0,001	0,86	0,003	0,001	0,63	0,001	0,003	-	-
Domicílio fica em área metropolitana	0,009	0,39	0,010	0,018	0,00	0,005	0,006	-	-
Domicílio fica na zona urbana	-0,051	0,02	0,022	-0,027	0,01	0,010	-0,011	-	-
No domicílio há água encanada	0,023	0,35	0,025	0,029	0,02	0,013	0,025	-	-
No domicílio há coleta de lixo	0,007	0,66	0,017	-0,011	0,21	0,009	-0,007	-	-
No domicílio há esgotamento sanitário	-0,017	0,20	0,013	-0,010	0,19	0,008	-0,007	-	-
No domicílio há eletricidade	0,049	0,19	0,037	0,028	0,29	0,027	0,003	-	-
R2		0,043			0,016			0,009	
Nº de observações		4.290			12.905			23.153	
Nº de estratos		199			242			252	
Nº de unidades primárias de amostra		1.122			1.541			1.789	

Fonte dos dados básicos: PNAD - IBGE.

**TABELA A 4 – Exercício I – Matrícula - Modelo logit com desenho complexo – MG e SP**

Variável	2001-2005			2001-2006			2001-2007		
	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão
Ano	11,795	0,00	3,869	2,658	0,00	0,294	2,262	-	-
Tratamento	0,888	0,33	0,109	0,894	0,19	0,076	0,942	-	-
Interação	0,743	0,49	0,319	0,590	0,00	0,087	0,541	-	-
Tipo de família	0,949	0,69	0,127	0,968	0,63	0,066	0,998	-	-
Raça	0,965	0,77	0,115	0,961	0,59	0,071	0,994	-	-
Sexo	1,100	0,39	0,121	1,025	0,70	0,065	1,109	-	-
Condição de ocupação do chefe do domicílio	1,044	0,63	0,092	1,120	0,02	0,053	0,973	-	-
Escolaridade do chefe do domicílio	1,004	0,78	0,014	1,015	0,06	0,008	1,007	-	-
Raça do chefe do domicílio	1,014	0,91	0,130	0,953	0,51	0,070	0,993	-	-
Sexo do chefe do domicílio	1,232	0,15	0,180	0,959	0,59	0,075	1,029	-	-
Renda domiciliar per capita	1,007	0,86	0,042	1,012	0,66	0,027	1,061	-	-
Domicílio fica em área metropolitana	1,114	0,39	0,138	1,269	0,00	0,092	1,088	-	-
Domicílio fica na zona urbana	0,492	0,02	0,148	0,694	0,01	0,098	0,852	-	-
No domicílio há água encanada	1,377	0,31	0,434	1,463	0,01	0,220	1,428	-	-
No domicílio há coleta de lixo	1,105	0,62	0,225	0,872	0,26	0,105	0,905	-	-
No domicílio há esgotamento sanitário	0,809	0,18	0,129	0,876	0,20	0,090	0,896	-	-
No domicílio há eletricidade	1,618	0,13	0,513	1,339	0,24	0,333	1,037	-	-
Nº de observações		4.290			12.905			23.153	
Nº de estratos		199			242			252	
Nº de unidades primárias de amostra		1.122			1.541			1.789	

Fonte dos dados básicos: PNAD - IBGE.

**TABELA A 5 – Exercício I – Matrícula - Modelo de probabilidade linear – MG e BA**

Variável	2001-2005			2001-2006			2001-2007		
	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão
Ano	0,138	0,00	0,018	0,076	0,00	0,009	0,051	0,00	0,006
Tratamento	0,103	0,00	0,015	0,085	0,00	0,009	0,069	0,00	0,006
Interação	-0,010	0,67	0,023	-0,036	0,00	0,012	-0,035	0,00	0,008
Casal com mais de um filho	-0,039	0,39	0,046	0,002	0,94	0,025	-0,012	0,48	0,017
Mãe com filhos	-0,031	0,51	0,047	-0,010	0,71	0,026	-0,020	0,26	0,018
Outros tipos de família	-0,112	0,04	0,055	-0,064	0,03	0,030	-0,034	0,09	0,020
Raça	0,006	0,63	0,013	0,002	0,81	0,007	-0,001	0,90	0,005
Sexo	0,026	0,02	0,011	0,014	0,02	0,006	0,017	0,00	0,004
Condição de ocupação do chefe do domicílio	0,003	0,67	0,008	0,003	0,50	0,004	-0,005	0,08	0,003
Escolaridade do chefe do domicílio	0,002	0,17	0,002	0,000	0,65	0,001	0,001	0,08	0,001
Raça do chefe do domicílio	-0,003	0,82	0,013	-0,015	0,03	0,007	0,000	0,92	0,005
Sexo do chefe do domicílio	0,002	0,88	0,015	0,001	0,84	0,008	0,003	0,59	0,005
Renda domiciliar per capita	-0,001	0,85	0,006	0,001	0,68	0,003	0,002	0,44	0,002
Domicílio fica em área metropolitana	-0,010	0,48	0,014	-0,002	0,80	0,008	-0,011	0,05	0,005
Domicílio fica na zona urbana	0,020	0,32	0,020	-0,004	0,71	0,012	0,000	0,97	0,008
No domicílio há água encanada	0,014	0,44	0,019	0,023	0,03	0,011	0,027	0,00	0,007
No domicílio há coleta de lixo	0,017	0,29	0,016	0,014	0,13	0,009	0,006	0,36	0,006
No domicílio há esgotamento sanitário	-0,029	0,06	0,015	-0,015	0,08	0,008	-0,011	0,07	0,006
No domicílio há eletricidade	0,107	0,00	0,023	0,101	0,00	0,013	0,078	0,00	0,009
R2		0,071			0,032			0,022	
R2 ajustado		0,067			0,030			0,021	
Nº de observações		4.214			12.532			23.615	

Fonte dos dados básicos: PNAD - IBGE.

**TABELA A 6 – Exercício I – Matrícula - Modelo logit – MG e BA**

Variável	2001-2005			2001-2006			2001-2007		
	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão
Ano	2,533	0,00	0,348	1,688	0,00	0,122	1,495	0,00	0,083
Tratamento	1,880	0,00	0,207	1,838	0,00	0,133	1,798	0,00	0,104
Interação	3,552	0,00	1,157	0,925	0,50	0,106	0,823	0,03	0,072
Casal com mais de um filho	0,707	0,41	0,300	1,023	0,91	0,214	0,894	0,52	0,154
Mãe com filhos	0,762	0,53	0,333	0,926	0,72	0,201	0,829	0,29	0,147
Outros tipos de família	0,418	0,07	0,201	0,624	0,06	0,155	0,725	0,11	0,147
Raça	1,069	0,54	0,117	1,017	0,80	0,065	0,995	0,92	0,049
Sexo	1,244	0,02	0,112	1,126	0,03	0,061	1,183	0,00	0,050
Condição de ocupação do chefe do domicílio	1,028	0,68	0,068	1,024	0,53	0,039	0,954	0,08	0,026
Escolaridade do chefe do domicílio	1,019	0,15	0,013	1,003	0,67	0,008	1,010	0,08	0,006
Raça do chefe do domicílio	0,976	0,83	0,108	0,877	0,04	0,056	1,008	0,88	0,050
Sexo do chefe do domicílio	1,015	0,91	0,127	1,012	0,85	0,067	1,027	0,60	0,052
Renda domiciliar per capita	0,991	0,84	0,041	1,015	0,62	0,031	1,022	0,37	0,025
Domicílio fica em área metropolitana	0,918	0,45	0,103	0,986	0,83	0,065	0,894	0,03	0,046
Domicílio fica na zona urbana	1,157	0,35	0,180	0,967	0,72	0,091	1,002	0,98	0,075
No domicílio há água encanada	1,115	0,47	0,168	1,205	0,03	0,105	1,281	0,00	0,087
No domicílio há coleta de lixo	1,148	0,27	0,143	1,125	0,10	0,081	1,059	0,31	0,060
No domicílio há esgotamento sanitário	0,807	0,08	0,099	0,888	0,10	0,064	0,913	0,11	0,052
No domicílio há eletricidade	1,699	0,00	0,266	1,791	0,00	0,173	1,635	0,00	0,125
Pseudo R2		0,090			0,036			0,027	
Log pseudolikelihood		-1680,35			-4932,49			-8412,31	
Nº de observações		4.214			12.532			23.615	

**Fonte dos dados básicos: PNAD - IBGE.**

**TABELA A 7 – Exercício I – Matrícula - Modelo de probabilidade linear com desenho complexo – MG e BA**

Variável	2001-2005			2001-2006			2001-2007		
	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão
Ano	0,137	0,00	0,025	0,076	0,00	0,014	0,051	0,00	0,010
Tratamento	0,103	0,00	0,019	0,085	0,00	0,012	0,069	0,00	0,009
Interação	-0,010	0,73	0,028	-0,036	0,03	0,016	-0,035	0,00	0,012
Tipo de família	-0,016	0,22	0,013	-0,019	0,01	0,007	-0,010	0,03	0,004
Raça	0,006	0,64	0,013	0,002	0,84	0,008	-0,001	0,91	0,005
Sexo	0,026	0,02	0,011	0,014	0,03	0,006	0,017	0,00	0,004
Condição de ocupação do chefe do domicílio	0,003	0,68	0,008	0,003	0,49	0,004	-0,005	0,11	0,003
Escolaridade do chefe do domicílio	0,002	0,17	0,002	0,000	0,58	0,001	0,001	0,08	0,001
Raça do chefe do domicílio	-0,003	0,80	0,013	-0,016	0,04	0,008	0,000	0,92	0,005
Sexo do chefe do domicílio	0,012	0,40	0,014	0,003	0,71	0,007	0,003	0,55	0,005
Renda domiciliar per capita	-0,001	0,78	0,004	0,001	0,69	0,003	0,002	0,41	0,002
Domicílio fica em área metropolitana	-0,010	0,46	0,014	-0,002	0,80	0,008	-0,011	0,06	0,006
Domicílio fica na zona urbana	0,021	0,43	0,027	-0,004	0,75	0,013	0,000	0,98	0,010
No domicílio há água encanada	0,013	0,55	0,022	0,023	0,04	0,012	0,027	0,00	0,008
No domicílio há coleta de lixo	0,018	0,35	0,019	0,013	0,17	0,010	0,006	0,39	0,007
No domicílio há esgotamento sanitário	-0,028	0,06	0,015	-0,014	0,11	0,009	-0,011	0,09	0,006
No domicílio há eletricidade	0,107	0,00	0,031	0,101	0,00	0,019	0,078	0,00	0,015
R2		0,069			0,031			0,022	
Nº de observações		4.204			12.524			23.615	
Nº de estratos		151			168			174	
Nº de unidades primárias de amostra		901			1.252			1.471	

Fonte dos dados básicos: PNAD - IBGE.

**TABELA A 8 – Exercício I – Matrícula - Modelo logit com desenho complexo – MG e BA**

Variável	2001-2005			2001-2006			2001-2007		
	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão
Ano	2,514	0,00	0,503	1,690	0,00	0,168	1,496	0,00	0,120
Tratamento	1,876	0,00	0,234	1,844	0,00	0,157	1,798	0,00	0,134
Interação	3,515	0,00	1,206	0,926	0,58	0,128	0,822	0,08	0,093
Tipo de família	0,884	0,23	0,090	0,853	0,00	0,046	0,912	0,02	0,037
Raça	1,067	0,57	0,123	1,015	0,83	0,069	0,995	0,92	0,052
Sexo	1,243	0,02	0,113	1,128	0,03	0,061	1,183	0,00	0,051
Condição de ocupação do chefe do domicílio	1,023	0,71	0,064	1,026	0,52	0,040	0,954	0,10	0,027
Escolaridade do chefe do domicílio	1,019	0,15	0,014	1,004	0,61	0,008	1,010	0,09	0,006
Raça do chefe do domicílio	0,974	0,82	0,111	0,876	0,05	0,058	1,007	0,88	0,050
Sexo do chefe do domicílio	1,097	0,44	0,131	1,023	0,72	0,064	1,029	0,56	0,051
Renda domiciliar per capita	0,990	0,83	0,045	1,013	0,68	0,031	1,022	0,42	0,028
Domicílio fica em área metropolitana	0,921	0,52	0,116	0,987	0,85	0,071	0,894	0,06	0,053
Domicílio fica na zona urbana	1,167	0,41	0,217	0,969	0,76	0,101	1,004	0,97	0,084
No domicílio há água encanada	1,096	0,56	0,174	1,205	0,04	0,108	1,280	0,00	0,093
No domicílio há coleta de lixo	1,158	0,32	0,171	1,123	0,15	0,090	1,059	0,36	0,067
No domicílio há esgotamento sanitário	0,813	0,10	0,102	0,891	0,14	0,070	0,913	0,16	0,059
No domicílio há eletricidade	1,698	0,00	0,267	1,789	0,00	0,186	1,636	0,00	0,147
Nº de observações		4.204			12.524			23.615	
Nº de estratos		151			168			174	
Nº de unidades primárias de amostra		901			1.252			1.471	

Fonte dos dados básicos: PNAD - IBGE.

**TABELA A 9 – Exercício II – Matrícula - Modelo de probabilidade linear – MG e SP**

Variável	2005			2006			2007		
	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão
Ano	-0,095	0,00	0,011	-0,049	0,00	0,005	-0,009	0,06	0,005
Tratamento	0,013	0,42	0,015	-0,002	0,75	0,006	0,005	0,34	0,006
Interação	-0,023	0,21	0,018	-0,040	0,00	0,009	-0,043	0,00	0,008
Casal com mais de um filho	-0,046	0,30	0,044	-0,004	0,84	0,022	0,047	0,01	0,017
Mãe com filhos	-0,053	0,25	0,046	-0,010	0,65	0,023	0,029	0,12	0,018
Outros tipos de família	-0,026	0,62	0,051	-0,010	0,69	0,025	0,046	0,02	0,020
Raça	-0,003	0,80	0,011	-0,004	0,50	0,005	-0,001	0,78	0,005
Sexo	0,011	0,16	0,008	-0,001	0,77	0,004	0,007	0,05	0,004
Condição de ocupação do chefe do domicílio	0,000	1,00	0,006	-0,008	0,01	0,003	0,000	0,95	0,003
Escolaridade do chefe do domicílio	-0,001	0,49	0,001	0,001	0,38	0,001	0,002	0,00	0,000
Raça do chefe do domicílio	0,033	0,00	0,011	0,010	0,06	0,005	0,008	0,08	0,005
Sexo do chefe do domicílio	0,023	0,11	0,014	0,010	0,14	0,007	0,006	0,34	0,006
Renda domiciliar per capita	0,003	0,26	0,003	0,002	0,17	0,001	0,004	0,01	0,001
Domicílio fica em área metropolitana	0,031	0,00	0,009	0,011	0,01	0,005	-0,006	0,17	0,004
Domicílio fica na zona urbana	0,001	0,96	0,020	-0,014	0,16	0,010	0,014	0,11	0,009
No domicílio há água encanada	-0,018	0,39	0,021	0,023	0,04	0,011	0,003	0,75	0,010
No domicílio há coleta de lixo	-0,003	0,83	0,016	-0,012	0,12	0,008	-0,008	0,28	0,007
No domicílio há esgotamento sanitário	-0,002	0,86	0,013	0,002	0,73	0,006	-0,005	0,42	0,006
No domicílio há eletricidade	0,008	0,86	0,042	0,039	0,08	0,022	0,046	0,04	0,022
R2		0,041			0,037			0,014	
R2 ajustado		0,037			0,034			0,012	
Nº de observações		4.191			8.640			11.167	

**Fonte dos dados básicos: PNAD - IBGE.**

**TABELA A 10 – Exercício II – Matrícula - Modelo logit – MG e SP**

Variável	2005			2006			2007		
	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão
Ano	0,096	0,00	0,033	0,103	0,00	0,029	0,767	0,07	0,110
Tratamento	1,338	0,59	0,716	0,521	0,05	0,176	1,158	0,38	0,192
Interação	0,708	0,53	0,387	1,140	0,71	0,407	0,435	0,00	0,088
Casal com mais de um filho	0,372	0,35	0,390	0,913	0,88	0,526	2,251	0,02	0,788
Mãe com filhos	0,326	0,29	0,347	0,771	0,66	0,457	1,438	0,35	0,555
Outros tipos de família	0,495	0,53	0,557	0,805	0,74	0,517	2,166	0,07	0,906
Raça	0,974	0,86	0,148	0,906	0,50	0,131	0,962	0,76	0,120
Sexo	1,177	0,18	0,143	0,975	0,82	0,111	1,210	0,05	0,119
Condição de ocupação do chefe do domicílio	1,004	0,97	0,096	0,810	0,01	0,063	0,999	0,99	0,069
Escolaridade do chefe do domicílio	0,986	0,44	0,018	1,010	0,50	0,016	1,045	0,00	0,013
Raça do chefe do domicílio	1,593	0,00	0,243	1,284	0,09	0,190	1,221	0,11	0,155
Sexo do chefe do domicílio	1,456	0,10	0,334	1,371	0,11	0,268	1,166	0,41	0,217
Renda domiciliar per capita	1,064	0,24	0,056	1,091	0,09	0,055	1,150	0,02	0,068
Domicílio fica em área metropolitana	1,606	0,00	0,216	1,377	0,02	0,180	0,854	0,14	0,091
Domicílio fica na zona urbana	1,018	0,95	0,305	0,712	0,19	0,186	1,355	0,15	0,287
No domicílio há água encanada	0,752	0,33	0,218	1,705	0,06	0,476	1,102	0,68	0,256
No domicílio há coleta de lixo	0,961	0,87	0,225	0,725	0,09	0,139	0,819	0,33	0,168
No domicílio há esgotamento sanitário	0,973	0,89	0,191	1,037	0,84	0,180	0,886	0,40	0,128
No domicílio há eletricidade	1,123	0,82	0,583	1,813	0,16	0,773	1,943	0,09	0,760
Pseudo R2		0,094			0,112			0,036	
Log pseudolikelihood		-1049,95			-1271,05			-1831,49	
Nº de observações		4.191			8.640			11.167	

**Fonte dos dados básicos: PNAD - IBGE.**



**TABELA A 11 – Exercício II – Matrícula - Modelo de probabilidade linear com desenho complexo – MG e SP**

Variável	2005			2006			2007		
	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão
Ano	-0,096	0,00	0,010	-0,049	0,00	0,005	-0,009	0,07	0,005
Tratamento	0,012	0,07	0,007	-0,002	0,65	0,004	0,006	0,25	0,005
Interação	-0,020	0,15	0,014	-0,040	0,00	0,009	-0,043	0,00	0,009
Tipo de família	-0,001	0,89	0,009	-0,004	0,44	0,005	-0,003	0,52	0,004
Raça	-0,003	0,76	0,011	-0,003	0,60	0,006	-0,001	0,86	0,005
Sexo	0,010	0,20	0,008	-0,001	0,72	0,004	0,007	0,07	0,004
Condição de ocupação do chefe do domicílio	0,001	0,92	0,007	-0,008	0,02	0,004	0,000	0,89	0,003
Escolaridade do chefe do domicílio	-0,001	0,54	0,001	0,001	0,33	0,001	0,002	0,00	0,000
Raça do chefe do domicílio	0,032	0,00	0,011	0,009	0,12	0,006	0,007	0,12	0,005
Sexo do chefe do domicílio	0,019	0,09	0,011	0,008	0,15	0,006	-0,003	0,53	0,005
Renda domiciliar per capita	0,003	0,22	0,002	0,001	0,03	0,001	0,004	0,00	0,001
Domicílio fica em área metropolitana	0,031	0,00	0,010	0,011	0,01	0,004	-0,005	0,23	0,004
Domicílio fica na zona urbana	0,003	0,88	0,021	-0,013	0,17	0,010	0,013	0,22	0,011
No domicílio há água encanada	-0,021	0,29	0,020	0,021	0,11	0,013	0,004	0,76	0,012
No domicílio há coleta de lixo	-0,003	0,84	0,017	-0,011	0,14	0,008	-0,008	0,41	0,009
No domicílio há esgotamento sanitário	-0,002	0,88	0,013	0,002	0,76	0,007	-0,004	0,49	0,006
No domicílio há eletricidade	0,010	0,83	0,045	0,040	0,25	0,035	0,045	0,15	0,031
R2		0,040			0,036			0,012	
Nº de observações		4.137			8.539			11.056	
Nº de estratos		143			151			151	
Nº de unidades primárias de amostra		949			1.078			1.098	

**Fonte dos dados básicos: PNAD - IBGE.**

**TABELA A 12 – Exercício II – Matrícula - Modelo logit com desenho complexo – MG e SP**

Variável	2005			2006			2007		
	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão
Ano	0,097	0,00	0,034	0,105	0,00	0,031	0,764	0,06	0,111
Tratamento	1,337	0,59	0,718	0,543	0,08	0,189	1,166	0,36	0,197
Interação	0,728	0,56	0,391	1,085	0,82	0,398	0,430	0,00	0,092
Tipo de família	0,972	0,84	0,133	0,906	0,43	0,113	0,938	0,52	0,094
Raça	0,966	0,81	0,142	0,923	0,59	0,138	0,970	0,81	0,123
Sexo	1,157	0,21	0,135	0,968	0,76	0,104	1,200	0,07	0,120
Condição de ocupação do chefe do domicílio	1,014	0,90	0,109	0,808	0,01	0,068	0,993	0,93	0,075
Escolaridade do chefe do domicílio	0,986	0,49	0,019	1,011	0,46	0,015	1,046	0,00	0,013
Raça do chefe do domicílio	1,570	0,00	0,228	1,248	0,15	0,191	1,205	0,13	0,147
Sexo do chefe do domicílio	1,360	0,10	0,255	1,291	0,12	0,213	0,915	0,50	0,120
Renda domiciliar per capita	1,057	0,27	0,053	1,082	0,09	0,050	1,147	0,03	0,070
Domicílio fica em área metropolitana	1,600	0,00	0,233	1,366	0,02	0,183	0,858	0,19	0,099
Domicílio fica na zona urbana	1,049	0,88	0,317	0,731	0,23	0,190	1,331	0,21	0,305
No domicílio há água encanada	0,725	0,29	0,219	1,630	0,09	0,470	1,100	0,72	0,296
No domicílio há coleta de lixo	0,964	0,88	0,241	0,745	0,18	0,163	0,820	0,40	0,194
No domicílio há esgotamento sanitário	0,977	0,90	0,189	1,037	0,85	0,196	0,897	0,47	0,134
No domicílio há eletricidade	1,160	0,80	0,673	1,841	0,20	0,873	1,948	0,09	0,754
Nº de observações	4.137			8.539			11.056		
Nº de estratos	143			151			151		
Nº de unidades primárias de amostra	949			1078			1098		

**Fonte dos dados básicos: PNAD - IBGE.**

**TABELA A 13 – Exercício II – Matrícula - Modelo de probabilidade linear – MG e BA**

Variável	2005			2006			2007		
	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão
Ano	-0,174	0,00	0,016	-0,122	0,00	0,008	-0,066	0,00	0,007
Tratamento	0,040	0,03	0,018	0,000	0,97	0,008	0,008	0,27	0,007
Interação	0,058	0,01	0,022	0,032	0,00	0,011	0,015	0,12	0,009
Casal com mais de um filho	-0,117	0,01	0,043	0,007	0,74	0,022	0,030	0,09	0,018
Mãe com filhos	-0,110	0,02	0,046	0,000	0,99	0,023	0,017	0,39	0,019
Outros tipos de família	-0,152	0,00	0,051	-0,013	0,63	0,026	0,018	0,39	0,021
Raça	-0,011	0,41	0,014	-0,002	0,81	0,007	-0,005	0,37	0,006
Sexo	0,023	0,03	0,010	0,008	0,13	0,005	0,012	0,01	0,005
Condição de ocupação do chefe do domicílio	0,007	0,36	0,008	-0,006	0,13	0,004	-0,003	0,45	0,003
Escolaridade do chefe do domicílio	0,003	0,03	0,002	0,001	0,20	0,001	0,001	0,40	0,001
Raça do chefe do domicílio	0,035	0,01	0,014	0,009	0,24	0,007	0,011	0,08	0,006
Sexo do chefe do domicílio	0,001	0,98	0,019	-0,004	0,64	0,009	0,000	0,99	0,008
Renda domiciliar per capita	-0,003	0,54	0,005	0,003	0,23	0,002	0,004	0,04	0,002
Domicílio fica em área metropolitana	0,009	0,49	0,013	0,009	0,20	0,007	-0,007	0,27	0,006
Domicílio fica na zona urbana	-0,023	0,24	0,020	-0,013	0,22	0,011	0,001	0,89	0,009
No domicílio há água encanada	0,018	0,32	0,018	0,007	0,50	0,010	0,007	0,47	0,009
No domicílio há coleta de lixo	0,018	0,24	0,015	0,006	0,47	0,008	0,004	0,55	0,007
No domicílio há esgotamento sanitário	-0,013	0,33	0,014	-0,002	0,75	0,007	0,001	0,91	0,006
No domicílio há eletricidade	0,081	0,00	0,024	0,042	0,00	0,013	0,006	0,63	0,013
R2		0,071			0,050			0,019	
R2 ajustado		0,067			0,048			0,017	
Nº de observações		4.103			8.337			10.782	

Fonte dos dados básicos: PNAD - IBGE.

**TABELA A 14 – Exercício II – Matrícula - Modelo logit – MG e BA**

Variável	2005			2006			2007		
	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão
Ano	0,221	0,00	0,037	0,128	0,00	0,022	0,359	0,00	0,043
Tratamento	6,011	0,00	2,651	1,418	0,19	0,373	1,349	0,06	0,217
Interação	0,318	0,01	0,145	0,909	0,73	0,249	0,932	0,70	0,171
Casal com mais de um filho	0,113	0,03	0,114	1,121	0,73	0,373	1,550	0,11	0,426
Mãe com filhos	0,120	0,04	0,122	0,986	0,97	0,350	1,240	0,48	0,379
Outros tipos de família	0,084	0,02	0,086	0,816	0,62	0,331	1,253	0,50	0,417
Raça	0,897	0,43	0,123	0,979	0,87	0,125	0,905	0,37	0,101
Sexo	1,237	0,04	0,124	1,150	0,13	0,106	1,229	0,01	0,103
Condição de ocupação do chefe do domicílio	1,075	0,35	0,082	0,917	0,17	0,059	0,960	0,48	0,055
Escolaridade do chefe do domicílio	1,033	0,04	0,017	1,016	0,25	0,014	1,009	0,43	0,011
Raça do chefe do domicílio	1,448	0,01	0,216	1,161	0,26	0,155	1,225	0,08	0,142
Sexo do chefe do domicílio	1,013	0,94	0,183	0,939	0,67	0,138	0,994	0,96	0,136
Renda domiciliar per capita	0,975	0,57	0,043	1,093	0,11	0,062	1,111	0,04	0,056
Domicílio fica em área metropolitana	1,097	0,45	0,136	1,174	0,15	0,132	0,886	0,23	0,090
Domicílio fica na zona urbana	0,824	0,27	0,145	0,823	0,22	0,130	1,026	0,88	0,174
No domicílio há água encanada	1,139	0,40	0,176	1,101	0,52	0,165	1,102	0,53	0,171
No domicílio há coleta de lixo	1,177	0,22	0,158	1,100	0,44	0,137	1,076	0,51	0,118
No domicílio há esgotamento sanitário	0,883	0,33	0,114	0,951	0,68	0,117	1,017	0,88	0,117
No domicílio há eletricidade	1,706	0,01	0,344	1,659	0,01	0,325	1,085	0,71	0,240
Pseudo R2		0,105			0,109			0,042	
Log pseudolikelihood		-1422,49			-1851,50			-2414,90	
Nº de observações		4.103			8.337			10.782	

Fonte dos dados básicos: PNAD - IBGE.

**TABELA A 15 – Exercício II – Matrícula - Modelo de probabilidade linear com desenho complexo – MG e BA**

Variável	2005			2006			2007		
	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão
Ano	-0,175	0,00	0,012	-0,122	0,00	0,010	-0,067	0,00	0,010
Tratamento	0,040	0,01	0,015	0,001	0,87	0,006	0,008	0,22	0,007
Interação	0,062	0,00	0,015	0,032	0,01	0,013	0,014	0,25	0,013
Tipo de família	-0,021	0,10	0,013	-0,007	0,27	0,006	-0,005	0,39	0,005
Raça	-0,013	0,36	0,014	-0,001	0,93	0,008	-0,006	0,37	0,006
Sexo	0,021	0,05	0,011	0,008	0,10	0,005	0,012	0,01	0,005
Condição de ocupação do chefe do domicílio	0,008	0,27	0,008	-0,007	0,18	0,005	-0,003	0,45	0,004
Escolaridade do chefe do domicílio	0,003	0,05	0,002	0,001	0,16	0,001	0,001	0,29	0,001
Raça do chefe do domicílio	0,033	0,02	0,014	0,008	0,26	0,007	0,011	0,06	0,006
Sexo do chefe do domicílio	0,019	0,28	0,017	-0,004	0,61	0,008	-0,005	0,48	0,007
Renda domiciliar per capita	-0,003	0,50	0,004	0,003	0,12	0,002	0,004	0,01	0,002
Domicílio fica em área metropolitana	0,008	0,51	0,012	0,009	0,17	0,006	-0,006	0,34	0,006
Domicílio fica na zona urbana	-0,024	0,25	0,021	-0,012	0,21	0,010	0,001	0,92	0,012
No domicílio há água encanada	0,019	0,36	0,020	0,005	0,58	0,009	0,006	0,61	0,011
No domicílio há coleta de lixo	0,018	0,32	0,018	0,007	0,45	0,009	0,004	0,59	0,007
No domicílio há esgotamento sanitário	-0,013	0,35	0,014	-0,002	0,79	0,008	0,001	0,90	0,008
No domicílio há eletricidade	0,083	0,09	0,049	0,043	0,15	0,030	0,007	0,70	0,018
R2		0,070			0,051			0,019	
Nº de observações		4.073			8.282			10.723	
Nº de estratos		110			112			114	
Nº de unidades primárias de amostra		758			822			826	

**Fonte dos dados básicos: PNAD - IBGE.**

**TABELA A 16 – Exercício II – Matrícula - Modelo logit com desenho complexo – MG e BA**

Variável	2005			2006			2007		
	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão
Ano	0,219	0,00	0,049	0,128	0,00	0,024	0,358	0,00	0,056
Tratamento	5,908	0,00	2,925	1,470	0,17	0,412	1,369	0,09	0,251
Interação	0,331	0,02	0,155	0,878	0,65	0,252	0,921	0,71	0,203
Tipo de família	0,824	0,09	0,095	0,887	0,21	0,084	0,926	0,38	0,081
Raça	0,890	0,40	0,122	0,995	0,97	0,142	0,903	0,36	0,100
Sexo	1,210	0,06	0,123	1,156	0,08	0,097	1,226	0,01	0,101
Condição de ocupação do chefe do domicílio	1,084	0,29	0,082	0,907	0,17	0,065	0,953	0,44	0,059
Escolaridade do chefe do domicílio	1,031	0,06	0,017	1,017	0,21	0,013	1,012	0,32	0,012
Raça do chefe do domicílio	1,421	0,02	0,215	1,148	0,27	0,145	1,235	0,06	0,138
Sexo do chefe do domicílio	1,190	0,28	0,192	0,935	0,60	0,122	0,920	0,47	0,107
Renda domiciliar per capita	0,981	0,64	0,041	1,084	0,14	0,059	1,106	0,05	0,056
Domicílio fica em área metropolitana	1,088	0,51	0,139	1,168	0,19	0,138	0,894	0,34	0,105
Domicílio fica na zona urbana	0,819	0,27	0,147	0,828	0,22	0,127	1,022	0,91	0,191
No domicílio há água encanada	1,145	0,42	0,194	1,081	0,60	0,159	1,087	0,64	0,195
No domicílio há coleta de lixo	1,193	0,27	0,189	1,116	0,43	0,156	1,067	0,57	0,123
No domicílio há esgotamento sanitário	0,890	0,39	0,120	0,953	0,72	0,130	1,023	0,87	0,149
No domicílio há eletricidade	1,732	0,04	0,469	1,665	0,11	0,528	1,098	0,70	0,269
Nº de observações	4.073			8.282			10.723		
Nº de estratos	110			112			114		
Nº de unidades primárias de amostra	758			822			826		

**Fonte dos dados básicos: PNAD - IBGE.**

**TABELA A 17 – Exercício III – PPS - Modelo de probabilidade linear – MG e SP**

Variável	e <sub>0</sub>			e <sub>1</sub>			e <sub>2</sub>		
	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão
Ano	-0,024	0,00	0,006	0,002	0,41	0,002	0,005	0,23	0,004
Tratamento	0,008	0,28	0,008	-0,002	0,38	0,003	-0,018	0,00	0,005
Interação	-0,016	0,14	0,011	0,003	0,46	0,004	0,009	0,20	0,007
Raça	-0,007	0,28	0,006	0,002	0,46	0,002	0,005	0,25	0,004
Sexo	-0,005	0,34	0,005	-0,002	0,31	0,002	0,007	0,03	0,003
Renda domiciliar per capita	-0,001	0,60	0,002	0,001	0,34	0,001	0,001	0,31	0,001
Casal com mais de um filho	-0,017	0,43	0,022	0,003	0,66	0,008	-0,011	0,41	0,014
Mãe com filhos	-0,011	0,64	0,023	0,003	0,69	0,009	-0,011	0,47	0,015
Outros tipos de família	-0,022	0,38	0,025	-0,001	0,88	0,009	-0,019	0,24	0,016
Condição de ocupação do chefe do domicílio	0,002	0,54	0,003	0,001	0,65	0,001	0,000	0,85	0,002
Chefe do domicílio possui ensino fundamental completo ou incompleto	0,024	0,00	0,007	-0,002	0,54	0,003	0,018	0,00	0,005
Chefe do domicílio possui ensino médio completo ou incompleto	0,018	0,03	0,008	0,003	0,37	0,003	0,018	0,00	0,005
Chefe do domicílio possui ensino superior incompleto	0,039	0,09	0,023	0,004	0,65	0,008	0,023	0,11	0,014
Chefe do domicílio possui ensino superior completo ou pós-graduação	0,022	0,14	0,015	0,002	0,77	0,005	0,021	0,03	0,009
Raça do chefe do domicílio	-0,004	0,56	0,006	-0,004	0,08	0,002	0,002	0,66	0,004
Sexo do chefe do domicílio	-0,009	0,24	0,008	-0,003	0,38	0,003	-0,003	0,51	0,005
Domicílio fica na zona urbana	0,005	0,72	0,013	0,003	0,55	0,005	-0,019	0,02	0,008
Domicílio fica em área metropolitana	0,008	0,15	0,005	0,000	0,95	0,002	0,002	0,50	0,003
No domicílio há eletricidade	-0,029	0,44	0,037	0,042	0,00	0,014	0,041	0,08	0,024
No domicílio há coleta de lixo	0,011	0,31	0,011	-0,006	0,16	0,004	0,003	0,64	0,007
No domicílio há esgotamento sanitário	0,007	0,41	0,008	0,002	0,47	0,003	0,006	0,24	0,005
No domicílio há água encanada	-0,012	0,39	0,014	0,000	0,99	0,005	0,016	0,08	0,009
R <sup>2</sup>	0,013			0,006			0,017		
R <sup>2</sup> ajustado	0,009			0,001			0,012		
Nº de observações	5.011			4.853			4.831		

**Fonte dos dados básicos: PNAD - IBGE.**

TABELA A 17 – Continuação

Variável	e <sub>3</sub>			e <sub>4</sub>		
	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão
Ano	0,017	0,01	0,007	0,042	0,00	0,012
Tratamento	-0,034	0,00	0,008	-0,080	0,00	0,015
Interação	0,038	0,00	0,012	0,046	0,03	0,021
Raça	0,011	0,10	0,007	0,019	0,12	0,012
Sexo	0,016	0,00	0,005	0,035	0,00	0,010
Renda domiciliar per capita	0,006	0,00	0,002	0,015	0,00	0,004
Casal com mais de um filho	-0,017	0,47	0,023	-0,006	0,88	0,041
Mãe com filhos	-0,010	0,69	0,025	-0,036	0,42	0,044
Outros tipos de família	-0,043	0,11	0,027	-0,030	0,53	0,048
Condição de ocupação do chefe do domicílio	0,000	0,90	0,004	-0,002	0,79	0,007
Chefe do domicílio possui ensino fundamental completo ou incompleto	0,011	0,14	0,008	0,001	0,94	0,014
Chefe do domicílio possui ensino médio completo ou incompleto	0,031	0,00	0,009	0,028	0,07	0,015
Chefe do domicílio possui ensino superior incompleto	0,023	0,35	0,024	0,021	0,62	0,043
Chefe do domicílio possui ensino superior completo ou pós-graduação	0,035	0,03	0,016	0,014	0,62	0,028
Raça do chefe do domicílio	0,004	0,53	0,007	0,024	0,05	0,012
Sexo do chefe do domicílio	-0,008	0,35	0,008	0,013	0,37	0,015
Domicílio fica na zona urbana	0,052	0,00	0,014	-0,045	0,08	0,026
Domicílio fica em área metropolitana	0,009	0,11	0,006	0,014	0,17	0,010
No domicílio há eletricidade	0,059	0,16	0,042	0,001	0,99	0,078
No domicílio há coleta de lixo	-0,007	0,54	0,012	-0,002	0,93	0,022
No domicílio há esgotamento sanitário	0,004	0,65	0,009	0,015	0,35	0,016
No domicílio há água encanada	-0,006	0,70	0,015	0,056	0,04	0,028
R2		0,033			0,038	
R2 ajustado		0,029			0,034	
Nº de observações		4.759			4.558	

Fonte dos dados básicos: PNAD - IBGE.



TABELA A 18 – Exercício III – PPS - Modelo logit – MG e SP

Variável	e <sub>0</sub>			e <sub>1</sub>			e <sub>2</sub>		
	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão
Ano	0,467	0,00	0,106	1,764	0,42	1,252	1,892	0,21	0,953
Tratamento	1,611	0,13	0,501	0,688	0,48	0,367	0,344	0,00	0,117
Interação	0,531	0,08	0,190	1,581	0,67	1,711	0,990	0,99	0,582
Raça	0,789	0,26	0,165	1,377	0,40	0,529	1,429	0,30	0,494
Sexo	0,867	0,41	0,149	0,644	0,33	0,291	1,866	0,03	0,519
Renda domiciliar per capita	0,978	0,66	0,048	1,432	0,11	0,323	1,091	0,61	0,186
Casal com mais de um filho	0,481	0,48	0,501	1,975	0,55	2,224	0,000	0,00	0,000
Mãe com filhos	0,577	0,61	0,619	2,014	0,62	2,840	0,000	0,00	0,000
Outros tipos de família	0,421	0,44	0,467	0,921	0,95	1,318	0,000	0,00	0,000
Condição de ocupação do chefe do domicílio	1,070	0,55	0,122	1,161	0,62	0,354	1,056	0,74	0,174
Chefe do domicílio possui ensino fundamental completo ou incompleto	1,949	0,00	0,419	0,766	0,69	0,504	2,820	0,00	0,830
Chefe do domicílio possui ensino médio completo ou incompleto	1,520	0,08	0,360	3,379	0,31	4,083	3,363	0,01	1,520
Chefe do domicílio possui ensino superior incompleto	3,901	0,18	3,944	Variável retirada do modelo			Variável retirada do modelo		
Chefe do domicílio possui ensino superior completo ou pós-graduação	1,787	0,30	1,003	1,593	0,70	1,941	4,665	0,13	4,790
Raça do chefe do domicílio	0,887	0,57	0,186	0,375	0,03	0,167	1,131	0,72	0,384
Sexo do chefe do domicílio	0,772	0,28	0,186	0,545	0,37	0,371	0,818	0,57	0,289
Domicílio fica na zona urbana	1,209	0,62	0,465	1,501	0,56	1,047	0,244	0,01	0,122
Domicílio fica em área metropolitana	1,306	0,16	0,246	0,969	0,95	0,462	1,153	0,62	0,332
No domicílio há eletricidade	Variável retirada do modelo			1,202	0,02	1,279	2,045	0,42	1,814
No domicílio há coleta de lixo	1,420	0,37	0,558	0,274	0,09	0,209	1,194	0,61	0,415
No domicílio há esgotamento sanitário	1,251	0,44	0,359	1,511	0,47	0,853	1,499	0,20	0,476
No domicílio há água encanada	0,648	0,25	0,245	1,217	0,77	0,824	2,508	0,06	1,242
Pseudo R2	0,044			0,079			0,104		
Log pseudolikelihood	-703,18			-123,33			-295,91		
Nº de observações	4.983			4.789			4.767		

Fonte dos dados básicos: PNAD - IBGE.

TABELA A 18 – Continuação

Variável	e <sub>3</sub>			e <sub>4</sub>		
	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão
Ano	1,968	0,01	0,526	1,596	0,00	0,224
Tratamento	0,614	0,01	0,113	0,584	0,00	0,067
Interação	1,583	0,20	0,571	1,160	0,43	0,218
Raça	1,322	0,14	0,250	1,163	0,18	0,132
Sexo	1,588	0,00	0,254	1,388	0,00	0,128
Renda domiciliar per capita	1,667	0,00	0,272	1,354	0,00	0,118
Casal com mais de um filho	0,494	0,30	0,337	0,974	0,95	0,423
Mãe com filhos	0,602	0,48	0,437	0,726	0,49	0,339
Outros tipos de família	0,296	0,11	0,224	0,769	0,59	0,370
Condição de ocupação do chefe do domicílio	1,011	0,92	0,104	0,996	0,95	0,062
Chefe do domicílio possui ensino fundamental completo ou incompleto	1,352	0,12	0,262	1,045	0,74	0,136
Chefe do domicílio possui ensino médio completo ou incompleto	3,201	0,00	1,067	1,331	0,07	0,210
Chefe do domicílio possui ensino superior incompleto	1,902	0,54	1,981	1,190	0,70	0,544
Chefe do domicílio possui ensino superior completo ou pós-graduação	3,080	0,04	1,712	1,164	0,59	0,330
Raça do chefe do domicílio	1,096	0,62	0,205	1,225	0,07	0,138
Sexo do chefe do domicílio	0,784	0,33	0,194	1,165	0,34	0,187
Domicílio fica na zona urbana	2,320	0,02	0,831	0,661	0,06	0,146
Domicílio fica em área metropolitana	1,351	0,10	0,245	1,150	0,17	0,117
No domicílio há eletricidade	1,911	0,27	1,131	0,990	0,99	0,563
No domicílio há coleta de lixo	0,803	0,52	0,270	0,961	0,84	0,190
No domicílio há esgotamento sanitário	1,032	0,89	0,234	1,116	0,42	0,153
No domicílio há água encanada	0,872	0,72	0,335	1,539	0,05	0,331
Pseudo R2	0,100			0,053		
Log pseudolikelihood	-689,39			-1.651,97		
Nº de observações	4.759			4.558		

Fonte dos dados básicos: PNAD - IBGE.

**TABELA A 19 – Exercício III – PPS - Modelo de probabilidade linear com desenho complexo– MG e SP**

Variável	e <sub>0</sub>			e <sub>1</sub>			e <sub>2</sub>		
	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão
Ano	0,100	0,00	0,026	0,002	0,33	0,002	0,004	0,18	0,003
Tratamento	0,068	0,00	0,023	-0,003	0,44	0,003	-0,018	0,00	0,005
Interação	-0,015	0,63	0,032	0,003	0,50	0,004	0,009	0,28	0,008
Raça	0,029	0,13	0,019	0,002	0,35	0,002	0,005	0,22	0,004
Sexo	0,068	0,00	0,013	-0,002	0,29	0,002	0,007	0,03	0,003
Renda domiciliar per capita	0,029	0,00	0,006	0,001	0,01	0,000	0,000	0,77	0,001
Tipo de Família	-0,009	0,56	0,015	-0,001	0,63	0,003	-0,003	0,39	0,004
Condição de ocupação do chefe do domicílio	0,009	0,36	0,009	0,000	0,76	0,001	0,001	0,81	0,002
Escolaridade do chefe do domicílio	0,022	0,02	0,009	0,001	0,14	0,001	0,006	0,00	0,002
Raça do chefe do domicílio	0,043	0,02	0,019	-0,004	0,05	0,002	0,002	0,68	0,004
Sexo do chefe do domicílio	-0,023	0,29	0,022	-0,002	0,59	0,003	-0,002	0,69	0,005
Domicílio fica na zona urbana	-0,034	0,30	0,032	0,003	0,52	0,005	-0,019	0,05	0,010
Domicílio fica em área metropolitana	-0,011	0,47	0,015	0,000	0,99	0,002	0,002	0,47	0,003
No domicílio há eletricidade	0,076	0,26	0,068	0,042	0,31	0,042	0,041	0,47	0,058
No domicílio há coleta de lixo	0,040	0,07	0,022	-0,006	0,09	0,003	0,003	0,67	0,008
No domicílio há esgotamento sanitário	0,038	0,06	0,021	0,002	0,47	0,003	0,006	0,27	0,005
No domicílio há água encanada	0,047	0,14	0,032	0,000	1,00	0,004	0,016	0,21	0,013
R2	0,064			0,005			0,015		
Nº de observações	3.858			4.816			4.794		
Nº de estratos	158			210			210		
Nº de unidades primárias de amostra	1.286			1.639			1.636		

Fonte dos dados básicos: PNAD - IBGE.

TABELA A 19 – Continuação

Variável	e <sub>3</sub>			e <sub>4</sub>		
	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão
Ano	0,018	0,01	0,007	0,041	0,00	0,014
Tratamento	-0,034	0,00	0,010	-0,077	0,00	0,016
Interação	0,037	0,00	0,012	0,045	0,05	0,023
Raça	0,011	0,10	0,007	0,021	0,10	0,013
Sexo	0,018	0,00	0,005	0,034	0,00	0,010
Renda domiciliar per capita	0,006	0,00	0,001	0,015	0,00	0,003
Tipo de Família	-0,008	0,26	0,007	-0,018	0,12	0,012
Condição de ocupação do chefe do domicílio	-0,001	0,76	0,004	-0,001	0,94	0,007
Escolaridade do chefe do domicílio	0,012	0,00	0,003	0,010	0,08	0,006
Raça do chefe do domicílio	0,005	0,50	0,007	0,025	0,04	0,013
Sexo do chefe do domicílio	0,000	0,97	0,008	0,007	0,59	0,014
Domicílio fica na zona urbana	0,055	0,01	0,022	-0,041	0,13	0,027
Domicílio fica em área metropolitana	0,009	0,10	0,006	0,015	0,18	0,011
No domicílio há eletricidade	0,059	0,29	0,056	0,000	1,00	0,075
No domicílio há coleta de lixo	-0,007	0,63	0,015	-0,002	0,95	0,027
No domicílio há esgotamento sanitário	0,005	0,64	0,010	0,019	0,27	0,017
No domicílio há água encanada	-0,007	0,75	0,021	0,052	0,09	0,031
R2	0,033			0,037		
Nº de observações	4.720			4.514		
Nº de estratos	209			208		
Nº de unidades primárias de amostra	1.625			1.611		

Fonte dos dados básicos: PNAD - IBGE.

**TABELA A 20 – Exercício III – PPS - Modelo logit com desenho complexo– MG e SP**

Variável	$\epsilon_0$			$\epsilon_1$			$\epsilon_2$		
	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão
Ano	0,433	0,00	0,096	1,845	0,37	1,257	1,801	0,24	0,902
Tratamento	1,582	0,16	0,515	0,683	0,47	0,361	0,356	0,00	0,123
Interação	0,569	0,13	0,210	1,596	0,67	1,740	1,015	0,98	0,623
Raça	0,766	0,19	0,155	1,376	0,41	0,535	1,410	0,30	0,472
Sexo	0,898	0,55	0,159	0,659	0,30	0,263	1,884	0,03	0,562
Renda domiciliar per capita	0,951	0,24	0,041	1,586	0,05	0,379	1,056	0,73	0,170
Tipo de Família	0,966	0,82	0,152	0,783	0,66	0,432	0,818	0,37	0,184
Condição de ocupação do chefe do domicílio	1,106	0,35	0,119	1,113	0,69	0,300	1,057	0,75	0,183
Escolaridade do chefe do domicílio	1,199	0,13	0,144	1,423	0,20	0,393	1,916	0,01	0,482
Raça do chefe do domicílio	0,883	0,55	0,186	0,391	0,03	0,170	1,144	0,70	0,394
Sexo do chefe do domicílio	0,817	0,38	0,188	0,702	0,58	0,446	0,888	0,73	0,300
Domicílio fica na zona urbana	1,250	0,59	0,510	1,478	0,59	1,077	0,258	0,01	0,126
Domicílio fica em área metropolitana	1,295	0,17	0,246	1,012	0,98	0,446	1,160	0,61	0,336
No domicílio há eletricidade	Variável retirada do modelo			12,576	0,01	12,936	1,922	0,47	1,719
No domicílio há coleta de lixo	1,268	0,59	0,550	0,269	0,09	0,210	1,203	0,62	0,451
No domicílio há esgotamento sanitário	1,251	0,46	0,375	1,592	0,42	0,922	1,480	0,24	0,497
No domicílio há água encanada	0,668	0,33	0,276	1,205	0,79	0,857	2,406	0,08	1,186
Nº de observações	4.944			4.816			4.794		
Nº de estratos	210			210			210		
Nº de unidades primárias de amostra	1.260			1.639			1.636		

Fonte dos dados básicos: PNAD - IBGE.

TABELA A 20 – Continuação

Variável	e <sub>3</sub>			e <sub>4</sub>		
	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão
Ano	2,106	0,01	0,600	1,576	0,00	0,248
Tratamento	0,612	0,02	0,130	0,595	0,00	0,072
Interação	1,500	0,27	0,546	1,174	0,45	0,247
Raça	1,297	0,20	0,260	1,186	0,16	0,144
Sexo	1,651	0,00	0,266	1,379	0,00	0,125
Renda domiciliar per capita	1,716	0,00	0,296	1,362	0,00	0,116
Tipo de Família	0,827	0,25	0,137	0,844	0,09	0,084
Condição de ocupação do chefe do domicílio	0,997	0,98	0,106	1,009	0,89	0,062
Escolaridade do chefe do domicílio	1,516	0,00	0,187	1,094	0,13	0,064
Raça do chefe do domicílio	1,094	0,64	0,210	1,236	0,07	0,143
Sexo do chefe do domicílio	0,977	0,92	0,221	1,077	0,56	0,137
Domicílio fica na zona urbana	2,390	0,02	0,891	0,684	0,10	0,156
Domicílio fica em área metropolitana	1,351	0,12	0,260	1,156	0,19	0,126
No domicílio há eletricidade	1,830	0,25	0,953	0,984	0,97	0,449
No domicílio há coleta de lixo	0,803	0,51	0,267	0,961	0,85	0,199
No domicílio há esgotamento sanitário	1,043	0,87	0,259	1,146	0,31	0,153
No domicílio há água encanada	0,867	0,72	0,346	1,488	0,07	0,320
Nº de observações	4.720			4.514		
Nº de estratos	209			208		
Nº de unidades primárias de amostra	1.625			1.611		

Fonte dos dados básicos: PNAD - IBGE.

**TABELA A 21 – Exercício III – PPS - Modelo de probabilidade linear – MG e BA**

Variável	e <sub>0</sub>			e <sub>1</sub>			e <sub>2</sub>		
	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão
Ano	-0,018	0,03	0,008	0,021	0,00	0,006	0,043	0,00	0,010
Tratamento	0,014	0,07	0,008	0,036	0,00	0,006	0,047	0,00	0,009
Interação	-0,021	0,05	0,011	-0,017	0,04	0,008	-0,029	0,02	0,013
Raça	0,003	0,63	0,007	0,004	0,52	0,006	0,006	0,50	0,008
Sexo	-0,003	0,57	0,005	0,019	0,00	0,004	0,029	0,00	0,006
Renda domiciliar per capita	0,000	0,93	0,003	0,003	0,18	0,002	0,007	0,04	0,003
Casal com mais de um filho	-0,016	0,42	0,020	0,018	0,25	0,015	0,020	0,40	0,024
Mãe com filhos	-0,017	0,44	0,022	0,017	0,31	0,017	0,025	0,34	0,026
Outros tipos de família	-0,012	0,59	0,023	0,021	0,25	0,018	0,007	0,79	0,028
Condição de ocupação do chefe do domicílio	0,000	0,94	0,004	-0,013	0,00	0,003	-0,006	0,15	0,004
Chefe do domicílio possui ensino fundamental completo ou incompleto	0,012	0,07	0,007	0,006	0,25	0,005	0,047	0,00	0,008
Chefe do domicílio possui ensino médio completo ou incompleto	0,000	0,98	0,009	0,014	0,04	0,007	0,056	0,00	0,010
Chefe do domicílio possui ensino superior incompleto	-0,003	0,92	0,026	0,021	0,31	0,020	0,081	0,01	0,031
Chefe do domicílio possui ensino superior completo ou pós-graduação	-0,020	0,36	0,022	0,007	0,68	0,017	0,042	0,11	0,026
Raça do chefe do domicílio	-0,017	0,02	0,007	-0,005	0,39	0,006	0,012	0,16	0,008
Sexo do chefe do domicílio	0,005	0,53	0,008	0,004	0,59	0,007	-0,012	0,24	0,010
Domicílio fica na zona urbana	-0,006	0,57	0,010	0,013	0,11	0,008	-0,003	0,82	0,012
Domicílio fica em área metropolitana	0,006	0,36	0,007	0,002	0,72	0,005	0,000	0,96	0,008
No domicílio há eletricidade	-0,002	0,89	0,017	0,023	0,08	0,013	0,103	0,00	0,021
No domicílio há coleta de lixo	-0,007	0,39	0,008	0,000	0,95	0,006	-0,002	0,83	0,010
No domicílio há esgotamento sanitário	0,011	0,15	0,007	-0,003	0,58	0,006	0,017	0,05	0,009
No domicílio há água encanada	0,000	0,97	0,010	0,008	0,32	0,008	0,016	0,17	0,012
R2	0,012			0,036			0,058		
R2 ajustado	0,007			0,031			0,053		
Nº de observações	4.764			4.612			4.511		

**Fonte dos dados básicos: PNAD - IBGE.**

TABELA A 21 – Continuação

Variável	e <sub>3</sub>			e <sub>4</sub>		
	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão
Ano	0,042	0,00	0,014	0,098	0,00	0,021
Tratamento	0,054	0,00	0,013	0,069	0,00	0,020
Interação	0,011	0,54	0,018	-0,013	0,64	0,027
Raça	0,010	0,40	0,012	0,025	0,15	0,017
Sexo	0,057	0,00	0,009	0,070	0,00	0,013
Renda domiciliar per capita	0,015	0,00	0,004	0,030	0,00	0,006
Casal com mais de um filho	0,051	0,13	0,033	-0,001	0,99	0,051
Mãe com filhos	0,079	0,03	0,036	-0,033	0,56	0,056
Outros tipos de família	0,033	0,39	0,038	-0,004	0,95	0,059
Condição de ocupação do chefe do domicílio	0,009	0,14	0,006	0,008	0,39	0,009
Chefe do domicílio possui ensino fundamental completo ou incompleto	0,032	0,00	0,011	0,032	0,06	0,017
Chefe do domicílio possui ensino médio completo ou incompleto	0,087	0,00	0,014	0,073	0,00	0,021
Chefe do domicílio possui ensino superior incompleto	0,106	0,01	0,041	0,043	0,48	0,060
Chefe do domicílio possui ensino superior completo ou pós-graduação	0,015	0,67	0,036	0,006	0,91	0,054
Raça do chefe do domicílio	0,019	0,09	0,012	0,041	0,02	0,017
Sexo do chefe do domicílio	-0,042	0,00	0,014	-0,011	0,60	0,021
Domicílio fica na zona urbana	0,003	0,84	0,017	-0,034	0,19	0,026
Domicílio fica em área metropolitana	0,024	0,03	0,011	-0,012	0,48	0,016
No domicílio há eletricidade	0,096	0,00	0,031	0,075	0,14	0,051
No domicílio há coleta de lixo	0,004	0,76	0,013	0,038	0,06	0,020
No domicílio há esgotamento sanitário	0,013	0,27	0,012	0,033	0,08	0,018
No domicílio há água encanada	0,013	0,42	0,017	0,049	0,06	0,026
R2		0,070			0,065	
R2 ajustado		0,066			0,060	
Nº de observações		4.274			3.871	

Fonte dos dados básicos: PNAD - IBGE.



TABELA A 22 – Exercício III – PPS - Modelo logit – MG e BA

Variável	$\epsilon_0$			$\epsilon_1$			$\epsilon_2$		
	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão
Ano	0,601	0,03	0,144	1,734	0,04	0,471	1,825	0,00	0,351
Tratamento	2,055	0,02	0,622	4,860	0,00	1,808	2,111	0,00	0,434
Interação	0,421	0,02	0,153	1,753	0,50	1,452	1,015	0,97	0,368
Raça	1,096	0,68	0,241	1,364	0,32	0,427	1,275	0,30	0,300
Sexo	0,917	0,62	0,160	2,952	0,00	0,730	1,973	0,00	0,300
Renda domiciliar per capita	1,009	0,90	0,067	2,042	0,18	1,086	1,483	0,29	0,555
Casal com mais de um filho	0,575	0,46	0,427	1,824	0,33	1,118	1,560	0,30	0,671
Mãe com filhos	0,563	0,47	0,443	1,741	0,44	1,243	1,774	0,24	0,869
Outros tipos de família	0,640	0,60	0,548	1,957	0,36	1,431	1,238	0,68	0,633
Condição de ocupação do chefe do domicílio	0,993	0,95	0,113	0,569	0,00	0,076	0,887	0,21	0,084
Chefe do domicílio possui ensino fundamental completo ou incompleto	1,443	0,08	0,307	1,068	0,78	0,255	2,134	0,00	0,349
Chefe do domicílio possui ensino médio completo ou incompleto	1,016	0,95	0,275	3,789	0,02	2,179	3,888	0,00	1,232
Chefe do domicílio possui ensino superior incompleto	0,970	0,97	0,729	Variável retirada do modelo			Variável retirada do modelo		
Chefe do domicílio possui ensino superior completo ou pós-graduação	0,645	0,44	0,367	1,198	0,864	1,266	1,942	0,39	1,510
Raça do chefe do domicílio	0,609	0,02	0,128	0,747	0,30	0,212	1,383	0,14	0,300
Sexo do chefe do domicílio	1,161	0,58	0,313	1,159	0,66	0,391	0,718	0,17	0,171
Domicílio fica na zona urbana	0,886	0,69	0,268	1,572	0,24	0,611	0,895	0,68	0,245
Domicílio fica em área metropolitana	1,221	0,33	0,253	1,048	0,89	0,342	0,914	0,65	0,182
No domicílio há eletricidade	0,921	0,88	0,520	1,350	0,46	0,543	2,105	0,01	0,580
No domicílio há coleta de lixo	0,783	0,38	0,217	1,069	0,84	0,340	1,010	0,96	0,216
No domicílio há esgotamento sanitário	1,355	0,22	0,339	0,908	0,78	0,307	1,547	0,06	0,351
No domicílio há água encanada	1,007	0,98	0,260	1,269	0,55	0,500	1,257	0,38	0,329
Pseudo R2	0,042			0,179			0,133		
Log pseudolikelihood	-684,36			-379,70			-772,63		
Nº de observações	4.764			4.560			4.459		

Fonte dos dados básicos: PNAD - IBGE.

TABELA A 22 – Continuação

Variável	e <sub>3</sub>			e <sub>4</sub>		
	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão
Ano	1,451	0,01	0,216	1,664	0,00	0,199
Tratamento	1,555	0,00	0,234	1,350	0,01	0,145
Interação	2,099	0,01	0,590	1,115	0,53	0,194
Raça	1,216	0,24	0,201	1,185	0,14	0,137
Sexo	2,112	0,00	0,253	1,541	0,00	0,129
Renda domiciliar per capita	1,964	0,00	0,352	1,459	0,00	0,136
Casal com mais de um filho	2,030	0,04	0,696	1,056	0,86	0,335
Mãe com filhos	2,954	0,01	1,164	0,865	0,68	0,303
Outros tipos de família	1,633	0,23	0,662	1,000	1,00	0,379
Condição de ocupação do chefe do domicílio	1,127	0,13	0,089	1,054	0,36	0,061
Chefe do domicílio possui ensino fundamental completo ou incompleto	1,361	0,02	0,173	1,218	0,06	0,126
Chefe do domicílio possui ensino médio completo ou incompleto	5,068	0,00	1,424	1,599	0,00	0,226
Chefe do domicílio possui ensino superior incompleto	8,044	0,05	8,676	1,259	0,55	0,480
Chefe do domicílio possui ensino superior completo ou pós-graduação	1,071	0,89	0,517	1,055	0,87	0,336
Raça do chefe do domicílio	1,290	0,12	0,210	1,283	0,03	0,146
Sexo do chefe do domicílio	0,541	0,00	0,097	0,933	0,65	0,140
Domicílio fica na zona urbana	0,948	0,80	0,203	0,807	0,19	0,132
Domicílio fica em área metropolitana	1,408	0,02	0,211	0,912	0,36	0,092
No domicílio há eletricidade	1,640	0,07	0,450	1,330	0,31	0,370
No domicílio há coleta de lixo	1,095	0,58	0,181	1,245	0,07	0,151
No domicílio há esgotamento sanitário	1,115	0,49	0,177	1,186	0,12	0,132
No domicílio há água encanada	1,079	0,69	0,208	1,277	0,12	0,198
Pseudo R2	0,133			0,067		
Log pseudolikelihood	-1.158,89			-1.921,18		
Nº de observações	4.274			3.871		

Fonte dos dados básicos: PNAD - IBGE.

**TABELA A 23 – Exercício III – PPS - Modelo de probabilidade linear com desenho complexo– MG e BA**

Variável	e <sub>0</sub>			e <sub>1</sub>			e <sub>2</sub>		
	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão
Ano	-0,019	0,05	0,010	0,022	0,02	0,010	0,042	0,00	0,013
Tratamento	0,016	0,01	0,006	0,036	0,00	0,007	0,050	0,00	0,011
Interação	-0,021	0,10	0,013	-0,017	0,10	0,010	-0,030	0,04	0,015
Raça	0,003	0,68	0,007	0,004	0,48	0,005	0,006	0,41	0,008
Sexo	-0,003	0,64	0,006	0,019	0,00	0,004	0,029	0,00	0,007
Renda domiciliar per capita	-0,001	0,76	0,003	0,003	0,02	0,001	0,005	0,14	0,003
Tipo de Família	-0,001	0,93	0,006	0,002	0,58	0,004	-0,002	0,81	0,008
Condição de ocupação do chefe do domicílio	0,000	0,93	0,004	-0,014	0,00	0,004	-0,007	0,15	0,005
Escolaridade do chefe do domicílio	-0,002	0,73	0,005	0,005	0,02	0,002	0,022	0,00	0,005
Raça do chefe do domicílio	-0,016	0,03	0,007	-0,005	0,41	0,006	0,012	0,11	0,007
Sexo do chefe do domicílio	0,003	0,65	0,007	0,002	0,67	0,005	-0,009	0,37	0,010
Domicílio fica na zona urbana	-0,005	0,62	0,011	0,013	0,21	0,010	-0,002	0,91	0,014
Domicílio fica em área metropolitana	0,007	0,31	0,007	0,002	0,56	0,004	0,000	0,97	0,006
No domicílio há eletricidade	0,000	1,00	0,014	0,023	0,35	0,025	0,108	0,00	0,035
No domicílio há coleta de lixo	-0,007	0,41	0,008	0,000	0,95	0,006	-0,002	0,85	0,011
No domicílio há esgotamento sanitário	0,011	0,28	0,010	-0,003	0,61	0,006	0,017	0,12	0,011
No domicílio há água encanada	0,000	1,00	0,009	0,008	0,48	0,011	0,017	0,27	0,015
R2	0,010			0,035			0,054		
Nº de observações	4.752			4.601			4.500		
Nº de estratos	159			159			159		
Nº de unidades primárias de amostra	1.351			1.322			1.318		

**Fonte dos dados básicos: PNAD - IBGE.**

TABELA A 23 – Continuação

Variável	e <sub>3</sub>			e <sub>4</sub>		
	Coefficiente	P-valor	Desvio padrão	Coefficiente	P-valor	Desvio padrão
Ano	0,043	0,03	0,020	0,100	0,00	0,026
Tratamento	0,052	0,00	0,016	0,068	0,00	0,023
Interação	0,010	0,64	0,022	-0,015	0,63	0,032
Raça	0,011	0,33	0,012	0,029	0,13	0,019
Sexo	0,058	0,00	0,010	0,068	0,00	0,013
Renda domiciliar per capita	0,016	0,00	0,003	0,029	0,00	0,006
Tipo de Família	0,005	0,57	0,009	-0,009	0,56	0,015
Condição de ocupação do chefe do domicílio	0,007	0,28	0,006	0,009	0,36	0,009
Escolaridade do chefe do domicílio	0,030	0,00	0,006	0,022	0,02	0,009
Raça do chefe do domicílio	0,020	0,08	0,011	0,043	0,02	0,019
Sexo do chefe do domicílio	-0,027	0,03	0,012	-0,023	0,29	0,022
Domicílio fica na zona urbana	0,004	0,83	0,021	-0,034	0,30	0,032
Domicílio fica em área metropolitana	0,025	0,01	0,009	-0,011	0,47	0,015
No domicílio há eletricidade	0,097	0,07	0,054	0,076	0,26	0,068
No domicílio há coleta de lixo	0,004	0,79	0,014	0,040	0,07	0,022
No domicílio há esgotamento sanitário	0,016	0,28	0,015	0,038	0,06	0,021
No domicílio há água encanada	0,015	0,44	0,020	0,047	0,14	0,032
R2	0,066			0,064		
Nº de observações	4.261			3.858		
Nº de estratos	158			158		
Nº de unidades primárias de amostra	1.307			1.286		

Fonte dos dados básicos: PNAD - IBGE.

TABELA A 24 – Exercício III – PPS - Modelo logit com desenho complexo– MG e BA

Variável	e <sub>0</sub>			e <sub>1</sub>			e <sub>2</sub>		
	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão
Ano	0,588	0,04	0,150	1,798	0,04	0,524	1,821	0,00	0,350
Tratamento	2,149	0,02	0,676	4,562	0,00	1,994	2,136	0,00	0,442
Interação	0,427	0,03	0,166	1,728	0,52	1,478	0,997	1,00	0,399
Raça	1,077	0,72	0,219	1,343	0,35	0,422	1,290	0,25	0,287
Sexo	0,923	0,64	0,159	2,953	0,00	0,719	1,967	0,00	0,313
Renda domiciliar per capita	0,982	0,77	0,061	2,063	0,13	0,990	1,454	0,30	0,529
Tipo de Família	0,982	0,92	0,171	1,114	0,59	0,221	0,979	0,89	0,156
Condição de ocupação do chefe do domicílio	0,995	0,96	0,114	0,563	0,00	0,080	0,870	0,16	0,086
Escolaridade do chefe do domicílio	0,954	0,74	0,137	1,366	0,06	0,230	1,905	0,00	0,278
Raça do chefe do domicílio	0,628	0,02	0,124	0,751	0,32	0,217	1,384	0,12	0,286
Sexo do chefe do domicílio	1,106	0,67	0,259	1,128	0,69	0,337	0,803	0,31	0,175
Domicílio fica na zona urbana	0,888	0,70	0,277	1,577	0,25	0,627	0,915	0,73	0,235
Domicílio fica em área metropolitana	1,242	0,32	0,270	1,122	0,74	0,383	0,917	0,67	0,184
No domicílio há eletricidade	0,992	0,99	0,490	1,312	0,54	0,576	2,155	0,00	0,530
No domicílio há coleta de lixo	0,800	0,42	0,220	1,070	0,83	0,332	1,001	1,00	0,228
No domicílio há esgotamento sanitário	1,356	0,27	0,371	0,902	0,77	0,317	1,535	0,09	0,389
No domicílio há água encanada	0,992	0,98	0,266	1,270	0,54	0,494	1,271	0,32	0,306
Nº de observações	4.752			4.601			4.500		
Nº de estratos	159			159			159		
Nº de unidades primárias de amostra	1.341			1.322			1.318		

Fonte dos dados básicos: PNAD - IBGE.

TABELA A 24 – Continuação

Variável	e <sub>3</sub>			e <sub>4</sub>		
	Razão de chance	P-valor	Desvio padrão	Razão de chance	P-valor	Desvio padrão
Ano	1,491	0,02	0,258	1,679	0,00	0,229
Tratamento	1,461	0,02	0,236	1,345	0,01	0,159
Interação	2,076	0,01	0,587	1,097	0,64	0,217
Raça	1,218	0,23	0,200	1,219	0,12	0,156
Sexo	2,094	0,00	0,265	1,524	0,00	0,127
Renda domiciliar per capita	2,037	0,00	0,345	1,448	0,00	0,133
Tipo de Família	1,072	0,56	0,128	0,940	0,51	0,089
Condição de ocupação do chefe do domicílio	1,098	0,24	0,088	1,057	0,34	0,061
Escolaridade do chefe do domicílio	1,550	0,00	0,158	1,152	0,02	0,071
Raça do chefe do domicílio	1,297	0,11	0,209	1,306	0,03	0,161
Sexo do chefe do domicílio	0,691	0,02	0,110	0,871	0,30	0,117
Domicílio fica na zona urbana	0,953	0,83	0,210	0,809	0,24	0,147
Domicílio fica em área metropolitana	1,453	0,01	0,222	0,918	0,40	0,094
No domicílio há eletricidade	1,632	0,12	0,517	1,331	0,33	0,386
No domicílio há coleta de lixo	1,084	0,63	0,181	1,254	0,06	0,149
No domicílio há esgotamento sanitário	1,139	0,46	0,199	1,230	0,08	0,146
No domicílio há água encanada	1,107	0,59	0,207	1,269	0,15	0,212
Nº de observações	4.261			3.858		
Nº de estratos	158			158		
Nº de unidades primárias de amostra	1.307			1.286		

Fonte dos dados básicos: PNAD - IBGE.

## APÊNDICE B – Desenho da amostra – Exercício I

Variáveis	Categorias	2001-2005			2001-2006			2001-2007		
		Bahia	Minas Gerais	São Paulo	Bahia	Minas Gerais	São Paulo	Bahia	Minas Gerais	São Paulo
Tipo de família	Casal com 1 filho	1,55	1,13	0,53	1,80	1,19	0,65	1,83	1,16	0,91
	Casal com mais de um filho	73,80	76,27	80,90	74,32	76,33	79,92	74,34	76,54	78,38
	Mãe com filhos	21,27	19,65	16,25	20,42	19,81	17,10	20,11	19,22	18,15
	Outros tipos de família	3,38	2,95	2,33	3,46	2,67	599,61	3,71	3,08	2,56
Raça	Branco ou amarelo	19,17	47,76	68,69	19,25	46,31	67,26	19,06	44,95	66,72
	Negro, pardo ou indígena	80,83	52,24	31,31	80,74	53,69	32,74	80,92	55,05	33,28
Sexo	Mulher	49,93	49,98	48,13	49,57	49,75	49,31	49,52	49,26	48,83
	Homem	50,07	50,02	51,87	50,43	50,25	50,69	50,48	50,74	51,17
Condição de ocupação do chefe	Ocupado	79,01	79,27	78,17	77,05	78,72	75,99	76,11	76,37	74,95
	Desocupado	20,99	20,73	21,83	22,95	21,28	24,01	23,89	23,63	25,05
Escolaridade do chefe (em anos de estudo)	0	31,36	15,46	10,36	30,30	16,17	9,69	31,03	17,25	11,20
	1	5,61	3,55	1,84	4,37	3,41	2,43	4,43	3,28	2,04
	2	7,35	6,50	4,30	7,07	5,43	4,58	7,07	5,88	4,24
	3	7,99	7,48	6,50	7,24	8,89	6,57	7,66	7,80	5,95
	4	13,46	21,47	16,95	13,76	19,32	17,66	12,40	21,72	18,49
	5	4,84	5,91	6,41	5,33	5,45	6,42	5,06	5,49	6,81
	6	2,83	3,89	5,45	3,00	4,16	4,60	2,70	3,83	4,58
	7	2,69	4,73	4,57	3,41	4,68	4,16	2,87	4,08	4,72
	8	6,07	8,91	12,08	6,19	9,53	11,74	5,86	9,28	10,73
	9	1,87	1,58	1,93	1,78	1,65	2,09	1,81	1,58	1,94
	10	1,73	1,58	2,33	1,83	1,47	2,72	2,16	1,52	2,42
	11	11,32	13,54	16,34	12,80	13,54	18,04	14,02	13,28	17,54
	12	0,14	0,39	0,66	0,37	0,40	0,69	0,25	0,40	0,89
	13	0,41	0,39	1,45	0,21	0,57	0,81	0,30	0,48	0,92
	14	0,18	0,54	1,49	0,27	0,60	1,15	0,32	0,47	1,05
	15	1,87	3,84	6,32	1,75	4,33	5,91	1,74	3,19	4,82
	16	0,27	0,25	1,01	0,30	0,38	0,73	0,32	0,49	1,66
Sexo do chefe	Mulher	26,38	22,16	23,32	27,89	24,41	24,91	27,69	25,15	26,03
	Homem	73,62	77,84	76,68	72,11	75,59	75,09	72,31	74,85	73,97
Raça do chefe	Branco ou amarelo	18,71	45,15	64,51	19,16	43,98	65,15	17,50	43,67	66,18
	Negro, pardo ou indígena	81,29	54,85	35,49	80,83	56,02	34,85	82,50	56,33	33,82
Metropolitano	Área metropolitana	36,70	33,83	44,58	36,42	33,50	47,29	37,40	33,00	47,80
	Área não-metropolitana	63,30	66,17	55,42	63,58	66,50	52,71	62,60	67,00	52,20
Urbano	Área urbana	71,84	84,64	92,62	71,26	84,99	92,80	71,16	84,59	92,86
	Área rural	28,16	15,36	7,38	28,74	15,01	7,20	28,84	15,41	7,14
Água	Possui água encanada	65,68	82,72	94,07	67,95	83,35	94,63	68,90	83,37	94,73
	Não possui água encanada	34,32	17,28	5,93	32,05	16,65	5,37	31,10	16,63	4,14
Lixo	Possui coleta de lixo	46,51	74,64	90,56	48,19	76,43	90,44	49,87	76,76	90,95
	Não possui coleta de lixo	53,49	25,36	9,44	51,81	23,57	9,56	50,13	23,24	8,89
Esgoto	Possui esgotamento	33,41	68,34	80,63	37,64	68,97	81,09	39,08	69,67	81,46
	Não possui esgotamento	66,59	31,66	19,37	62,36	31,03	18,91	60,92	30,33	18,00
Eletricidade	Possui eletricidade	88,04	97,59	99,52	89,13	97,46	99,65	90,21	97,61	99,70
	Não possui eletricidade	11,91	2,41	0,48	10,87	2,54	0,35	9,79	2,39	0,14
Matrícula no EF	Matriculado	78,55	89,22	90,82	81,88	89,42	92,05	84,68	90,85	93,18
	Não-matriculado	21,45	10,78	9,18	18,12	10,58	7,95	15,32	9,15	6,82
Renda	Mínimo	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Máximo	14,55	27,33	22,83	22,00	23,43	71,96	20,26	27,89	66,28
	Média	0,47	0,74	1,17	0,49	0,74	1,06	0,48	0,73	1,00
	Desvio padrão	0,82	1,29	1,85	0,91	1,20	1,77	0,87	1,20	1,51

Fonte dos dados básicos: PNAD - IBGE.

## APÊNDICE C – Desenho da amostra – Exercício II

Variáveis	Categorias	2005			2006			2007		
		Bahia	Minas Gerais	São Paulo	Bahia	Minas Gerais	São Paulo	Bahia	Minas Gerais	São Paulo
Tipo de família	Casal com 1 filho	1,28	1,36	0,66	1,89	1,18	0,77	2,28	1,11	1,28
	Casal com mais de um filho	71,85	75,76	76,91	72,58	74,82	77,74	71,90	75,46	74,94
	Mãe com filhos	22,71	20,02	20,01	21,83	21,11	18,81	21,21	19,22	20,59
	Outros tipos de família	4,16	2,87	2,42	3,71	2,90	2,67	4,60	4,21	3,20
Raça	Branco ou amarelo	18,51	42,18	62,71	17,94	42,31	64,70	17,99	39,96	63,10
	Negro, pardo ou indígena	81,44	57,82	37,29	82,06	57,69	35,30	81,97	60,04	36,90
Sexo	Mulher	48,72	48,75	48,90	48,70	49,14	49,07	49,67	48,61	47,80
	Homem	51,28	51,25	51,10	51,30	50,86	50,93	50,33	51,39	52,20
Condição de ocupação do chefe	Ocupado	80,90	83,99	82,32	80,01	83,82	83,36	79,42	83,16	80,74
	Desocupado	19,10	16,01	17,68	19,99	16,18	16,64	20,58	16,84	19,26
Escolaridade do chefe (em anos de estudo)	0	26,46	11,31	8,00	25,70	11,85	5,75	28,33	16,08	16,94
	1	6,26	3,44	1,54	4,72	3,25	1,83	4,55	3,20	1,85
	2	7,59	4,95	3,39	6,75	4,78	3,25	6,78	5,11	3,18
	3	7,50	7,04	6,11	8,36	8,08	4,93	8,43	7,67	4,21
	4	13,07	21,58	15,66	12,71	19,00	15,58	11,63	18,66	13,81
	5	5,58	6,31	7,48	6,40	7,15	8,16	5,64	7,39	7,38
	6	3,34	5,53	5,15	3,34	4,63	5,40	2,84	4,61	4,31
	7	3,56	4,64	5,36	3,73	5,78	4,67	3,04	4,06	4,02
	8	6,17	11,16	12,27	5,89	10,28	13,00	6,87	11,96	12,25
	9	2,97	2,35	2,42	2,30	2,00	2,50	2,30	1,70	1,83
	10	2,33	1,30	2,33	2,33	2,30	2,37	2,71	1,51	2,61
	11	12,48	14,75	19,61	14,21	14,55	22,06	14,36	14,95	20,05
	12	0,23	0,52	0,53	0,39	0,65	0,67	0,47	0,40	0,64
	13	0,32	0,63	1,36	0,28	0,55	1,08	0,34	0,19	0,37
	14	0,23	0,63	1,32	0,46	0,55	1,38	0,67	0,19	0,30
	15	1,78	3,49	6,64	2,03	4,28	6,41	0,70	0,96	1,65
16	0,14	0,36	0,84	0,41	0,33	0,97	30,24	1,36	4,61	
Sexo do chefe	Mulher	28,75	22,26	23,61	29,75	25,83	24,84	69,76	26,47	28,09
	Homem	71,25	77,74	76,39	70,25	74,17	75,16	17,11	73,53	71,91
Raça do chefe	Branco ou amarelo	16,54	38,01	58,75	17,02	39,18	61,71	82,85	37,53	59,68
	Negro, pardo ou indígena	83,46	61,99	41,25	82,98	60,74	38,29	38,31	62,47	40,32
Metropolitano	Área metropolitana	38,89	34,57	46,88	36,52	32,66	47,50	2,28	32,84	47,84
	Área não-metropolitana	61,11	65,43	53,12	63,48	67,34	52,50	61,69	67,16	52,16
Urbano	Área urbana	72,21	84,88	92,88	70,71	85,00	93,63	71,42	83,96	93,44
	Área rural	27,79	15,12	7,12	28,67	15,00	6,37	28,58	16,04	6,56
Água	Possui água encanada	70,75	84,10	93,84	71,33	85,55	95,18	73,67	84,34	95,34
	Não possui água encanada	29,25	15,90	6,16	28,67	14,45	4,82	26,33	15,66	4,66
Lixo	Possui coleta de lixo	52,70	76,59	92,17	51,19	78,54	91,95	53,96	78,14	92,46
	Não possui coleta de lixo	47,30	23,41	7,83	48,81	21,46	8,05	46,04	21,86	7,54
Esgoto	Possui esgotamento	39,81	69,86	82,98	42,11	71,32	81,10	44,29	72,42	82,32
	Não possui esgotamento	60,19	30,14	17,02	57,89	28,68	18,90	55,71	27,58	17,68
Eletricidade	Possui eletricidade	92,05	98,12	99,52	92,40	98,22	99,59	94,86	98,60	99,61
	Não possui eletricidade	7,95	1,88	0,48	7,60	1,78	0,41	5,14	1,40	0,39
Matrícula no EF	Matriculado	82,45	91,14	92,52	91,90	94,40	96,97	92,43	94,64	96,59
	Não-matriculado	17,55	8,86	7,48	8,10	5,60	3,03	7,57	5,36	3,41
	Mínimo	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Renda	Máximo	14,56	27,33	30,00	34,29	25,33	71,96	26,32	25,33	19,87
	Média	0,66	0,92	1,34	0,67	0,97	1,33	0,64	0,92	1,24
	Desvio padrão	1,04	1,42	1,88	1,25	1,41	2,24	1,05	1,29	1,58

Fonte dos dados básicos: PNAD - IBGE.



## APÊNDICE D – Desenho da amostra – Exercício III: e0, e1, e2

Variáveis	Categorias	e <sub>0</sub>			e <sub>1</sub>			e <sub>2</sub>		
		Bahia	Minas Gerais	São Paulo	Bahia	Minas Gerais	São Paulo	Bahia	Minas Gerais	São Paulo
Tipo de família	Casal com 1 filho	2,21	1,35	1,36	2,20	1,39	1,37	2,15	1,36	1,37
	Casal com mais de um filho	69,63	73,86	75,60	69,70	73,61	75,75	69,79	73,69	75,81
	Mãe com filhos	22,95	20,14	19,67	22,83	20,34	19,55	22,80	20,32	19,51
	Outros tipos de família	5,21	4,65	3,37	5,27	4,66	3,32	5,25	4,64	3,30
Raça	Branco ou amarelo	16,60	39,27	60,73	16,57	39,16	60,46	16,56	60,68	60,38
	Negro, pardo ou indígena	83,40	60,73	39,27	83,43	60,84	39,54	83,44	39,32	39,62
Sexo	Mulher	48,94	48,69	46,31	49,07	48,39	46,30	50,13	48,42	46,21
	Homem	51,06	51,31	53,69	50,93	51,61	53,70	49,87	51,58	53,79
Condição de ocupação do chefe	Ocupado	77,35	81,93	79,76	77,47	81,79	79,82	78,14	81,87	79,78
	Desocupado	22,65	18,07	20,24	22,53	18,21	20,18	21,86	18,13	20,22
	Analfabeto	30,87	18,67	18,88	30,69	18,47	18,34	30,03	18,48	18,34
Escolaridade do chefe (em anos de estudo)	Ensino Fundamental	46,02	59,42	48,35	46,48	1369,00	48,89	46,45	59,54	48,80
	Ensino Médio	20,78	18,58	26,79	20,59	18,55	26,71	21,18	18,65	26,78
	Ensino Superior incompleto	1,13	1,14	1,44	1,12	1,13	1,49	1,17	1,14	1,49
	Ensino Superior completo	1,21	2,20	4,54	1,12	2,22	4,58	1,17	2,19	4,59
Sexo do chefe	Mulher	34,08	29,56	30,95	33,92	29,83	30,62	33,84	29,82	30,55
	Homem	65,92	70,44	69,05	66,08	70,17	69,38	66,16	70,18	69,45
Raça do chefe	Branco ou amarelo	16,56	39,57	59,86	16,53	39,16	59,76	16,34	39,23	59,64
	Negro, pardo ou indígena	83,40	60,43	40,14	83,43	60,84	40,24	83,62	60,77	40,36
Metropolitano	Área metropolitana	38,13	34,67	48,20	38,37	34,80	48,38	38,91	34,85	48,37
	Área não-metropolitana	61,87	65,33	51,80	61,63	65,20	51,62	61,09	65,15	51,63
Urbano	Área urbana	70,55	84,29	94,74	70,48	1936,00	94,80	71,23	84,41	94,78
	Área rural	29,45	15,71	5,26	29,52	15,68	5,20	28,77	15,59	5,22
Água	Possui água encanada	74,89	84,38	94,63	74,79	84,41	94,64	75,63	84,46	94,66
	Não possui água encanada	25,11	15,63	5,37	25,21	15,59	5,36	24,37	15,54	5,34
Lixo	Possui coleta de lixo	54,90	80,57	93,45	54,60	80,57	93,55	55,21	80,56	93,56
	Não possui coleta de lixo	45,10	19,43	6,55	45,40	19,43	6,45	44,79	19,44	6,44
Esgoto	Possui esgotamento	43,60	74,32	85,09	43,76	74,39	85,10	44,30	74,47	85,12
	Não possui esgotamento	56,40	25,68	14,91	56,24	25,61	14,90	55,70	25,53	14,88
Eletricidade	Possui eletricidade	95,62	99,24	99,62	95,64	99,22	99,61	95,87	99,21	99,65
	Não possui eletricidade	4,38	0,76	0,38	4,36	0,78	0,39	4,13	0,79	0,35
Matrícula no EF	Matriculado	97,29	97,72	97,92	98,40	1,39	99,10	98,74	98,56	99,14
	Não-matriculado	2,71	2,28	2,08	1,60	1,39	0,90	1,26	1,44	0,86
Renda	Mínimo	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	Máximo	11,01	25,69	21,41	11,01	25,69	21,41	11,01	25,69	21,41
	Média	0,63	0,96	1,29	0,63	0,96	1,28	0,64	0,97	1,28
	Desvio padrão	0,89	1,28	1,58	0,89	1,30	1,57	0,90	1,30	1,57

Fonte dos dados básicos: PNAD - IBGE.

## APÊNDICE E – Desenho da amostra – Exercício III: e3, e4

Variáveis	Categorias	e <sub>3</sub>			e <sub>4</sub>		
		Bahia	Minas Gerais	São Paulo	Bahia	Minas Gerais	São Paulo
Tipo de família	Casal com 1 filho	2,00	1,39	1,38	1,81	1,38	1,42
	Casal com mais de um filho	70,07	73,77	75,84	70,49	73,67	76,00
	Mãe com filhos	22,79	20,27	19,49	22,91	20,33	19,49
	Outros tipos de família	5,13	4,57	3,28	4,80	4,62	3,09
Raça	Branco ou amarelo	16,87	39,82	60,34	16,87	40,95	60,37
	Negro, pardo ou indígena	83,13	60,18	39,66	83,13	59,05	39,63
Sexo	Mulher	51,20	48,97	46,22	53,84	49,67	46,54
	Homem	48,80	51,03	53,78	46,16	50,33	53,46
Condição de ocupação do chefe	Ocupado	78,53	82,11	79,68	21,50	82,10	79,82
	Desocupado	21,47	17,89	20,32	78,50	17,90	20,18
Escolaridade do chefe (em anos de estudo)	Analfabeto	28,26	17,98	18,23	26,13	17,67	18,06
	Ensino Fundamental	46,65	59,78	48,79	46,05	59,00	48,54
	Ensino Médio	22,59	18,88	26,85	25,17	19,90	27,18
	Ensino Superior incompleto	1,27	1,17	1,50	1,41	1,24	1,51
Sexo do chefe	Ensino Superior completo	1,22	2,20	4,63	1,24	2,19	4,72
	Mulher	33,79	29,64	30,53	33,30	29,67	30,43
Raça do chefe	Homem	66,21	70,36	69,47	66,70	70,33	69,57
	Branco ou amarelo	16,82	39,60	59,63	17,16	40,43	59,72
Metropolitano	Negro, pardo ou indígena	83,13	60,40	40,37	82,79	59,57	40,28
	Área metropolitana	39,71	35,16	48,32	42,04	35,86	48,58
Urbano	Área não-metropolitana	60,29	64,84	51,68	57,96	64,14	51,42
	Área urbana	72,67	84,53	94,74	74,38	85,38	95,04
Água	Área rural	27,33	15,47	5,26	25,62	14,62	4,96
	Possui água encanada	77,21	84,66	94,74	79,18	85,33	94,91
Lixo	Não possui água encanada	22,79	15,34	5,26	20,82	14,67	5,09
	Possui coleta de lixo	56,09	80,90	93,55	57,34	81,62	93,69
Esgoto	Não possui coleta de lixo	43,91	19,10	6,45	42,66	18,38	6,31
	Possui esgotamento	45,77	74,93	85,13	48,25	75,76	85,35
Eletricidade	Não possui esgotamento	54,23	25,07	376,00	51,75	24,24	14,65
	Possui eletricidade	96,63	99,24	99,68	97,23	99,38	99,67
Matrícula no EF	Não possui eletricidade	3,37	0,76	0,32	2,77	0,62	0,33
	Matriculado	98,83	98,70	99,13	98,93	98,86	99,15
Renda	Não-matriculado	1,17	1,30	0,87	1,07	1,14	0,85
	Mínimo	0,00	0,00	0,00	0,00	0,00	0,00
	Máximo	11,01	25,69	21,41	11,01	25,69	21,41
	Média	0,66	0,97	1,28	0,70	1,00	1,29
	Desvio padrão	0,93	1,30	21,41	0,96	1,32	1,59

Fonte dos dados básicos: PNAD - IBGE.

## APÊNDICE F – Codificação das variáveis usadas nas regressões dos três exercícios

Metropolitano	1	domicílio pertence à região metropolitana
	0	domicílio não pertence à região metropolitana
Urbano	1	domicílio pertence à zona urbana
	0	domicílio pertence à zona rural
Água	1	domicílio possui água encanada
	0	domicílio não possui água encanada
Luz	1	domicílio possui energia elétrica
	0	domicílio não possui energia elétrica
Esgoto	1	domicílio possui esgotamento sanitário
	0	domicílio não possui esgotamento sanitário
Lixo	1	domicílio possui coleta de lixo
	0	domicílio não possui coleta de lixo
Tipo de família	1	casal sem filhos
	2	casal com filhos
	3	mãe com filhos
	4	outros tipos de família
Raça	1	branco ou amarelo
	0	negro, pardo ou indígena
Sexo	1	feminino
	0	masculino
Condição de ocupação do chefe	1	ocupado
	0	desocupado
Escolaridade do chefe do domicílio	0	analfabeto
	1	ensino fundamental completo ou incompleto
	2	ensino médio completo ou incompleto
	3	ensino superior incompleto
	4	ensino superior completo ou pós-graduação
Raça do chefe do domicílio	1	branco ou amarelo
	0	negro, pardo ou indígena
Sexo do chefe do domicílio	1	feminino
	0	masculino
Ano	1	2011
	0	2007
Tratamento	1	MG
	0	SP ou BA
Interação	1	Observação de MG em 2011
	0	Observação de SP ou BA em 2007