

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS
ANÍSIO TEIXEIRA

NOTA TÉCNICA Nº 8/2017/CGCQTI/DEED

Brasília, 27 de junho de 2017.

Assunto: Estimativas de fluxo escolar a partir do acompanhamento longitudinal dos registros de aluno do Censo Escolar do período 2007-2016.

INTRODUÇÃO

O presente estudo descreve a metodologia de tratamento e adequação dos dados do Censo da Educação Básica, a partir da sua edição de 2007, com fins de possibilitar o desenvolvimento de análises longitudinais e o cálculo de indicadores referentes à trajetória escolar dos alunos, baseados nas informações dos registros escolares e administrativos captados pela pesquisa. Além de possibilitarem análises mais detalhadas, permitindo, por exemplo, a desagregação das taxas de transição (promoção, repetência, evasão e migração para EJA) por características dos estudantes e níveis territoriais anteriormente inviáveis, o trabalho realizado abre possibilidades para novas linhas de pesquisas científicas, que poderão por em prova novas hipóteses sobre a relação entre a trajetória escolar e os resultados educacionais, em diferentes contextos, além da possibilidade de avaliar a eficiência dos sistemas de ensino em manter os alunos na escola e na sua capacidade de promover uma trajetória regular nas diferentes etapas de ensino.

Cabe destacar que esse estudo e os produtos derivados dele materializam parte dos objetivos iniciais já vislumbrados na proposição de alteração da metodologia do Censo da Educação Básica, ocorrida em 2007. Alguns textos, à época, já apontavam que a mudança metodológica possibilitaria o acompanhamento da permanência do aluno na escola, de seu rendimento escolar, do descompasso na relação idade/série e o estudo das relações desses fenômenos com outras condições internas e externas ao sistema de ensino¹.

Os indicadores de transição, equivalentemente denominados como indicadores de fluxo escolar, têm por objetivo informar sobre a trajetória do estudante na educação básica, que por sua vez, depende da estrutura organizacional dos programas de educação formal (sistemática), ou educação escolar², normalmente operacionalizada em níveis de ensino. Os níveis de ensino representam grandes etapas de progressão educativa em termos de complexidade do conteúdo educacional e estão relacionados aos estágios do desenvolvimento cognitivo do indivíduo, os quais estão associados à sua idade (maturação) biológica. Os níveis de ensino, portanto, constituem-se em um conjunto de categorias ordenadas, definidas em relação a gradações de experiências e conhecimentos, habilidades e competências projetadas para serem vivenciadas e aprendidas em um determinado programa educacional em uma determinada idade. Dessa forma,

¹ INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. Experiência: Sistema de coleta on-line do Censo Escolar da Educação Básica – Educacenso. ENAP. Brasil 2010. Disponível em: <http://repositorio.enap.gov.br/handle/1/280>

² A educação escolar compõe-se de: I - educação básica, formada pela educação infantil, ensino fundamental e ensino médio; II - educação superior. (Lei 9.394, de 20 de dezembro de 1996. Art. 21.).

a trajetória do estudante no sistema de ensino, que pode ser regular (transição de séries de acordo com a idade do estudante sem repetência ou evasão) ou não, traz informações tanto sobre o nível de conhecimento agregado como sobre a eficiência dos sistemas de ensino sendo, portanto, objeto de estudo acadêmico e de análise de políticas públicas educacionais.

A seguir são apresentadas as referências legais relativas ao tema e na seção de metodologia são descritos o processo de deduplicação de registros – que é realizado regularmente no cadastro de pessoas do Inep e tem importância para o acompanhamento dos estudantes ao longo dos anos –, as bases de dados e variáveis utilizadas na construção da base de dados longitudinal, o processo de imputação e as definições e metodologia de cálculo dos indicadores de fluxo escolar. A seção de resultados apresenta dados gerais dos indicadores calculados e nas considerações finais são destacados alguns aspectos da divulgação.

I- Referência legal

A relação entre nível de ensino e idade do indivíduo, no sistema educacional brasileiro, materializa-se na forma da lei da seguinte maneira:

- 1) O dever do Estado com educação escolar pública será efetivado mediante a garantia de educação básica obrigatória e gratuita dos 4 (quatro) aos 17 (dezessete) anos de idade [...] (Artigo 4º da Lei 9.394/1996 em seu inciso primeiro, com redação dada pela Lei nº 12.796/2013; e inciso primeiro do artigo 208 da Constituição da República Federativa do Brasil de 1998, com redação modificada pela Emenda Constitucional nº 59/2009).
- 2) A educação infantil será oferecida em: creches [...] para crianças de até três anos de idade e pré-escolas para crianças de quatro e cinco anos de idade (Artigo 30, incisos I e II da Lei 9.394/1996, com redação dada pela Lei nº 12.796/2013).
- 3) O ensino fundamental obrigatório, com duração de 9 (nove) anos, gratuito na escola pública, iniciando-se aos 6 (seis) anos de idade [...] (Artigo 32 da Lei 9.394/1996, com redação dada pela Lei nº 11.274/2006).
- 4) O ensino médio, etapa final da educação básica, com duração mínima de três anos [...] (Artigo 35 da Lei 9.394/1996).

Portanto, a trajetória do aluno entre o primeiro e o nono ano do ensino fundamental se dá entre os 6 aos 14 anos de idade, se não houver atraso na trajetória escolar do aluno. Já o percurso entre a primeira e terceira série do ensino médio, como nível subsequente ao ensino fundamental, ocorreria, portanto, dos 15 aos 17 anos de idade.

Diante do exposto, a efetivação do direito do cidadão e do dever do Estado com a educação formal dar-se-ia, em um sistema educacional eficiente, mediante a garantia de igualdade de condições para o acesso e a permanência no sistema e uma trajetória escolar regular na educação básica obrigatória, gratuita e de qualidade, com ingresso aos 4 anos e conclusão aos 17 anos de idade.^{3,4,5}

Antes de apresentar a metodologia e os resultados do estudo do fluxo escolar empreendido, é importante esclarecer duas questões que permeiam o mesmo:

³ Fernandes, Reynaldo. Índice de Desenvolvimento da Educação Básica (Ideb). Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Série Documental. Textos para Discussão), 2007. 26 p. Disponível em: <<http://www.publicacoes.inep.gov.br/portal/download/503>>. Acesso em: 6 maio 2017.

⁴ As metas 1, 2, 3, 5, 7 e 8 do Plano Nacional de Educação e parte das suas respectivas estratégias, estão diretamente relacionadas à avaliação do vínculo idade e nível de ensino - além da referência ao tema em outras metas (Lei nº 13.005, de 25 de junho de 2014).

⁵ TODOS, PELA EDUCAÇÃO. Todos pela educação – 5 anos, 5 metas, 5 bandeiras. 2012. Disponível em: <https://www.todospelaeducacao.org.br/arquivos/biblioteca/tpe_5anos_final.pdf>. Acesso em: 6 maio 2017.

- 1) Em que pese a flexibilidade da organização da educação básica, possibilitada em lei⁶, a forma de organização mais comum no país é a organização da oferta de ensino por meio de séries anuais. A metodologia da coleta do Censo da Educação Básica aponta, inclusive, para a necessidade de ajustar o dado da situação real declarada, quando a organização não se dá em séries anuais, para a forma de séries e etapas anuais⁷.
- 2) É comum a divisão do ensino fundamental em dois segmentos, anos iniciais do ensino fundamental (composto pelos cinco primeiros anos) e anos finais (composto pelos quatro últimos anos). Essa separação, por exemplo, pode ser útil para a comparação internacional dos sistemas de ensino, uma vez que a taxa de escolarização do primeiro nível da Classificação Internacional Normalizada da Educação (ISCED 1) é representada, no Brasil, por indivíduos que concluíram as cinco primeiras séries do ensino fundamental e o segundo nível (ISCED 2), por aqueles que concluíram as quatro últimas séries anuais⁸.

Por fim, a Lei 9.394/1996 ainda dispõe que a classificação do aluno em qualquer série ou etapa, exceto a primeira do ensino fundamental, se dará por promoção dos alunos que cursaram, com aproveitamento, na própria escola, a etapa ou fase anterior; por transferência de candidatos procedentes de outras escolas; e, independente da escolarização anterior, mediante avaliação feita pela escola que defina o grau de desenvolvimento e experiência do candidato e permita a sua inscrição na série ou etapa adequada.

II- Metodologia

- a. Processo de manutenção do registro único do cadastro de alunos e docentes do Censo Escolar: deduplicação

A deduplicação é um processo realizado regularmente na base de dados cadastrais de alunos e docentes, que tem por objetivo a identificação e tratamento de registros duplicados (dois ou mais registros que se referem a um mesmo indivíduo). Em um cadastro de pessoas ideal todos os indivíduos devem preencher em algum campo o seu código de identificação único, que pode ser validado por uma base externa ou, minimamente, por algum dígito verificador. Isso permitiria que o sistema alertasse sobre a entrada de dados errôneos e também indicasse de forma precisa a pré-existência do cadastro na base – impedindo a duplicidade. O Censo Escolar já coleta alguns campos que podem ser utilizados como identificador único, como o CPF, a nova certidão de nascimento de 32 dígitos, o NIS e a própria identificação única do EDUCACENSO - sendo alguns melhores que outros como identificadores. Entretanto, nem todos os alunos possuem esses documentos e, assim, esses campos não podem ser obrigatórios no sistema.

Antes de cadastrar um aluno, o sistema EDUCACENSO obriga o usuário a realizar uma busca pelo aluno na base nacional, caso esse aluno não seja encontrado, o cadastro é liberado.

⁶ A educação básica poderá organizar-se em séries anuais, períodos semestrais, ciclos, alternância regular de períodos de estudos, grupos não-seriados, com base na idade, na competência e em outros critérios, ou por forma diversa de organização, sempre que o interesse do processo de aprendizagem assim o recomendar (artigo 23 da Lei nº 9.394, de 20 de dezembro de 1996).

⁷ Caderno de Instruções, Inep, 2017. Disponível em: <http://download.inep.gov.br/educacao_basica/censo_escolar/caderno_de_instrucoes/caderno_de_instrucoes_censo_escolar_2017.pdf>. Acesso em: 6 maio 2017).

⁸ International Standard Classification of Education - ISCED 2011. Instituto for Statistics (UIS)/Unesco, 2011. Disponível em: <http://uis.unesco.org/sites/default/files/documents/isced_2011_mapping_en_brazil_0.xlsx>. Acesso em: 6 maio 2017.

Quando um usuário preenche os campos de pesquisa sem o devido cuidado, este pode acabar não encontrando um aluno já cadastrado, fazendo com que o sistema libere a inserção de um aluno que já consta na base de dados. É possível que o sistema barre a entrada desse novo cadastro, por exemplo, se um CPF pré-existente for inserido. Entretanto, se isso não ocorrer, um cadastro duplo será criado.

O processo de deduplicação é aplicado na base de cadastro ao menos uma vez por ano e vem sendo aperfeiçoado desde 2007. O processo consiste na busca de pares de registros cujas informações em vários campos apresentam grande similaridade. Diversos campos são avaliados nesse processo, sendo alguns exemplos: nome, nome dos pais, data de nascimento, município de nascimento, município de residência e campos de documentação. Por exemplo, para avaliação da similaridade de dois nomes, são realizados os seguintes passos: 1) Testa-se a igualdade dos nomes após simplificação fonética (utiliza-se uma adaptação do algoritmo “BUSCABR”⁹, voltado para a língua portuguesa); 2) Utiliza-se a distância “levenshtein” entre os dois nomes, que avalia o número de modificações - inserção, supressão ou substituição de uma letra - necessárias para transformar um nome em outro. Atualmente, verifica-se até mesmo a similaridade entre as etapas frequentadas ao longo dos anos pelo par de registros suspeito de ser do mesmo aluno.

O resultado final desse processo é uma tabela, denominada **DEPARA**, que identifica pares de cadastros considerados como duplicidades. A primeira variável desta tabela indica o código da pessoa que será desconsiderada do cadastro e a segunda variável, o código da pessoa que será mantida no cadastro. Após esse processo, as matrículas dos alunos eventualmente desconsiderados serão atribuídas para o aluno mantido.

É importante destacar que, como esse é um procedimento recorrente, os microdados da pesquisa não são influenciados por esse processo, apenas as tabelas internas do Inep. Além disso, a deduplicação impacta diretamente o acompanhamento, já que quando se cria um cadastro duplo aquele mais antigo usualmente deixa de ser utilizado, criando assim, uma falsa impressão de evasão – quando o processo de deduplicação identifica esse caso, as trajetórias dos dois cadastros são reunificadas e a evasão não é mais perceptível.

b. Bases de dados e variáveis utilizadas

O primeiro passo para a construção da base de dados longitudinal com as informações do período 2007-2016 foi a padronização das variáveis e categorias coletadas na primeira e segunda etapa do Censo Escolar. Esse processo permitiu armazenar todas as informações de uso estatístico coletadas nesses 10 anos em cinco Tabelas Simplificadas (TS). As informações da matrícula inicial¹⁰ coletadas na primeira etapa do Censo Escolar, que tem como referência a última quarta feira do mês de maio, estão separadas nas tabelas de escolas (TS_CENSO_BASICO_ESCOLA), turmas (TS_CENSO_BASICO_TURMA), docentes (TS_CENSO_BASICO_DOCENTE) e matrículas (TS_CENSO_BASICO_MATRICULA). Os dados de rendimento (aprovação ou reprovação) e movimento (abandono, transferência, falecimento dentre outros) dos alunos ao final do ano letivo (segunda etapa do Censo Escolar¹¹) referentes a cada uma das matrículas declaradas na primeira etapa do Censo, ficam registrados na tabela TS_CENSO_BASICO_SITUACAO.

⁹ LUCENA, Fred. Desenvolvimento de uma função em pl/sql para busca fonética em português do Brasil. 2006. 23 f. Monografia (Especialização) - Curso de Especialização em Banco de Dados, Faculdades Integradas Barros Melo. Disponível em:

<<http://www.pedraagroindustrial.com.br/adminsite/uploads/observador/observadorgdfg1238761300.pdf>>.

Acesso em: 6 maio 2017.

¹⁰ Mais informações em: <http://inep.gov.br/web/guest/matricula-inicial>

¹¹ Mais informações em: <http://inep.gov.br/web/guest/situacao-do-aluno>

As quatro tabelas referentes à primeira etapa do Censo Escolar possuem praticamente a mesma estrutura dos microdados de 2016 (<http://inep.gov.br/web/guest/microdados>), já que o processo de padronização buscou adaptar o formato das tabelas de anos anteriores ao formato do ano mais recente. Existem, no entanto, peculiaridades em cada ano, podendo haver diferenças relativas ao conjunto de variáveis coletadas (variáveis não coletadas em determinado ano possuem apenas valores do tipo *missing*) ou quanto ao conjunto de categorias de resposta de uma variável. Na maioria dos casos, as categorias de resposta foram ajustadas para corresponder ao dicionário mais recente, porém, quando isso não foi possível, foram estabelecidas categorias diferentes ao longo dos anos.

Considerando o interesse na trajetória dos alunos, a construção da base de dados longitudinal considera apenas os dados constantes nas tabelas simplificadas de matrícula (TS_CENSO_BASICO_MATRICULA) e de situação do aluno (TS_CENSO_BASICO_SITUACAO). Dessas tabelas foram selecionadas apenas as matrículas de escolarização (não contemplando matrículas em turmas de atividade complementar e atendimento educacional especializado) e o conjunto de variáveis apresentadas na Tabela 1.

Tabela 1 - Conjunto de variáveis selecionadas das tabelas TS_CENSO_BASICO_MATRICULA e TS_CENSO_BASICO_SITUACAO.

Nome da variável	Descrição
NU_ANO_CENSO	Ano do Censo
CO_PESSOA_FISICA	Código do aluno (identificação única)
ID_MATRICULA	Código da matrícula
CO_UF	Código da UF da escola
CO_MUNICIPIO	Código do município da escola
CO_ENTIDADE	Código da escola
TP_DEPENDENCIA	Dependência administrativa da escola - categorias: federal, estadual, municipal ou privada.
TP_LOCALIZACAO	Localização da escola - categorias: urbana ou rural.
TP_MEDIACAO_DIDATICO_PEDAGO	Tipo de mediação didático-pedagógica - categorias: presencial, semipresencial ou Educação a Distância.
IN_ESPECIAL_EXCLUSIVA	Aluno de turma exclusiva de alunos com deficiência, transtorno global do desenvolvimento ou altas habilidades/superdotação (Classes Especiais) - categorias: 0-Não; 1-Sim.
IN_REGULAR	Modo, maneira ou metodologia de ensino correspondente às turmas com etapas de escolarização consecutivas, Creche ao Ensino Médio. Etapas consideradas: TP_ETAPA_ENSINO igual a 1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 14, 15, 16, 17, 18, 19, 20, 21, 41, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37 ou 38 - categorias: 0-Não; 1-Sim.
IN_EJA	Modo, maneira ou metodologia de ensino correspondente às turmas destinadas a pessoas que não cursaram o ensino fundamental e/ou médio em idade própria. Etapas consideradas: TP_ETAPA_ENSINO igual a 65, 67, 69, 70, 71, 73 ou 74 - categorias: 0-Não; 1-Sim.
IN_PROFISSIONALIZANTE	Modo profissionalizante de ensino correspondente às turmas de cursos de formação inicial e continuada ou de qualificação profissional (Cursos FIC) articulados à EJA ou concomitantes; ou de cursos técnicos de nível médio nas formas articulada (integrada ou concomitante) ou subsequente ao ensino médio e de normal/magistério. Etapas consideradas: TP_ETAPA_ENSINO igual a 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 65, 67, 68, 73 ou 74 - categorias: 0-Não; 1-Sim.
TP_ETAPA_ENSINO	Etapas de ensino (ver Tabela 2)
TP_SITUACAO*	Situação de rendimento ou movimento do aluno ao final do ano letivo - categorias: aprovado, reprovado, abandono, falecido ou Sem Informação de Rendimento (SIR).
IN_CONCLUINTE*	Situação de conclusão – categorias: 0-Não; 1-Sim.

Nota: *variáveis presentes apenas na tabela TS_CENSO_BASICO_SITUACAO - demais variáveis estão presentes nas duas tabelas.

Embora não seja usual, alguns alunos mudam de escola ou de etapa de ensino entre a data de referência da pesquisa e o final do ano letivo. Essas mudanças, que são relevantes para avaliar a transição de etapas de um aluno entre o final de um ano e o início de outro, podem ser

avaliadas comparando as informações prestadas na primeira (TS_CENSO_BASICO_MATRICULA) e segunda etapa do Censo Escolar (TS_CENSO_BASICO_SITUACAO). A variável ID_MATRICULA é importante nesses casos, já que mesmo que o aluno seja transferido para outra escola o seu código de matrícula é preservado – de modo que a quantidade de matrículas de escolarização na tabela TS_CENSO_BASICO_SITUACAO é mesma da tabela TS_CENSO_BASICO_MATRICULA.

Para simplificar o posterior cálculo das taxas de transição, a variável TP_ETAPA_ENSINO (que indica a etapa de ensino que cada aluno está matriculado) foi recodificada de acordo com a tabela 2, buscando manter uma ordenação numérica que corresponda de alguma forma a evolução natural dos alunos. As etapas do ensino fundamental de 8 e 9 anos foram compatibilizadas igualando a primeira série do fundamental de 8 anos ao segundo ano do fundamental de 9 anos. Na nova codificação também não há mais distinção entre as categorias normal/magistério e integrado à educação profissional do ensino médio. As demais etapas, que são de menor interesse para o cálculo das taxas de transição, foram agregadas, mantendo um menor detalhamento dos códigos.

Tabela 2 - Recodificação da variável TP_ETAPA_ENSINO para auxiliar a construção da base de dados longitudinal e o cálculo dos indicadores de fluxo escolar (etapas seriadas de interesse, para as quais as taxas de transição serão divulgadas, são destacadas em cinza).

Variável auxiliar TP_ETAPA_FLUXO		Codificação original da variável TP_ETAPA_ENSINO	
Cód.	Descrição	Cód.	Descrição
1101	Educação Infantil – Creche	1	Educação Infantil - Creche (0 a 3 anos)
1102	Educação Infantil - Pré-escola	2	Educação Infantil - Pré-escola (4 e 5 anos)
1211	Ensino Fundamental - 1º Ano	14	Ensino Fundamental de 9 anos - 1º Ano
1212	Ensino Fundamental - 2º Ano	4	Ensino Fundamental de 8 anos - 1ª Série
		15	Ensino Fundamental de 9 anos - 2º Ano
1213	Ensino Fundamental - 3º Ano	5	Ensino Fundamental de 8 anos - 2ª Série
		16	Ensino Fundamental de 9 anos - 3º Ano
1214	Ensino Fundamental - 4º Ano	6	Ensino Fundamental de 8 anos - 3ª Série
		17	Ensino Fundamental de 9 anos - 4º Ano
1215	Ensino Fundamental - 5º Ano	7	Ensino Fundamental de 8 anos - 4ª Série
		18	Ensino Fundamental de 9 anos - 5º Ano
1221	Ensino Fundamental - 6º Ano	8	Ensino Fundamental de 8 anos - 5ª Série
		19	Ensino Fundamental de 9 anos - 6º Ano
1222	Ensino Fundamental - 7º Ano	9	Ensino Fundamental de 8 anos - 6ª Série
		20	Ensino Fundamental de 9 anos - 7º Ano
1223	Ensino Fundamental - 8º Ano	10	Ensino Fundamental de 8 anos - 7ª Série
		21	Ensino Fundamental de 9 anos - 8º Ano
1224	Ensino Fundamental - 9º Ano	11	Ensino Fundamental de 8 anos - 8ª Série
		41	Ensino Fundamental de 9 anos - 9º Ano
1301	Ensino Médio - 1ª Série	25	Ensino Médio - 1ª Série
		30	Curso Técnico Integrado (Ensino Médio Integrado) 1ª Série
		35	Ensino Médio - Normal/Magistério 1ª Série
1302	Ensino Médio - 2ª Série	26	Ensino Médio - 2ª Série
		31	Curso Técnico Integrado (Ensino Médio Integrado) 2ª Série
		36	Ensino Médio - Normal/Magistério 2ª Série
1303	Ensino Médio - 3ª Série	27	Ensino Médio - 3ª Série
		32	Curso Técnico Integrado (Ensino Médio Integrado) 3ª Série
		37	Ensino Médio - Normal/Magistério 3ª Série
1304	Ensino Médio - 4ª Série	28	Ensino Médio - 4ª Série
		33	Curso Técnico Integrado (Ensino Médio Integrado) 4ª Série
		38	Ensino Médio - Normal/Magistério 4ª Série
1305	Ensino Médio - Não Seriada	29	Ensino Médio - Não Seriada
		34	Curso Técnico Integrado (Ensino Médio Integrado) Não Seriada
2001	Educação Profissional - Concomitante	39	Curso Técnico - Concomitante
		68	Curso FIC Concomitante

Variável auxiliar TP_ETAPA_FLUXO		Codificação original da variável TP_ETAPA_ENSINO	
Cód.	Descrição	Cód.	Descrição
2002	Educação Profissional - Subsequente	40	Curso Técnico - Subsequente
3200	EJA - Ensino Fundamental	65	EJA - Ensino Fundamental - Projovem Urbano
		69	EJA - Ensino Fundamental - Anos iniciais
		70	EJA - Ensino Fundamental - Anos finais
		72	EJA - Ensino Fundamental - Anos iniciais e Anos finais ¹
		73	Curso FIC integrado na modalidade EJA - Nível Fundamental
3300	EJA - Ensino Médio	67	Curso FIC integrado na modalidade EJA - Nível Médio
		71	EJA - Ensino Médio
		74	Curso Técnico Integrado na Modalidade EJA

Nota: 1) Código existente na tabela de matrículas apenas em 2007 (de 2008 a 2015 os alunos de turmas com código "72 - EJA Anos iniciais e Anos finais" foram registrados separadamente como "69 - EJA anos iniciais" ou "70 - EJA anos finais").

Nos casos eventuais de alunos com mais de uma matrícula de escolarização no mesmo ano – que representam menos de 1% da base de dados – optou-se por manter na base longitudinal apenas uma delas para evitar que múltiplas trajetórias fossem geradas. A escolha da matrícula foi realizada por meio da tabela TS_CENSO_BASICO_SITUACAO, cujos dados foram primeiramente ordenados pelas seguintes variáveis na sequência apresentada (sentido da ordenação entre parênteses): NU_ANO_CENSO (crescente), CO_PESSOA_FISICA (crescente), IN_REGULAR (decrecente), IN_EJA (crescente), IN_ESPECIAL_EXCLUSIVA (crescente), variável indicadora (valores zero ou um) das etapas seriadas de interesse (ver Tabela 2 – decrecente), TP_SITUACAO (aprovado > reprovado > abandono), TP_ETAPA_FLUXO (decrecente), ID_MATRICULA (crescente).

Após a ordenação, foi mantido na base apenas o primeiro registro de cada aluno em cada ano. Implicitamente, na existência de mais de uma matrícula para o mesmo aluno, esse procedimento determina uma sequência de prioridades adotadas para a seleção da matrícula que permanecerá na base: 1) a matrícula na modalidade regular é preferida em relação àquela da EJA; 2) a matrícula em turma não exclusiva de alunos com deficiência é preferida em relação àquela exclusiva; 3) matrícula das etapas seriadas do ensino fundamental ou médio têm preferência sobre as demais etapas; 4) é preferida a matrícula em que o aluno foi aprovado, em detrimento daquela em que ele foi reprovado ou deixou de frequentar; 5) são priorizadas as matrículas de etapas de ensino mais avançadas; 6) caso as regras anteriores não definam a ordenação, foi mantida a primeira matrícula registrada no sistema (menor código de matrícula). A base de dados resultante deste processo agrega as informações do período 2007-2016 e permitiu a avaliação das trajetórias dos alunos por meio da identificação única dos mesmos (CO_PESSOA_FISICA).

c. Processo de imputação

Após os tratamentos realizados na base do fluxo escolar (remoção de alunos duplos e de múltiplos vínculos de matrícula em cada ano) verificou-se a presença de inconsistências na trajetória de alguns alunos na base no período 2007-2015. Essas inconsistências estão basicamente relacionadas a irregularidades na trajetória dos alunos que são perceptíveis pelo acompanhamento da evolução das etapas de ensino no período. Os dados inconsistentes foram submetidos a um procedimento de imputação denominado “hot deck”¹², que busca preservar a distribuição original e a coerência dos dados. De forma resumida, as informações ausentes de um aluno são substituídas por informações válidas de outros alunos (“doadores”) que possuem características similares a esse.

¹² LITTLE, Roderick JA; RUBIN, Donald B. **Statistical analysis with missing data**. John Wiley & Sons, 2014.

No caso específico do estudo do fluxo escolar, a adoção dessa abordagem de imputação é importante por preservar a consistência das informações de escola, modalidade e etapas imputadas em cada ano. O procedimento impõe restrições que garantem, por exemplo, que as informações imputadas de um determinado aluno sejam obtidas de outro aluno que em algum momento esteve na mesma escola, etapa de ensino e modalidade do primeiro – garantindo assim, que a trajetória imputada seja uma trajetória plausível para o aluno.

Como primeiro passo para aplicação do procedimento proposto, foram separados da base do fluxo escolar os alunos cuja trajetória foi considerada inconsistente e passível de tratamento por meio de imputação. De forma sucinta, essas trajetórias passíveis de imputação contemplam duas situações: a) alunos que aparecem pela primeira vez na base após 2007 em etapa igual ou superior ao terceiro ano do fundamental¹³ com idade adequada para etapa – para os quais se esperaria uma trajetória regular (frequência à escola regular sem repetência ou evasão) até a etapa em que aparece pela primeira vez no período 2007-2015; b) alunos que frequentavam a escola em alguma etapa seriada de interesse, deixaram de frequentar por um ou mais anos e voltaram a frequentar em uma etapa muito superior, que sugere frequência à escola com promoção em pelo menos um dos anos em que não há registros. A Tabela 3 apresenta em detalhes os cinco tipos de trajetórias consideradas como passíveis de imputação.

A segunda etapa consistiu na definição de uma segunda base de dados, denominada **ELEGÍVEIS**, contemplando os dados de alunos com trajetórias sem inconsistências e que são possíveis “doadores” de informações. Para a criação da base **ELEGÍVEIS** foram desconsideradas trajetórias classificadas como não usuais, como aquelas que possuíam pulos entre as etapas seriadas (por exemplo, alunos que passam do sexto ano do fundamental para o oitavo ano do fundamental) ou ainda, aquelas em que o aluno regride de etapa de ensino (por exemplo, passando do terceiro ano do médio para o segundo ano do médio).

O processo de imputação pode ser resumido nas seguintes etapas:

- 1) Para cada aluno com dados inconsistentes, foi selecionado da base **ELEGÍVEIS** um conjunto de alunos que cumpriam os critérios de elegibilidade¹⁴ definidos (ver Tabela 3) - que são potenciais “doadores” de informação para o aluno com inconsistência.
- 2) Dentre os alunos selecionados da base **ELEGÍVEIS**, um único aluno foi sorteado aleatoriamente, sendo esse denominado “doador”.
- 3) As informações de matrícula cada aluno da base que estavam ausentes em um ou mais anos (código da escola, código da modalidade, código da etapa e possivelmente a situação final do aluno) foram substituídas pelas informações do aluno “doador”.

A Tabela 3 também apresenta para cada tipo de inconsistência o percentual de casos em que o processo de imputação foi realizado com sucesso, isto é, para os quais foi possível encontrar ao menos um aluno “doador” de informações que cumpriam os critérios estabelecidos.

¹³ Como o 2º ano do ensino fundamental de 9 anos é colocado como equivalente à primeira série do ensino fundamental de 8 anos optou-se por tratar como inconsistentes apenas entradas a partir do terceiro ano do fundamental de 9 anos.

¹⁴ Caso nenhum aluno cumpra os critérios estabelecidos, a imputação não é realizada e o aluno, portanto, permanece com a trajetória inconsistente.

Tabela 3 - Trajetórias passíveis de imputação, total de ocorrências, critérios de elegibilidade e total de casos imputados, Censo Escolar, 2007-2015.

Tipos de trajetórias passíveis de imputação	Total de ocorrências ¹	Crítérios de elegibilidade	Casos imputados	%
1) No ano <i>t</i> o aluno <i>i</i> está matriculado em uma das etapas de ensino de interesse, deixa de frequentar a escola por um ou mais anos e volta a frequentar (ano <i>t+n</i>) em etapa de ensino, tal que, comparada com a etapa inicial, indicaria uma trajetória regular do aluno na escola, isto é, sem evasão ou repetência.	2.820.871	Alunos da base ELEGÍVEIS que, quando comparados ao perfil do aluno inconsistente, possuam: mesma escola, modalidade e etapa no ano de retorno (<i>t+n</i>) e mesma etapa no ano de evasão (<i>t</i>).	2.678.170	94,9
2) No ano <i>t</i> o aluno <i>i</i> está matriculado em uma das etapas de ensino de interesse, deixa de frequentar a escola por dois ou mais anos e volta a frequentar (ano <i>t+n</i>) em etapa pelo menos duas séries acima da etapa inicial, de tal forma que se presume que o aluno estava matriculado em pelo menos um dos anos do período de ausência, mas que sua trajetória escolar não foi regular, isto é, há ao menos uma evasão ou repetência no período.	235.908	Alunos da base ELEGÍVEIS que, quando comparados ao perfil do aluno inconsistente, possuam: mesma escola, modalidade e etapa no ano de retorno (<i>t+n</i>) e mesma etapa e situação de rendimento no ano de evasão (<i>t</i>).	159.747	67,7
3) No ano <i>t</i> o aluno <i>i</i> está matriculado em uma das etapas de ensino de interesse, possui situação final de reprovado no mesmo ano, deixa de frequentar a escola por um ano e volta a frequentar em etapa superior.	268.779	Alunos da base ELEGÍVEIS que, quando comparados ao perfil do aluno inconsistente, possuam: mesma escola, modalidade e etapa no ano de retorno (<i>t+2</i>) e que no ano anterior (<i>t+1</i>) frequentava a escola em etapa de ensino imediatamente anterior.	262.937	97,8
4) O aluno <i>i</i> aparece pela primeira vez na base após 2007 em etapa igual ou superior ao terceiro ano do fundamental com evidência de inconsistência ² .	2.560.754	Alunos da base ELEGÍVEIS que, quando comparados ao perfil do aluno inconsistente, possuam: mesma escola, modalidade, etapa no ano de entrada (<i>t</i>) e grupo de idade ³ compatível.	2.326.464	90,9
5) No ano <i>t</i> o aluno <i>i</i> está matriculado na creche ou na pré-escola deixa de frequentar a escola por um ou mais anos e volta a frequentar (ano <i>t+n</i> ; <i>n</i> >1) em etapa igual ou superior ao terceiro ano do ensino fundamental.	214.610	Alunos da base ELEGÍVEIS que, quando comparados ao perfil do aluno inconsistente, possuam: mesma escola, modalidade, etapa no ano de retorno (ano <i>t+n</i> ; <i>n</i> >1) e grupo de idade ³ compatível.	191.537	89,2

Notas: 1) Número de vezes em que o tipo de trajetória inconsistente foi identificado - o mesmo aluno pode apresentar mais de uma inconsistência no período; 2) Para alunos com idade adequada para a etapa, a situação apresentada foi considerada como inconsistente e para aqueles com distorção idade-série, foi avaliado se o período em que o aluno permaneceu não matriculado correspondia ou não a uma situação de inconsistência (comparando esse período com o número de anos de distorção idade-série); 3) Foram considerados três grupos de idade: idade adequada, 1 ano de distorção, 2 a 4 anos de distorção para a etapa.

d. Definições e metodologia de cálculo dos indicadores de fluxo escolar

O acompanhamento da trajetória de um aluno entre o ano t , denominado aqui como ano base, e o ano $t + 1$, foi realizado por meio da comparação entre as informações de escola, modalidade e principalmente da etapa de ensino registrada na situação final do aluno do ano t e a etapa registrada na matrícula inicial do ano $t + 1$. As definições de promoção, repetência e evasão adotadas aqui, seguem os conceitos adotados na literatura^{15,16,17,18} relativa ao tema. Separa-se aqui, no entanto, a evasão de escola, que considera apenas os alunos que deixaram de frequentar a escola no ano seguinte, e a migração para a EJA, que engloba alunos que saíram das etapas seriadas do ensino regular e migraram para a Educação de Jovens e Adultos no ano seguinte.

É importante destacar que as informações registradas no ano base (relativas à situação final do aluno) são tomadas como referência nos cálculos das taxas de transição. Por exemplo, a avaliação dos promovidos de uma determinada etapa e escola considera o total de alunos que na situação final estavam nesta mesma etapa e escola e que na matrícula inicial do ano seguinte estavam em etapa superior à do ano base - independente da escola que registrou a matrícula desse ano seguinte. Ainda, a avaliação das taxas de transição dos alunos (promoção, repetência, evasão e migração para a EJA) foi feita apenas para as etapas seriadas de interesse, ou seja, do primeiro ao nono ano do ensino fundamental e da primeira a terceira série do ensino médio. Assim, os alunos que na situação final do ano base estavam matriculados fora das etapas de interesse (ou seja, educação infantil, quarta série ou não seriado do ensino médio, educação profissional e EJA) não possuem informação de fluxo na base longitudinal (não aplicável).

Finalmente, dada a etapa de ensino k (em alguma das etapas de interesse¹⁹) da situação final de um aluno no ano t e a sua etapa da matrícula inicial do ano $t + 1$, foram adotadas as seguintes definições:

¹⁵ KLEIN, Ruben. Produção e Utilização de Indicadores Educacionais: Metodologia de Cálculo de Indicadores do Fluxo Escolar da Educação Básica. Revista Brasileira de Estudos Pedagógicos, v. 84, p. 107-157, 2003. Disponível em: <<http://rbep.inep.gov.br/index.php/rbep/article/view/893/868>>. Acesso em: 6 maio 2017.

¹⁶ INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. Geografia da Educação Brasileira 2001. Brasília, 2002. Disponível em: <<http://www.publicacoes.inep.gov.br/portal/download/254>>. Acesso em: 6 maio 2017.

¹⁷ RIBEIRO, Sérgio Costa. A pedagogia da repetência. Estud. Av., São Paulo, v. 5, n. 12, p. 07-21, Aug. 1991. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-40141991000200002&lng=em&nrm=isso>. Acesso em: 6 maio 2017.

¹⁸ THONSTAD, T. Analysing and projecting school enrolment in developing countries: a manual of methodology. Paris: Unesco, 1980. (Statistical Reports and Studies, n. 24). Disponível em: <<http://unesdoc.unesco.org/images/0004/000447/044736eo.pdf>>. Acesso em: 6 maio 2017.

¹⁹ Variável auxiliar TP_ETAPA_FLUXO igual a 1211, 1212, 1213, 1214, 1215, 1221, 1222, 1223, 1224, 1301, 1302 ou 1303.

- Aluno promovido:
 - i. Aluno na etapa k , no ano t , que no ano $t + 1$ esteja matriculado em etapa superior a k ;
 - ii. Aluno matriculado na terceira série do ensino médio no ano t , aprovado, concluinte ou que no ano $t + 1$ esteja matriculado na quarta série do ensino médio ou na educação profissional subsequente.

- Aluno repetente:

Aluno na etapa k , no ano t , que no ano $t + 1$ esteja matriculado em etapa igual ou inferior a k .

- Aluno evadido:

Aluno matriculado no ano t nas etapas seriadas de interesse que não se matricula no ano $t + 1$.

- Aluno que migrou para EJA:

Aluno matriculado no ano t nas etapas seriadas de interesse, que no ano $t + 1$ se matricula na EJA.

Situações excepcionais²⁰, como a transição de etapas seriadas para não seriadas, que apresentam baixa frequência na base de dados, foram tratadas caso a caso.

As taxas de transição²¹ (fluxo escolar) de uma dada etapa são calculadas considerando a proporção de alunos em cada uma das situações acima. Por exemplo, a taxa de promoção do primeiro ano do ensino fundamental para o Brasil, referente à transição entre 2014 e 2015, corresponde à divisão entre o número de alunos promovidos e o total de alunos dessa etapa em 2014 (que equivale à soma de promovidos, repetentes, evadidos e de alunos que migraram para EJA). Ou seja, corresponde à proporção de alunos que, estando matriculados no primeiro ano do ensino fundamental em 2014, foram promovidos. Valendo o mesmo raciocínio para as taxas de repetência, evasão e migração para a EJA.

²⁰ Observando os valores da variável auxiliar TP_ETAPA_FLUXO (ver Tabela 2), considera-se promovido: a) aluno matriculado no ensino fundamental no ano t que no ano $t + 1$ se matricule nas etapas 1305, 2001 ou 2002 – que foram tratadas como etapas de nível superior; b) aluno matriculado nas etapas 1301 ou 1302 no ano t que no ano $t + 1$ se matricule na etapa 2002 ou 1305 - sendo que para essa última é exigida a aprovação no ano t ; c) aluno matriculado na etapa 1303 no ano t , aprovado e que no ano $t + 1$ se matricule nas etapas 2001 ou na EJA – considerou-se que aprovação na etapa final prevalece. Considera-se repetente o aluno matriculado nas etapas 1301, 1302 ou 1303 no ano t , não aprovado que no ano $t + 1$ se matricula na etapa 1305. Considera-se evadido: a) aluno matriculado nas etapas 1301 ou 1302 no ano t , que no ano $t + 1$ se matricule na etapa 2001; b) aluno matriculado na etapa 1303 no ano t , não aprovado e que no ano $t + 1$ se matricule na etapa 2001. Finalmente são desconsiderados do cálculo das taxas: a) alunos matriculados no ano t , com situação de falecido e que no ano $t + 1$ não se matriculam em qualquer etapa; b) os alunos matriculados na etapa 1303 no ano t , sem informação de rendimento e movimento (SIR) que não se matriculam no ano $t + 1$ – devido à impossibilidade de avaliar a conclusão da etapa final.

²¹ Seguindo o padrão de divulgação das taxas de rendimento (aprovação, reprovação e abandono) as taxas de transição (promoção, repetência, migração para a EJA e evasão) não incluem no cálculo alunos que no ano base estavam em turmas exclusivas de alunos com deficiência, transtorno global do desenvolvimento ou altas habilidades/superdotação (Classes Especiais).

Tendo como enfoque os alunos matriculados no ano t e observando onde eles estavam no ano $t - 1$, podemos ainda classificá-los em dois tipos de alunos novos:

- Aluno novo:
 - Aluno não matriculado no ano $t - 1$, que no ano t esteja matriculado em qualquer etapa.
- Aluno novo na modalidade:
 - i. Aluno matriculado em qualquer etapa da Educação de Jovens e Adultos no ano $t - 1$, que no ano t esteja matriculado no ensino regular.
 - ii. Aluno matriculado em qualquer etapa do ensino regular no ano $t - 1$, que no ano t esteja matriculado na Educação de Jovens e Adultos.

Nos registros de um dado ano na base de dados longitudinal (identificado pela variável NU_ANO_CENSO), é possível identificar qual foi o fluxo de um aluno desse mesmo ano para o ano seguinte na variável TP_FLUXO (que indica a promoção, repetência, evasão ou migração para a EJA) e, ainda, a variável TP_NOVO indica se, nesse ano, ele é um aluno novo ou um aluno novo na modalidade.

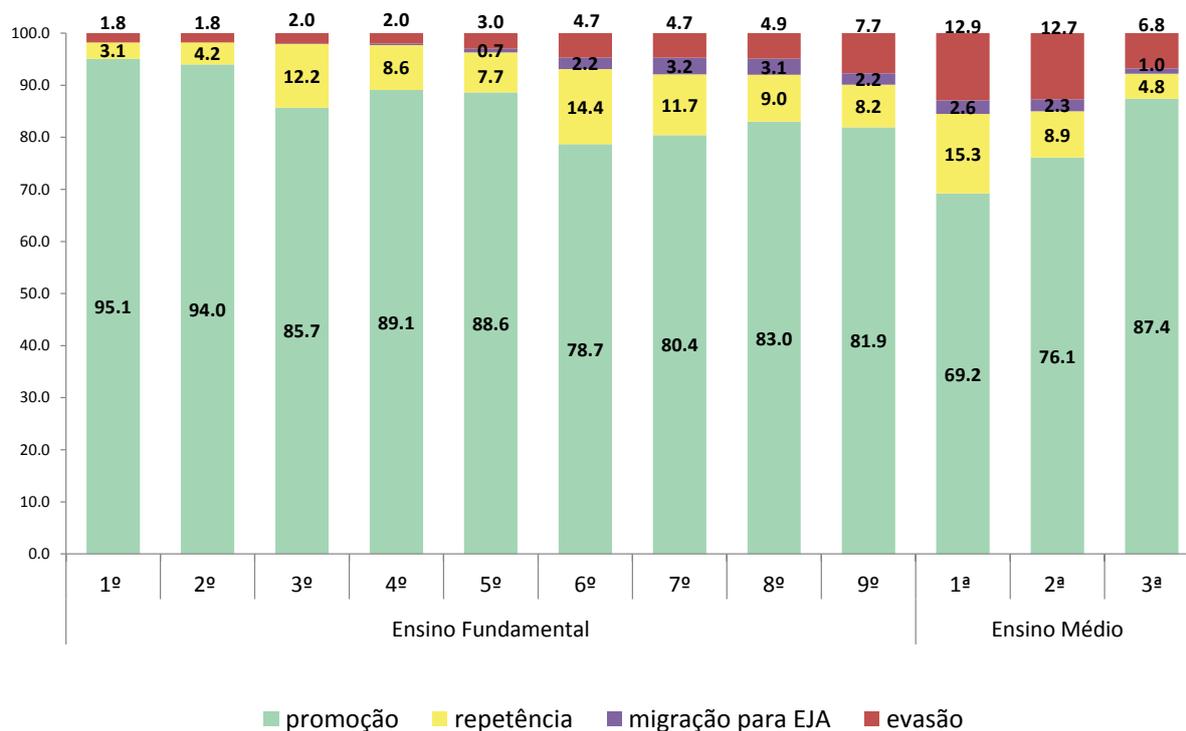
III- Resultados

A partir da base de dados longitudinal foram calculadas as taxas de transição (promoção, repetência, migração para a EJA e evasão) das séries do ensino fundamental e médio para os períodos 2007/2008, 2008/2009, 2010/2011, 2012/2013, 2013/2014 e 2014/2015 nos seguintes níveis territoriais: Brasil, regiões, Unidades da Federação e Municípios. Na divulgação inicial das taxas de transição no site do Inep (planilhas disponíveis em: <http://portal.inep.gov.br/web/guest/indicadores-educacionais>) os resultados para o Brasil e regiões foram desagregados pela dependência administrativa (Federal, Estadual, Municipal e Privada) e localização da escola (urbana e rural). Já os dados das Unidades da Federação e dos Municípios foram desagregados por rede de ensino (pública e privada) e localização da escola.

A Figura 1 apresenta as taxas de transição do Brasil por série do ensino fundamental e médio período 2014/2015. Observa-se uma taxa de promoção de 95,1% no primeiro ano do ensino fundamental no período 2014/2015, ou seja, de 2014 para 2015 95,1% dos alunos do primeiro ano avançaram para a etapa seguinte. No terceiro ano do ensino fundamental há uma queda brusca na taxa de promoção, que passa para 85,7%, o mesmo ocorrendo no sexto ano do ensino fundamental e na primeira série do ensino médio, onde as taxas de promoção são de 78,7% e 69,2%. A primeira e segunda série do ensino médio apresentam as maiores taxas de evasão (alunos que não se matriculam em qualquer escola no ano seguinte) com 12,9% e 12,7%, respectivamente. A migração para a EJA começa a ocorrer com maior força a partir sexto ano do ensino fundamental, etapa na qual se verifica uma taxa de migração de 2,2% - ou seja, aproximadamente 2 de cada 100 alunos que estavam matriculados nessa etapa em 2014 se matricularam na EJA em 2015.

O Apêndice 1 apresenta maiores detalhes sobre a transição entre as etapas seriadas de 2014 para 2015 e ainda permite comparar as taxas de transição antes e após os tratamentos realizados na base (deduplicação e imputação).

Figura 1 – Taxas de promoção, repetência, migração para EJA e evasão por ano/série - Brasil - Censo Escolar 2014/2015.



Apenas a título de ilustração, são apresentadas na Figura 2 as taxas de aprovação, reprovação e abandono de 2014 (cujos dados também estão disponíveis em: <http://portal.inep.gov.br/web/guest/indicadores-educacionais>). Observa-se, como já esperado, que as taxas de promoção apresentadas anteriormente na Figura 1 seguem o mesmo comportamento das taxas de aprovação da Figura 2, afinal, a maior parte dos alunos aprovados tendem a se matricular em uma etapa mais avançada no ano seguinte. Todavia, embora não seja usual, parte dos alunos aprovados podem acabar não se matriculando no ano seguinte (como nos casos dos evadidos aprovados). Outras situações ainda mais raras podem acontecer, como nos casos em que um aluno é aprovado, muda de escola, é reclassificado e acaba repetindo a série – algumas escolas adotam esse tipo de processo. Além disso, existem ainda alunos que são aprovados e acabam migrando para a EJA. Assim, não é estranho que as taxas de promoção (que medem a progressão efetiva entre as etapas seriadas) sejam invariavelmente iguais ou inferiores às taxas de aprovação.

As Figuras 3 e 4 apresentam as taxas de transição do Brasil para os anos iniciais e finais do ensino fundamental e para o ensino médio, desagregadas por dependência administrativa da escola e localização. Para essas taxas, é importante observar que algumas redes de ensino têm maior participação em determinadas etapas do que outras e também observar que o número de alunos matriculados em escolas da zona rural é muito inferior ao quantitativo nas escolas urbanas.

Figura 2 – Taxas de aprovação, reprovação e abandono por série (todas as redes de ensino) - Brasil - Censo Escolar 2014.

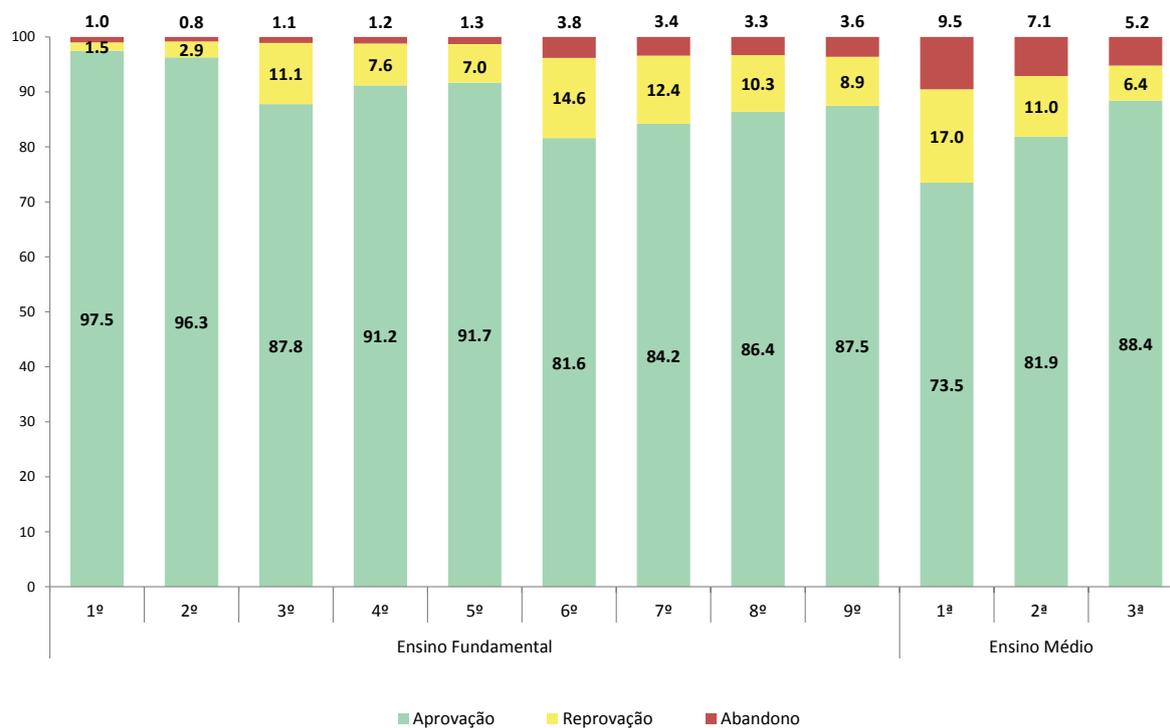


Figura 3 – Taxas de promoção, repetência, migração para EJA e evasão por etapa de ensino e dependência administrativa - Brasil - Censo Escolar 2014/2015.

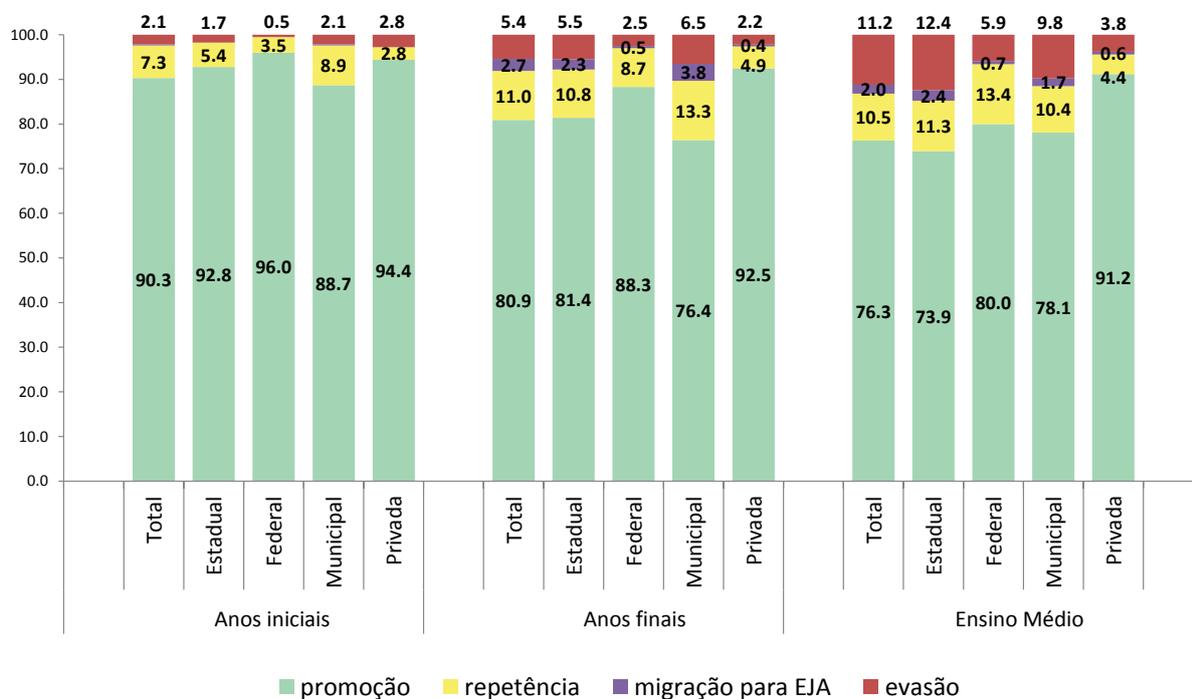
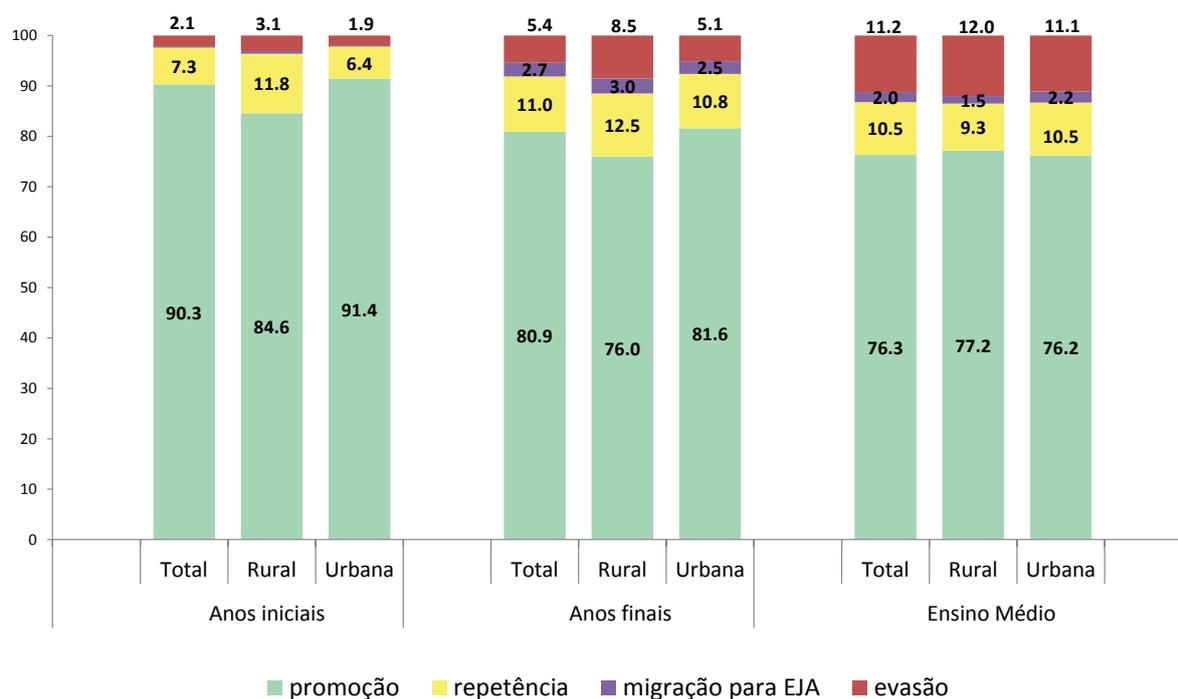


Figura 4 – Taxas de promoção, repetência, migração para EJA e evasão por etapa de ensino e localização (todas as redes de ensino) - Brasil - Censo Escolar 2014/2015.



É perceptível a evolução das taxas de promoção no Brasil nos últimos anos, embora haja uma tendência recente de estagnação (Figura 5). As taxas de migração para a EJA apresentam uma leve tendência de queda no período. Verifica-se também que os Estados possuem taxas de transição muito heterogêneas (Figuras 6, 7 e 8). A evasão do ensino médio, no entanto, é um problema comum a todos eles, com taxas variando entre 9,0% e 15,9%.

Figura 5 – Taxas de promoção, repetência, migração para EJA e evasão por etapa de ensino e dependência administrativa (todas as redes de ensino) - Brasil - Censo Escolar 2014/2015.

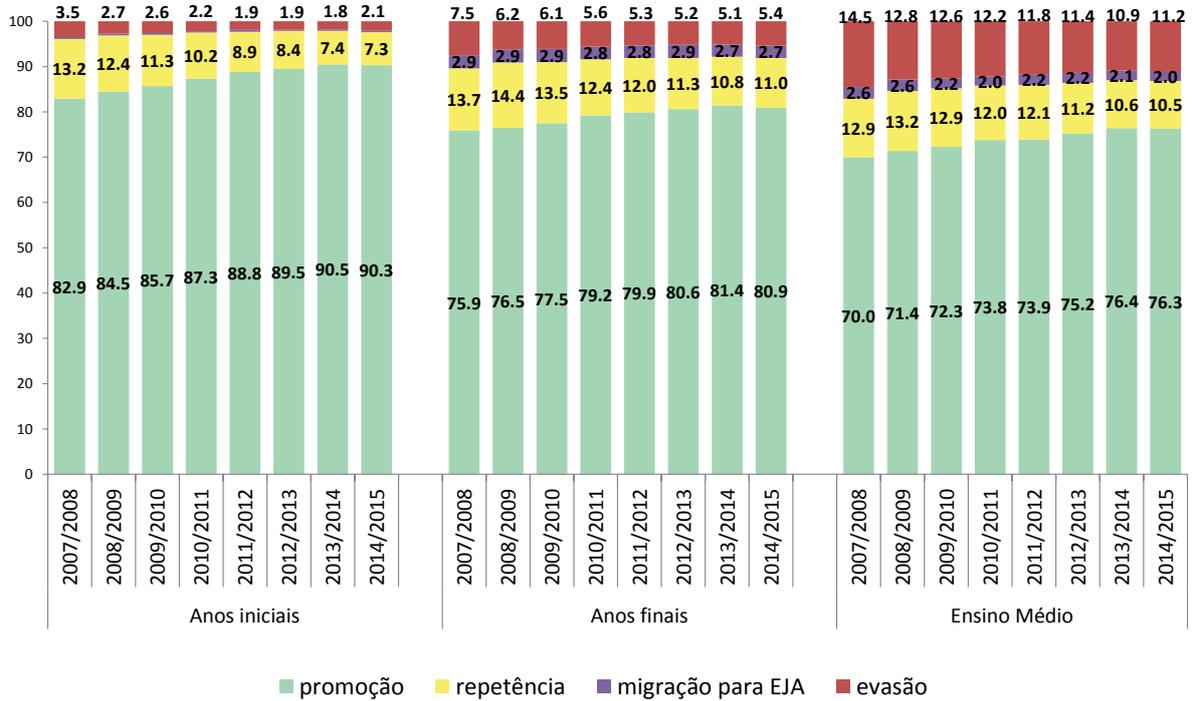


Figura 6 – Taxas de promoção, repetência, migração para EJA e evasão dos anos iniciais (todas as redes de ensino) - Brasil e Unidades da Federação - Censo Escolar 2014/2015.

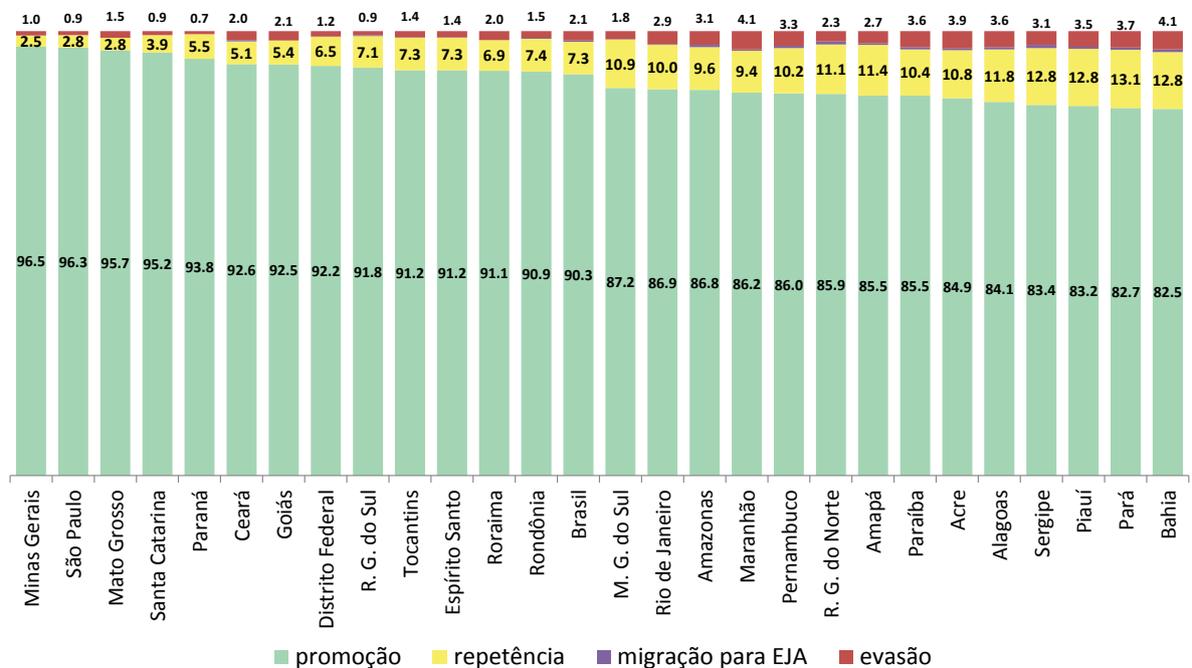


Figura 7 – Taxas de promoção, repetência, migração para EJA e evasão dos anos finais (todas as redes de ensino) - Brasil e Unidades da Federação - Censo Escolar 2014/2015.

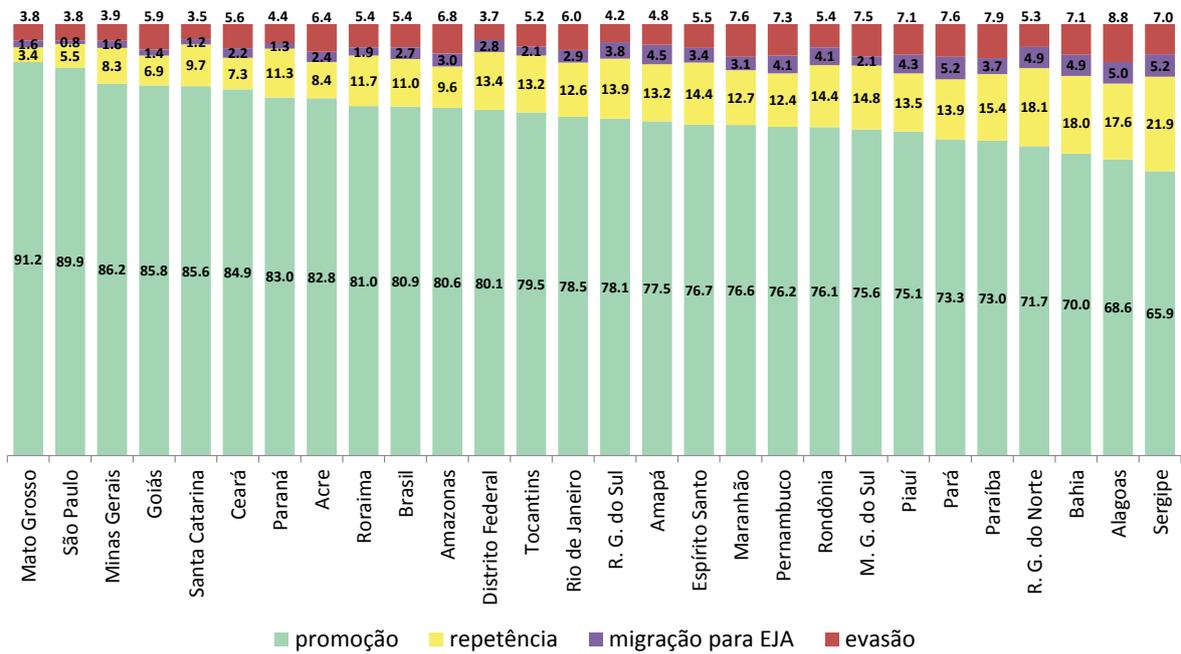
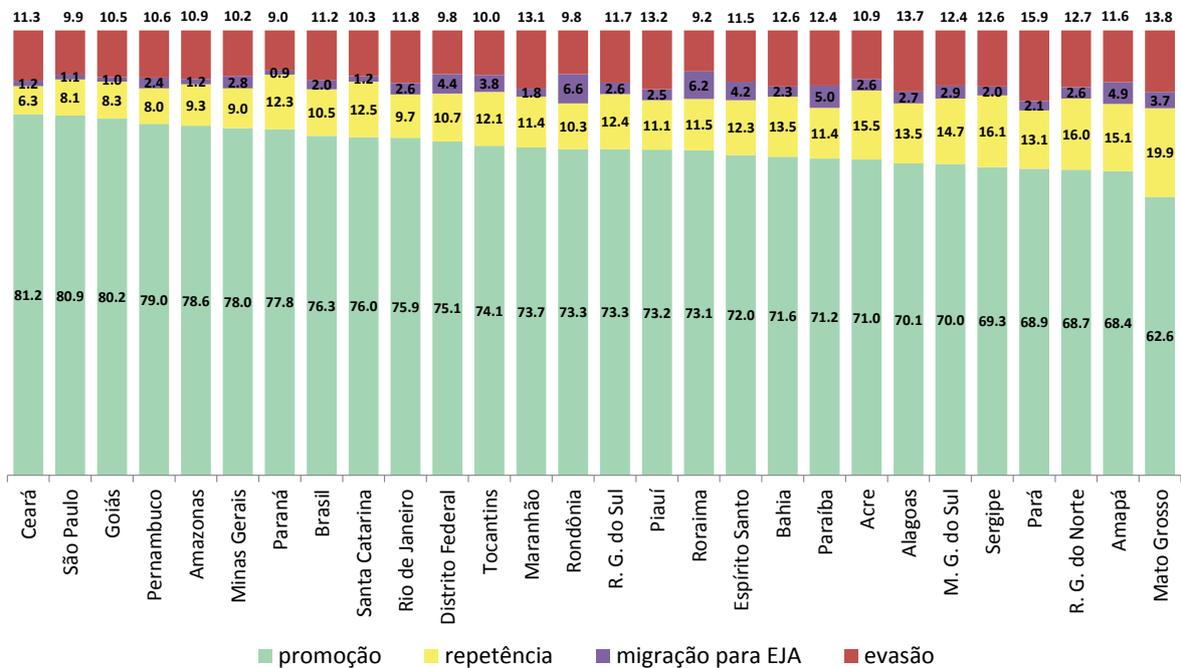


Figura 8 – Taxas de promoção, repetência, migração para EJA e evasão do ensino médio (todas as redes de ensino) - Brasil e Unidades da Federação - Censo Escolar 2014/2015.



IV- Considerações finais

Este documento tem como principal objetivo detalhar os tratamentos aplicados aos dados da primeira e segunda etapa do Censo Escolar da Educação Básica do período 2007 a 2016, visando o acompanhamento da trajetória dos estudantes. Além disso, busca-se apresentar claramente todas as definições adotadas quanto às taxas de transição: promoção, repetência, migração para a EJA e evasão.

Ressalta-se que os tratamentos de deduplicação e de imputação têm impacto importante na correção de falsas evasões – que, como explicado anteriormente, podem tanto ser ocasionadas pela criação de cadastros duplos na base, como pela existência de irregularidades na trajetória dos estudantes. Devido à necessidade desses tratamentos, a divulgação dos resultados das taxas de transição ocorrerá sempre com um ano de atraso, ou seja, a divulgação das taxas de transição do período 2015/2016 ocorrerá apenas após a divulgação da matrícula inicial de 2017 – já que o processo de deduplicação relativo por vezes não ocorre imediatamente após a divulgação dos dados da matrícula inicial e ainda os dados de 2017 são necessários para a correção de falsas evasões.

As taxas de transição por escola não foram divulgadas para respeitar a lógica do próprio indicador, que busca avaliar se as garantias de acesso, de permanência e de trajetória escolar regular estão sendo cumpridas nas redes de ensino (pública ou privada) e no território (escolas urbanas e rurais dos municípios, estados, regiões do Brasil). A divulgação dos microdados, no entanto, permitirá a desagregação das taxas de transição por características dos alunos (sexo, cor/raça e idade) ou mesmo de acordo com a situação final declarada pela escola (aprovação, reprovação e abandono).

Por fim, cabe ressaltar a importância da divulgação das taxas de transição para a formulação, monitoramento e avaliação de políticas públicas que visam assegurar o acesso à educação de qualidade no país. Além disso, a produção e disponibilização de uma base longitudinal, que permite o acompanhamento de milhões de alunos ao longo do tempo, favorece o desenvolvimento de estudos e pesquisas que possam contribuir, dentre outros aspectos, para compreender os fatores determinantes para uma trajetória regular.

Apêndice 1 – Nota sobre a análise de transição entre as etapas seriadas de interesse e o impacto dos tratamentos realizados na base de dados.

A tabela 4 apresenta as taxas de transição dos alunos nas etapas seriadas de interesse (primeiro ano do ensino fundamental até a terceira série do ensino médio) entre 2014 e 2015 além de apresentar as taxas de transição gerais do mesmo período - mesmos dados apresentados na Figura 1. Observa-se que nem todos os alunos possuem um fluxo natural entre as etapas, podendo haver tanto alunos que avançam duas ou mais séries no ano seguinte – que são tratados como promovidos – como também alunos que retrocedem uma ou mais séries no ano seguinte – que são tratados como repetentes. Essas situações, no entanto, apresentam uma baixa frequência na base de dados, por exemplo, dos alunos matriculados no primeiro ano do ensino fundamental em 2014, 95,1% foram considerados promovidos entre 2014 e 2015, sendo que, em 2015, 94,3% estão matriculados no segundo ano, 0,6% estão no terceiro ano e 0,1% estão no quarto ano do ensino fundamental e apenas 0,1% se encaixam em outras condições que também são classificadas como promoção.

Tabela 4 – Taxas de transição entre as séries de interesse (primeiro ano do ensino fundamental até a terceira série do ensino médio) e taxas de transição gerais (todas as redes de ensino) - base de dados do fluxo escolar - Brasil - Censo Escolar 2014-2015.

Taxas de transição (%) entre as séries de interesse 2014/2015 (percentuais em relação ao total de alunos matriculados na etapa em 2014 – valores arredondados na primeira casa decimal)													Taxas de transição gerais (%) 2014/2015			
Etapa de ensino ¹ 2014	Etapa de ensino ² 2015												Promoção ³	Repetência ⁴	Migração para EJA	Evasão
	1º EF	2º EF	3º EF	4º EF	5º EF	6º EF	7º EF	8º EF	9º EF	1ª EM	2ª EM	3ª EM				
1º EF	3,0	94,3	0,6	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	95,1	3,1	0,0	1,8
2º EF	0,2	4,0	93,4	0,5	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	94,0	4,2	0,0	1,8
3º EF	0,1	0,1	12,0	85,0	0,5	0,1	0,0	0,0	0,0	0,0	0,0	0,0	85,7	12,2	0,1	2,0
4º EF	0,0	0,0	0,1	8,4	88,4	0,6	0,0	0,0	0,0	0,0	0,0	0,0	89,1	8,6	0,3	2,0
5º EF	0,0	0,0	0,0	0,1	7,4	88,0	0,4	0,1	0,1	0,0	0,0	0,0	88,6	7,7	0,7	3,0
6º EF	0,0	0,0	0,0	0,0	0,1	14,3	77,8	0,7	0,2	0,0	0,0	0,0	78,7	14,4	2,2	4,7
7º EF	0,0	0,0	0,0	0,0	0,0	0,2	11,5	79,6	0,7	0,1	0,0	0,0	80,4	11,7	3,2	4,7
8º EF	0,0	0,0	0,0	0,0	0,0	0,0	0,1	8,8	82,1	0,9	0,0	0,0	83,0	9,0	3,1	4,9
9º EF	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,2	7,8	81,5	0,2	0,0	81,9	8,2	2,2	7,7
1ª EM	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	15,2	68,7	0,3	69,2	15,3	2,6	12,9
2ª EM	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,2	8,7	76,0	76,1	8,9	2,3	12,7
3ª EM	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	4,6	87,4	4,8	1,0	6,8

Nota: 1) Etapa de ensino registrada na situação final Censo Escolar da Educação Básica de 2014; 2) Etapa de ensino registrada na matrícula inicial do Censo Escolar da Educação Básica 2015; 3) O valor da taxa de promoção geral para cada uma das etapas de 2014 não bate com a soma dos percentuais onde a etapa de 2015 é superior à etapa de 2014 devido ao arredondamento dos valores na primeira casa decimal e às situações excepcionais tratadas na seção “Definições e metodologia de cálculo dos indicadores de fluxo escolar”; 4) O valor da taxa de repetência geral para cada uma das etapas de 2014 não bate com a soma dos percentuais onde a etapa de 2015 é igual ou inferior à etapa de 2014 devido ao arredondamento dos valores na primeira casa decimal e a situações excepcionais tratadas na seção “Definições e metodologia de cálculo dos indicadores de fluxo escolar”.

Destaca-se que todas as informações de matrícula presentes na base de dados do fluxo escolar correspondem às informações declaradas pelas redes de ensino e escolas na primeira e segunda etapa do Censo Escolar, ou seja não houve nenhum processo de edição das etapas ou modalidades de ensino informada. Todavia, como apresentado neste documento os processos de imputação (estudantes cuja informação de matrícula provém de outros estudantes de perfil similar), de deduplicação e de remoção de duplos vínculos impactam no quantitativo de alunos e matrículas presentes na base do fluxo escolar. Assim, os dados da base de fluxo escolar, que tem enfoque no acompanhamento longitudinal dos alunos, devem ser utilizados

com cautela para outras finalidades, já que, por exemplo, a contagem de matrículas por etapa de ensino, escolas, municípios ou outras agregações nessa base não correspondem aos dados oficiais do Inep – que seguem tendo como referência os microdados da matrícula inicial divulgado a cada ano.

Para fins de comparação, a Tabela 5 apresenta as taxas de transição dos alunos matriculados nas etapas seriadas de interesse (primeiro ano do ensino fundamental até a terceira série do ensino médio) entre 2014 e 2015 quando os tratamentos de deduplicação e de imputação não são realizados na base de dados do Censo Escolar. Comparando esses resultados com os apresentados na Tabela 4 - que refletem os dados oficiais de transição - verifica-se que os processos de deduplicação e imputação impactam diretamente na redução das taxas de evasão. A evasão do primeiro ano do ensino fundamental sem esses tratamentos é de 3,7% e, com esses tratamentos (dados oficiais), o valor passa para 1,8%. Como detalhado anteriormente, esses tratamentos corrigem falsas evasões tanto pelo reestabelecimento do correto código de identificação do aluno – processo de deduplicação – quanto pela correção de inconsistências na trajetória dos alunos – processo de imputação.

Tabela 5 – Taxas de transição - não oficiais - entre as séries de interesse (primeiro ano do ensino fundamental até a terceira série do ensino médio) e taxas de transição gerais (todas as redes de ensino) - base de dados sem tratamentos de duplicação ou imputação - Brasil - Censo Escolar, 2014-2015.

Taxas de transição (%) entre as séries de interesse 2014/2015 (percentuais em relação ao total de alunos matriculados na etapa em 2014 – valores arredondados na primeira casa decimal)													Taxas de transição gerais (%) 2014/2015			
Etapa de ensino ¹ 2014	Etapa de ensino ² 2015												Promoção ³	Repetência ⁴	Migração para EJA	Evasão
	1º EF	2º EF	3º EF	4º EF	5º EF	6º EF	7º EF	8º EF	9º EF	1ª EM	2ª EM	3ª EM				
1º EF	3.0	92.5	0.6	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	93.3	3.0	0.0	3.7
2º EF	0.2	4.0	91.8	0.5	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	92.3	4.2	0.0	3.5
3º EF	0.1	0.1	11.9	83.5	0.5	0.1	0.0	0.0	0.0	0.0	0.0	0.0	84.2	12.1	0.1	3.6
4º EF	0.0	0.0	0.1	8.3	86.9	0.6	0.0	0.0	0.0	0.0	0.0	0.0	87.5	8.5	0.3	3.7
5º EF	0.0	0.0	0.0	0.1	7.4	86.3	0.4	0.1	0.1	0.0	0.0	0.0	86.8	7.6	0.7	4.9
6º EF	0.0	0.0	0.0	0.0	0.1	14.2	76.5	0.7	0.2	0.0	0.0	0.0	77.4	14.3	2.2	6.1
7º EF	0.0	0.0	0.0	0.0	0.0	0.2	11.4	78.3	0.7	0.1	0.0	0.0	79.1	11.6	3.1	6.2
8º EF	0.0	0.0	0.0	0.0	0.0	0.0	0.1	8.7	80.7	0.9	0.0	0.0	81.6	8.9	3.1	6.4
9º EF	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.2	7.7	79.6	0.2	0.0	80.0	8.1	2.1	9.8
1ª EM	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	15.0	67.4	0.3	68.0	15.0	2.6	14.4
2ª EM	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	8.6	75.1	75.7	8.8	2.2	13.3
3ª EM	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	4.4	86.0	4.7	0.9	8.4

Nota: 1) Etapa de ensino registrada na situação final Censo Escolar da Educação Básica de 2014; 2) Etapa de ensino registrada na matrícula inicial do Censo Escolar da Educação Básica 2015; 3) O valor da taxa de promoção geral para cada uma das etapas de 2014 não bate com a soma dos percentuais onde a etapa de 2015 é superior à etapa de 2014 devido ao arredondamento dos valores na primeira casa decimal e às situações excepcionais tratadas na seção “Definições e metodologia de cálculo dos indicadores de fluxo escolar”; 4) O valor da taxa de repetência geral para cada uma das etapas de 2014 não bate com a soma dos percentuais onde a etapa de 2015 é igual ou inferior à etapa de 2014 devido ao arredondamento dos valores na primeira casa decimal e a situações excepcionais tratadas na seção “Definições e metodologia de cálculo dos indicadores de fluxo escolar”.

Estudos de acompanhamento realizados em bases de dados do Censo Escolar não submetidas aos tratamentos de deduplicação e imputação devem apresentar resultados similares aos da Tabela 5, ou seja, com percentual de evasão superior aos dados oficiais – que

utilizam a base de dados do fluxo escolar. Marino e Klein (2015)²², em estudo recente, avaliam as taxas de transição obtidas por meio dos microdados do Censo Escolar e ressaltam a existência desse excesso de evasão, apontando inclusive o provável motivo dessa inconsistência:

“[...] Isso está ocorrendo por causa do problema de identificação de alunos que mudam de escola, sem levar seus códigos INEP. Não sendo identificados, recebem novos códigos no ano seguinte e, portanto, “saem” do sistema regular de ensino com um código e retornam no ano seguinte com outro código.[...]”

Os tratamentos realizados na base de dados do fluxo escolar reduzem significativamente essas inconsistências e permitem a realização de estudos de acompanhamento mais precisos.

²² MARINO, Leandro Lins, KLEIN, Ruben. Cálculo das Taxas de Transição entre Séries. Reuniões da ABAVE, n. 8, p. 291-312, 2015. Disponível em: <http://abave.com.br/ojs/index.php/Reunioes_da_Abave/article/viewFile/340/146>. Acesso em: 6 maio 2017.