

O QUE PODE DAR ERRADO QUANDO SE USAM OS MICRODADOS DA RAIS PARA ANÁLISES DO EMPREGO PÚBLICO POR ENTE FEDERADO? UM INVENTÁRIO DOS PROBLEMAS CONHECIDOS E ALGUMAS SOLUÇÕES¹

Cláudio Hamilton Matos dos Santos²

Érica Lima Ambrósio³

Felipe dos Santos Martins⁴

Igor Ribeiro Mendonça⁵

Lucas Martins de Mattos⁶

1 INTRODUÇÃO

Coletada desde os anos 1970 e praticamente censitária, a Relação Anual de Informações Sociais (Rais) do Ministério do Trabalho é reconhecidamente uma fonte riquíssima de informações sobre o emprego formal no país. Os dados da Rais para as administrações públicas (daqui em diante referidos coletivamente como Rais-APU) são, todavia, notoriamente problemáticos (Brasil, 2015, p. 3).

O propósito desta nota é estender e refinar o esforço de Santos *et al.* (2018) no sentido de quantificar e mitigar os problemas da Rais-APU. A seção 2 exemplifica alguns desses problemas.

Santos *et al.* (2018) limitaram-se a analisar o caso dos estados e dos 2.082 municípios com Regimes Próprios de Previdência Social (RPPS) ativos em 2017 no

1. Este projeto beneficiou-se do apoio do Banco Interamericano de Desenvolvimento (BID), viabilizado pela Secretaria de Previdência do Ministério da Fazenda (SPrev) (Termo de Cooperação BR-T1357).

2. Técnico de planejamento e pesquisa na Diretoria de Estudos e Políticas Macroeconômicas (Dimac) do Ipea. *E-mail*: <claudio.santos@ipea.gov.br>.

3. Consultora do BID contratada no âmbito do Termo de Cooperação BR-T1357. *E-mail*: <erica.ambrosio@ipea.gov.br>.

4. Bolsista na Dimac do Ipea e consultor do BID no projeto contratado no âmbito do Termo de Cooperação BR-T1357. *E-mail*: <felipe.martins@ipea.gov.br>.

5. Bolsista na Dimac do Ipea e consultor do BID no projeto contratado no âmbito do Termo de Cooperação BR-T1357. *E-mail*: <igor.mendonca@ipea.gov.br>.

6. Bolsista na Dimac do Ipea e consultor do BID no projeto contratado no âmbito do Termo de Cooperação BR-T1357. *E-mail*: <lucas.mattos@ipea.gov.br>.

período 2006-2015. Esta nota aplica a metodologia proposta pelos referidos autores a todos os estados e municípios no período 2004-2016.

A nota propõe, ainda, quatro refinamentos à metodologia de Santos *et al.* (2018). O primeiro desses refinamentos consiste no cruzamento dos vínculos da Rais-APU com os dados do Sistema de Controle de Óbitos (Sisobi) da Dataprev. O segundo refinamento consiste no cruzamento dos dados da Rais-APU com os dados dos servidores inativos dos entes federados com RPPS divulgados nos respectivos portais de transparência. O terceiro refinamento consiste na análise da robustez intertemporal das declarações da “natureza jurídica” dos estabelecimentos que se declaram órgãos das administrações públicas estaduais ou municipais na Rais em algum ano do período analisado. O quarto refinamento consiste na análise da robustez intertemporal das informações sobre a natureza dos vínculos empregatícios na Rais-APU. Esses refinamentos são discutidos na seção 4, mas – por motivos práticos – aplicados apenas aos dados dos estados e dos municípios com RPPS.

O texto propõe, por fim, uma metodologia – pensada para ser tão geral quanto possível – de identificação de municípios com microdados de boa qualidade em anos específicos na Rais-APU. Essa metodologia é exemplificada na seção 5, novamente com os dados dos municípios com RPPS.

Entre os refinamentos propostos, o cruzamento dos dados da Rais-APU com os dados dos servidores inativos dos portais de transparência e a análise da robustez das informações sobre as naturezas dos vínculos empregatícios mostraram-se quantitativamente bem mais importantes do que os demais. Somados, os impactos de todos os refinamentos propostos sobre a precisão das bases de servidores estatutários dos estados e dos municípios com RPPS ativos entre 2011 e 2016 foram da ordem de 5 pontos percentuais (p.p.) – valor esse que, além de ligeiramente sobrestimado, é bem menor do que o impacto (da ordem de 9 p.p.) dos tratamentos básicos propostos por Santos *et al.* (2018). Por fim, a aplicação da metodologia sugerida nesta nota leva à conclusão de que a cobertura dos servidores estatutários ativos na Rais-APU de 2015 é satisfatória em 1.195 dos 2.082 municípios com RPPS ativos em agosto de 2017. Aperfeiçoamentos em andamento poderão aumentar consideravelmente esse número já no futuro próximo.

2 DEFINIÇÕES BÁSICAS E EXEMPLOS DE PROBLEMAS NA RAIS-APU

Antes de prosseguir, cumpre lembrar alguns conceitos básicos relativos ao emprego público na Rais. Terminada a revisão conceitual, exemplos concretos de estados e municípios com dados problemáticos na Rais são apresentados. Os exemplos têm como objetivo motivar os objetivos e o escopo desta nota.

2.1 Conceitos básicos e quatro clivagens importantes

Em princípio, todos os “órgãos da administração direta e indireta dos governos federal, estadual ou municipal, inclusive as fundações supervisionadas” e empresas públicas – entre um conjunto bem maior de instituições (Brasil, 2018, p. 6) – devem preencher o questionário da Rais anualmente e enviá-lo para o Ministério do Trabalho. O ministério, então, processa, homologa e trata essas informações, divulgando-as “da forma mais

democrática possível, com vista, dentre outros objetivos, de permitir o monitoramento constante e metucioso das tendências do emprego e renda (...)” (*ibidem*, p. 5).

As informações contidas no questionário da Rais podem ser divididas em dois grupos: *i*) informações sobre o estabelecimento encarregado de preencher o referido questionário; e *ii*) informações referentes aos empregados/servidores com vínculos empregatícios com o estabelecimento em questão.

A natureza jurídica e o endereço do estabelecimento são variáveis-chave na identificação dos órgãos de governo. É natural, por exemplo, identificar como órgãos das administrações públicas estaduais ou municipais aqueles estabelecimentos que declaram ter essas naturezas jurídicas na Rais. Mas o endereço é importante para que se possa determinar de qual estado ou município se está falando. Se, por exemplo, um estabelecimento declara-se como órgão da administração pública estadual e tem sede no estado de Minas Gerais, faz sentido supor que se trata de um órgão da administração pública de Minas Gerais. Analogamente, se um estabelecimento declara-se como órgão da administração pública municipal e está localizado no município de Abaíra, na Bahia, a hipótese usual é que se trata de um órgão da administração pública do município de Abaíra, Bahia.

A identificação de qual ente federado é responsável por qual empresa pública é mais difícil. Daí que esta nota *trata apenas do emprego nas administrações públicas estaduais e municipais*, desconsiderando completamente o emprego nas empresas públicas.

Uma segunda clivagem importante é a existente entre os municípios com ou sem RPPS ativos. Essa diferenciação é importante porque os municípios com RPPS ativos devem prestar contas desses regimes mensalmente para a Secretaria de Previdência do Ministério da Fazenda (SPrev). Os registros administrativos derivados dessas prestações de contas cumprem um papel importante no exercício da seção 5. Em agosto de 2017 havia 2.082 municípios com RPPS ativos e 3.486 municípios sem RPPS ativos. Todos os estados e o Distrito Federal têm RPPS ativos.

Importa, ademais, diferenciar a natureza do vínculo empregatício do servidor/empregado público. Servidores estatutários são estáveis e – caso o município/estado tenha RPPS ativo – beneficiários dos RPPS. Servidores temporários não são nem uma coisa nem outra. Daí que a contratação de um e de outro tipo de servidor impacta as finanças públicas municipais/estaduais de formas bastante distintas.

Por fim, uma quarta clivagem crucial para os propósitos desta nota é a de pessoas com apenas um ou com mais de um vínculo empregatício com determinado estado/município. É perfeitamente possível, por exemplo, que um procurador estadual seja também professor da universidade pública mantida pelo estado em questão. Ou que uma mesma professora tenha dois vínculos empregatícios de vinte horas semanais na mesma prefeitura – por exemplo, dando aulas de português na educação fundamental de manhã e aulas de educação infantil na parte da tarde. Isso importa por pelo menos dois motivos. Primeiro porque implica que o número de servidores será menor ou igual ao de vínculos. Segundo porque, como se verá a seguir, muito mais coisas podem dar errado nos registros de uma pessoa com vários vínculos empregatícios no mesmo ente federado do que nos registros de uma pessoa com apenas um vínculo por ente federado.

2.2 Exemplos de dados problemáticos

Dados imperfeitos são a regra e não a exceção no trabalho aplicado de economistas e cientistas sociais em geral. O fato de a Rais-APU conter imprecisões não é, portanto, particularmente surpreendente. Mas dois fatores tornam o estudo das imprecisões da Rais-APU interessante. O primeiro é que – como evidenciado na tabela 1 – os dados da Rais-APU para entes federados específicos podem ser muito imperfeitos em um dado ano qualquer. O segundo é que – como discutido nas seções 3 e 4 – essas imperfeições podem ser atenuadas consideravelmente com um conjunto relativamente pequeno de tratamentos estatísticos.

TABELA 1

Quantitativos de vínculos empregatícios nas administrações públicas de municípios selecionados (2012-2016)

Nome	UF	2012	2013	2014	2015	2016
Marapanim	PA	759	10	1.126	1	1.242
Zé Doca	MA	1.377	1.390	17	1.427	13
Natal	RN	16.821	18.265	17.420	1.079	19.180
Carmo do Paranaíba	MG	23	1.256	1.284	1.292	1.373
Santa Luzia	MG	3.631	6.508	3.841	3.466	3.694
Japeri	RJ	4.157	4.038	5.555	4.428	3.131
Monte Alegre dos Campos	RS	215	228	2	243	245
Campo Grande	MS	24.837	26.918	21.080	31.176	30.622

Fonte: Rais-APU.

2.3 Quais perguntas se quer responder?

É possível que os municípios listados na tabela 1 tenham dados menos precisos do que a média nacional. Mas o simples fato de existirem casos como os retratados na tabela 1 já é, por si só, suficiente para alertar o analista de que análises do emprego público por ente federado baseadas nos dados brutos da Rais-APU estão sujeitas a erros muito fortes.

Essa constatação faz com que análises do tipo sejam raras no Brasil. Mais comuns são análises do emprego público “nos estados” ou “nos municípios”, nas quais os dados da Rais-APU são comparados aos – e validados pelos – dados da Pesquisa Nacional por Amostra de Domicílios do Instituto Brasileiro de Geografia e Estatística (Pnad/IBGE).⁷

Mas não haveria maneiras de se verificar se e quando os microdados da Rais-APU de um determinado ente federado são de boa ou má qualidade? O fato de haver algumas maçãs podres dentro de uma saca quer dizer que todas as maçãs da saca devem ser jogadas no lixo? São essas perguntas para cujas respostas os resultados desta nota visam contribuir.

7. Estudos do tipo podem ser encontrados, por exemplo, no volume organizado por Cardoso Jr. (2011).

3 A ABORDAGEM DE SANTOS *ET AL.* (2018): RESULTADOS PARA OS PERÍODOS 2004-2016 E 2011-2016

Santos *et al.* (2018) sugerem que os microdados da Rais por ente federado podem ser utilizados sempre que: *i)* forem logicamente consistentes ao longo do tempo; e *ii)* forem consistentes, em um dado ano, com dados agregados e/ou com as médias das distribuições relevantes publicados em outras fontes de dados. Trataremos nesta seção do primeiro tipo de consistência – que, em linha com Santos *et al.* (2018), chamaremos de “interna”. Voltaremos ao segundo tipo de consistência – dita “externa” – na seção 5.

Os critérios de consistência interna de Santos *et al.* (2018) estão resumidos no quadro 1. A ideia é que algo está errado quando o servidor “aparece” na Rais sem ter sido admitido (erro do tipo 1) ou “desaparece” da Rais sem ter sido desligado (erro do tipo 2). A combinação desses dois erros é chamada de erro do tipo 3. Supõe-se que algo está errado também quando o servidor “desaparece” da Rais em um dado ano e “reaparece” na Rais anos depois (erro do tipo 4). Os demais erros listados no quadro 1 são combinações dos erros 1-4. O quadro 1 denota a presença de um vínculo na Rais pública pela letra “V”, e a admissão e o desligamento de um determinado servidor, respectivamente, pelas letras “A” e “D”.

QUADRO 1

Uma tipologia de problemas nos dados da Rais pública para os vínculos de servidores estatutários e militares

	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Erro 1			V	V	V	V	V	D		
Erro 2	A	V	V	V	V	V				
Erro 3				V	V	V	V			
Erro 4	A	V	V	V			V	V	V	D
Erro 5 = erro 4 + erro 1			V	V			V	V	D	
Erro 6 = erro 4 + erro 2		A	V			V	V			
Erro 7 = erro 4 + erro 3			V	V			V	V		

Fonte: Santos *et al.* (2018).

Analisando o universo dos servidores estatutários (apenas) nos estados e nos municípios com RPPS ativos no período 2006-2015 com dados identificados da Rais-APU, Santos *et al.* (2018) estimaram que 39% de todas as observações (de vínculos de servidores estatutários) da Rais-APU no período em questão apresentavam algum dos erros listados no quadro 1 – sendo que o percentual de erros, no caso dos vínculos das pessoas com mais de um vínculo, foi superior a 50% tanto nos estados quanto nos municípios.

Os erros são mais comuns no caso de pessoas com mais de um vínculo porque, nesse caso, a identificação de cada vínculo individual é mais difícil. Quando uma pessoa tem apenas um vínculo, o CPF dela basta para identificá-lo. Mas se uma pessoa tem, por exemplo, dois vínculos, então a identificação de cada vínculo requer não apenas o CPF da pessoa em questão, mas também as datas de admissão em cada um dos vínculos que ela tem.

É comum, entretanto, que as datas de admissão sejam mal reportadas na Rais – ou que a tabulação das referidas datas pelo Ministério do Trabalho seja imperfeita. Contudo, se as datas de admissão são mal reportadas em um dado ano, então não é mais possível identificar os vínculos individuais das pessoas com mais de um vínculo nesse ano – o que, por seu turno, leva ao “desaparecimento” (espúrio) dos vínculos existentes nos anos anteriores e ao “aparecimento” (igualmente espúrio) de vínculos novos (sem data de admissão) no ano em questão.

Santos *et al.* (2018) propõem, assim, cinco “tratamentos” para mitigar as inconsistências lógicas supracitadas.

Inicialmente, propõe-se que a natureza jurídica dos estabelecimentos (CNPJs) identificados como administrações públicas em todos os anos da amostra menos um seja mudada no ano excepcional. A motivação desse tratamento foi o fato de a Secretaria Municipal de São Paulo ter se declarado equivocadamente como “órgão da administração direta estadual” em 2010 – provocando o “desaparecimento” (espúrio) de perto de 60 mil vínculo da prefeitura de São Paulo e o “aparecimento” (igualmente espúrio) desses vínculos no governo do estado de São Paulo no ano em questão.

Em segundo lugar, foram identificados os vínculos sem data de admissão – especificamente nos anos de 2011 e 2013 – na versão da Rais-APU utilizada por Santos *et al.* (2018) –, e as referidas datas foram imputadas. Se, por exemplo, o CPF 0034 tinha dois vínculos em 2010 – digamos o vínculo 0034-2008 de baixa remuneração e o vínculo 0034-1997 de alta remuneração – e tinha dois vínculos sem data de admissão em 2011, as datas de admissão foram imputadas a partir do ordenamento das remunerações em 2010. Ou seja, o vínculo de maior remuneração em 2011 foi identificado como o 0034-1997 e o vínculo de menor remuneração foi identificado como o 0034-2008.

Em terceiro lugar, todos os “buracos” – isto é, vínculos que desaparecem da série e voltam a aparecer – foram “corrigidos”/preenchidos.

Em quarto lugar, os vínculos que apareceram na série sem nunca terem sido admitidos foram identificados e, no caso da data de admissão ser recente (até três anos antes do “aparecimento” do vínculo na Rais), foram “corrigidos”.

Finalmente, todos os CPFs dos vínculos que desapareceram na série sem nunca terem sido desligados foram identificados e “procurados” na Rais inteira (inclusive setor privado) após o desaparecimento. Se encontrados, foram “corrigidos”.

Com esses cinco tratamentos (efetivados na ordem descrita acima), Santos *et al.* (2018) reportam um aumento do grau de precisão dos registros dos vínculos de servidores estatutários nos 27 estados (incluindo o Distrito Federal) e nos municípios com RPPS ativos no período 2006-2015, de 61% para 77%, com aumentos proporcionalmente bem maiores no caso dos vínculos de pessoas com mais de um vínculo.

TABELA 2

Quantidades de vínculos empregatícios consistentes e por tipo de erro antes dos e após os tratamentos: todos os municípios, estados e vínculos (2004-2016)

	Acertos	Tipo 1	Tipo 2	Tipo 3	Total	Precisão pré-tratamento (%)	Precisão pós-tratamento (%)	Aumento (p.p.)
Dados ajustados – municípios (um vínculo)	8.316.211	242.300	1.452.720	89.350	10.100.581	71,0	82,3	11,4
Dados ajustados – municípios (mais de um vínculo)	6.047.909	87.832	1.034.365	60.079	7.230.185	68,0	83,6	15,6
Dados ajustados – estadual (um vínculo)	4.023.564	172.172	878.750	76.451	5.150.937	61,9	78,1	16,2
Dados ajustados – estadual (mais de um vínculo)	2.820.272	108.389	792.404	631.702	3.782.767	66,3	74,6	8,2
Total	21.207.956	610.693	4.158.239	287.582	26.264.470	67,9	80,7	12,9

Elaboração dos autores.

TABELA 3

Quantidades de vínculos empregatícios consistentes e por tipo de erro antes dos e após os tratamentos: apenas vínculos estatutários de todos os estados e municípios com RPPS ativos (2004-2016)

	Acertos	Tipo 1	Tipo 2	Tipo 3	Total	Precisão pré-tratamento (%)	Precisão pós-tratamento (%)	Aumento (p.p.)
Dados ajustados – municípios (um vínculo)	2.868.162	206.042	550.601	56.983	3.681.788	61,3	77,9	16,6
Dados ajustados – municípios (mais de um vínculo)	718.814	48.885	155.027	23.249	945.975	57,1	76,0	18,9
Dados ajustados – estadual (um vínculo)	2.618.676	335.990	537.896	61.536	3.554.098	54,0	73,7	19,7
Dados ajustados – estadual (mais de um vínculo)	749.670	122.768	176.638	40.363	1.089.439	44,1	68,8	24,7
Total	6.955.322	713.685	1.420.162	182.131	9.271.300	55,9	75,0	19,1

Elaboração dos autores.

As tabelas 2 e 3 reportam os resultados da metodologia de Santos *et al.* (2018) aos dados: *i*) de todos os vínculos da Rais-APU em todos os estados e municípios do país no período 2004-2016 (tabela 2); *ii*) apenas dos vínculos estatutários da Rais-APU em todos os estados e nos municípios com RPPS ativos no período 2004-2016 (tabela 3). Os dados estão divididos por estados e municípios e entre vínculos de pessoas com apenas um vínculo e de pessoas com mais de um vínculo.

Os dados brutos da tabela 3 são um pouco mais imprecisos do que os reportados por Santos *et al.* (2018), essencialmente por conta do aumento da amostra analisada. É interessante notar, ainda, que a precisão da Rais-APU para todos os vínculos, estados e municípios (tabela 2) é um pouco maior do que a precisão da Rais-APU apenas para os vínculos estatutários de estados e municípios com RPPS ativos. Nada, entretanto, que permita o uso desavisado da base.

Os critérios de consistência de Santos *et al.* (2018) são rígidos. Considere o caso de um município que reportou dados com grande precisão em todos os anos menos 2016. Esse município certamente terá um nível de precisão muito baixo por conta do grande número de erros de desaparecimento não passíveis de correção. Mas isso não significa, naturalmente, que não se possa achar anos na Rais-APU nos quais os microdados do município em questão são de excelente qualidade. Não faz sentido, ademais, “punir” um município que, por exemplo, reportava dados muito ruins até 2009 e passou a reportar dados excelentes a partir de 2010. De outro modo, impor a necessidade de consistência lógica da Rais-APU no período 2004-2016 é um exagero se o objetivo é identificar microdados de boa qualidade na Rais-APU de anos recentes. Com efeito, a tabela 4 mostra que a precisão dos dados brutos e tratados aumenta consideravelmente quando se diminui o tamanho do período coberto pelos testes de consistência.

TABELA 4

Quantidades de vínculos empregatícios consistentes e por tipo de erro antes dos e após os tratamentos: apenas vínculos estatutários de todos os estados e municípios com RPPS ativos (2011-2016)

	Acertos	Tipo 1	Tipo 2	Tipo 3	Total	Precisão pré-tratamento (%)	Precisão pós-tratamento (%)	Aumento (p.p.)
Dados ajustados – municípios (um vínculo)	2.547.926	147.040	247.354	26.616	2.968.936	79,9	85,8	5,9
Dados ajustados – municípios (mais de um vínculo)	490.042	20.917	77.376	8.121	596.456	67,7	82,2	14,5
Dados ajustados – estadual (um vínculo)	2.470.563	179.793	255.416	12.934	2.918.706	75,6	84,6	9,1
Dados ajustados – estadual (mais de um vínculo)	431.028	31.810	69.588	6.440	538.866	60,5	80,0	19,5
Total	5.939.559	379.560	649.734	54.111	7.022.964	75,5	84,6	9,1

Elaboração dos autores.

4 REFINANDO A ABORDAGEM DE SANTOS *ET AL.* (2018)

O principal resultado de Santos *et al.* (2018) é que os erros por conta de “desaparecimentos” de vínculos da Rais-APU sem desligamentos são os mais numerosos, seguidos pelo número de vínculos que “aparecem” sem serem admitidos.

Ora, servidores estatutários podem “desaparecer” da Rais-APU porque: *i)* morreram e não foram desligados; *ii)* aposentaram-se mas não foram desligados; *iii)* a instituição na qual trabalham mudou de natureza jurídica; e *iv)* a instituição na qual trabalham passou a registrar esse servidor como não estatutário. As ocorrências *(iii)* e *(iv)* ajudam a solucionar também os “aparecimentos” sem admissões porque implicam que vínculos “desapareçam” em alguns lugares e “apareçam” em outros.

Esta seção tem como objetivo refinar os tratamentos básicos propostos por Santos *et al.* (2018) de quatro maneiras. O primeiro desses refinamentos consiste no cruzamento dos vínculos da Rais-APU com os dados do Sisobi da Dataprev. O segundo refinamento consiste no cruzamento dos dados da Rais-APU com os dados dos servidores inativos dos entes federados com RPPS divulgados nos respectivos portais de transparência. O terceiro refinamento consiste na análise da robustez intertemporal das declarações da “natureza jurídica” dos estabelecimentos que se declaram órgãos das administrações públicas estaduais ou municipais na Rais em algum ano do período analisado. O quarto refinamento consiste na análise da robustez intertemporal das informações sobre a natureza dos vínculos empregatícios na Rais-APU.

4.1 O que se ganha cruzando os “erros” de Santos *et. al* (2018) com os dados do Sisobi?

Do ponto de vista do aumento da precisão, ganha-se pouco cruzando os dados da Rais-APU com os dados do Sisobi (ver tabela 5). Os resultados desse cruzamento são, todavia, muito instrutivos do ponto de vista analítico.

TABELA 5

Impacto do cruzamento com o Sisobi sobre a precisão da Rais-APU: apenas vínculos estatutários de todos os estados e municípios com RPPS ativos (2011-2016)

	Erro 1	Erro 2	Erro 3	Erros	Sem erros	Total
Erros originais	379.560	649.734	54.111	1.083.405	5.939.559	7.022.964
	5,40%	9,25%	0,77%	15,43%	84,57%	-
Após tratamentos propostos por meio do Sisobi	378.406	635.198	52.638	1.066.242	5.948.613	7.014.855
	5,39%	9,06%	0,75%	15,20%	84,80%	-

Elaboração dos autores.

O número de (CPFs de) servidores estatutários de todos os estados e dos municípios com RPPS ativos que apareceram na Rais-APU entre 2011 e 2016 foi pouco superior a 6 milhões, totalizando pouco mais de 7 milhões de vínculos. O número de registros do Sisobi entre 2001 e 2016 foi de pouco menos de 18 milhões, mas apenas 10,3 milhões (ou 57,7%) desses registros continham os CPFs dos falecidos. Após o pareamento das duas bases – feito inclusive a partir do nome no caso dos registros do Sisobi sem CPFs – foram encontradas 74,1 mil pessoas (ou 1,23% dos 6 milhões de CPFs na Rais-APU) presentes em ambas bases.

Ocorre que a maior parte (71,9%) dessas pessoas não apresentava problemas na Rais-APU. A pessoa pode muito bem vir a falecer após ter sido desligada corretamente da Rais. Adicionalmente, o simples fato de ter falecido não necessariamente corrige um erro de registro na Rais – uma pessoa pode, por exemplo, ter desaparecido da Rais em 2008 e falecido em 2015.

Dito isso, há, de fato, casos nos quais o falecimento da pessoa coincide com o desaparecimento dela da Rais. A aproximação adotada aqui foi assumir que esse é o caso sempre que o servidor desaparece da Rais até dois anos antes ou depois do falecimento. Essa correção diminui os erros do tipo 2 e transforma erros do tipo 3 em erros do tipo 1. A despeito de conceitualmente interessante, a correção não é particularmente importante do ponto de vista quantitativo.

Dois outros fenômenos conceitualmente interessantes – mas igualmente desimportantes do ponto de vista quantitativo – são: *i*) o fato de que muitas pessoas “aparecem” na Rais justamente no ano em que morrem, frequentemente (mas nem sempre) sendo corretamente desligadas nesse ano; e *ii*) o fato de que muitas pessoas/vínculos continuam sendo reportadas na Rais bem depois de terem falecido. No primeiro caso, a opção foi por supor que essas pessoas nunca existiram na Rais-APU – diminuindo, assim, a ocorrência de erros do tipo 1. No segundo caso – um erro não mencionado por Santos *et al.* (2018) –, a opção foi por desligar a pessoa quando do seu falecimento.

Após todas essas correções, a precisão total da base passou de 84,57% para 84,80% (tabela 5).

4.2 O que se ganha “corrigindo” inconsistências nas declarações da natureza jurídica dos estabelecimentos?

Novamente, do ponto de vista do aumento da precisão, a resposta à pergunta do título é muito pouco (ver tabela 6). Mas, uma vez mais, o exercício é instrutivo.

TABELA 6

Impacto da correção dos CNPJs sobre a precisão da Rais-APU: apenas vínculos estatutários de todos os estados e municípios com RPPS ativos (2011-2016)

	Erro 1	Erro 2	Erro 3	Erros
Erros originais	379.560	649.734	54.111	1.083.405
Erros após correção dos CNPJs	373.643	640.423	52.999	1.067.065
Diminuição dos erros (%)	1,56	1,43	2,06	1,51

Elaboração dos autores.

De 2004 até 2016 há 42 mil estabelecimentos diferentes que se declaram pertencentes ao setor público. Desse total, pouco menos de 21 mil (49%) mudam de esfera ao menos um ano e pouco menos de 10 mil (23%) mudam em mais de um ano. Os números diminuem pouco quando se diminui o período analisado. Entre 2011 e 2016, 42% dos estabelecimentos mudam de esfera pelo menos uma vez, sendo que 21% de todos os estabelecimentos mudam de esfera em mais de um ano. Embora esses números pareçam altos e sugiram que a correção do problema possa aumentar fortemente a precisão da Rais-APU, isso não ocorre na prática.

São dois os motivos para isso. O primeiro é que o caso dos estabelecimentos que mudam de natureza jurídica apenas uma vez já tinha sido tratado por Santos *et al.* (2018). O segundo é que os estabelecimentos que mudam de esfera mais de uma vez tendem a ser pequenos.

Esses últimos estabelecimentos foram divididos em dois grupos. No caso dos estabelecimentos que ficaram mais do que 50% dos anos em uma mesma esfera de governo, supôs-se que a esfera na qual o estabelecimento “permaneceu” mais tempo fosse a verdadeira. No caso dos demais estabelecimentos, a esfera de governo foi determinada pela análise dos nomes dos estabelecimentos por meio de expressões regulares. A conclusão, nos dois casos, foi que a maior parte dos estabelecimentos que mudam de natureza jurídica ao longo do tempo na Rais-APU não é, de fato, integrante das administrações públicas, mas pequenos estabelecimentos privados que preenchem errado o campo da natureza jurídica na Rais.

4.3 O que se ganha cruzando os “erros” de Santos *et al.* (2018) com os dados dos servidores inativos dos portais de transparência?

Ao contrário dos refinamentos anteriores, o cruzamento da Rais-APU com os dados dos servidores inativos dos portais de transparência aumenta significativamente a precisão da Rais-APU, notadamente no caso dos estados (tabela 7). Tal cruzamento requer um significativo trabalho de coleta dos dados dos portais de transparência, que é mais fácil de ser feito no caso dos estados do que no caso dos municípios.

TABELA 7

Impacto do cruzamento com os dados dos inativos dos portais de transparência sobre a precisão da Rais-APU: apenas vínculos estatutários de todos os estados e municípios com RPPS ativos (2011-2016)

Bases	Acertos	Tipo 1	Tipo 2	Tipo 3	Total	Precisão (%)	Precisão após tratamento básico (%)	Aumento (p.p.)
Municípios (um vínculo)	2.556.741	147.615	238.539	26.041	2.968.936	86,1	85,8	0,3
Municípios (mais de um vínculo)	492.106	21.178	75.312	7.860	596.456	82,5	82,2	0,3
Estados (um vínculo)	2.570.173	181.709	155.806	11.018	2.918.706	88,1	84,6	3,5
Estados (mais de um vínculo)	441.633	32.855	58.983	5.395	538.866	82,0	80,0	2,0
Total	6.060.653	383.357	528.640	50.314	7.022.964	86,3	84,6	1,7

Elaboração dos autores.

No caso dos estados, foram coletados perto de 1 milhão e 100 mil nomes listados como inativos nos portais de 22 estados relativos a dezembro de 2017. Perto de 352 mil dessas pessoas foram encontradas na Rais-APU entre 2011 e 2016. Essas pessoas estão associadas a 375 mil vínculos.

No caso dos municípios, os dados de servidores inativos nem sempre são disponibilizados nos portais – quando há portais. Além disso, o número de municípios com dados não

manipuláveis não é pequeno. Ainda assim, foram encontrados perto de 188 mil nomes de servidores inativos nos portais de 615 municípios relativos a dezembro de 2017. Dessas 188 mil pessoas, 57 mil associadas a 62 mil vínculos foram encontradas na Rais-APU entre 2011 e 2016.

Não é impossível que um servidor inativo volte à atividade – por exemplo, por conta da aceitação de um cargo comissionado. A interpretação dada aqui, entretanto, é que os inativos encontrados na Rais-APU entre 2011 e 2016 são pessoas que estiveram em atividade nesses anos antes de entrarem para a inatividade.

Somando estados e municípios, foram encontrados 407 mil inativos (em 2017) associados a 438 mil vínculos de estatutários ativos em pelo menos um ano entre 2011 e 2016. A maior parte desses 438 mil vínculos tinha registros livres de problemas. Mas perto de 121 mil desses vínculos apresentaram erros do tipo 2 – isto é, contavam como tendo desaparecido da Rais-APU em um dos anos em questão. Outros 3 mil vínculos tinham registros com erros do tipo 3 – isto é, constavam como tendo aparecido na Rais-APU sem serem admitidos e como tendo desaparecido da Rais-APU. A hipótese adotada nesses casos é que o último ano da presença dessas pessoas na Rais-APU foi o ano de aposentadoria. Com isso, foi possível corrigir os erros supracitados e aumentar a precisão da base em 1,7 p.p., com o grosso do aumento ocorrendo nos estados (tabela 7).

4.4 O que se ganha “corrigindo” inconsistências nas declarações da natureza dos vínculos dos servidores?

A correção de inconsistências nas declarações da natureza dos vínculos foi, de longe, o refinamento mais efetivo em termos quantitativos (ver tabela 8).

TABELA 8

Impacto do cruzamento com os dados dos inativos dos portais de transparência sobre a precisão da Rais-APU: apenas vínculos estatutários de todos os estados e municípios com RPPS ativos (2011-2016)

Bases	Acertos	Tipo 1	Tipo 2	Tipo 3	Total	Precisão (%)	Precisão após tratamento básico (%)	Aumento (p.p.)
Municípios (um vínculo)	2.652.531	76.394	220.472	19.539	2.968.936	89,3	85,8	3,5
Municípios (mais de um vínculo)	497.908	16.356	75.675	6.517	596.456	83,5	82,2	1,3
Estados (um vínculo)	2.570.863	89.804	248.445	9.594	2.918.706	88,1	84,6	3,5
Estados (mais de um vínculo)	440.144	23.905	69.235	5.582	538.866	81,68	80,0	1,7
Total	6.161.446	206.459	613.827	41.232	7.022.964	87,7	84,6	3,1

Elaboração dos autores.

Tal como nos demais casos, hipóteses tiveram que ser realizadas sobre a natureza do vínculo empregatício para que correções pudessem ser feitas. Inicialmente foram

identificados todos os vínculos de servidores estatutários cujo tipo de vínculo mudou na Rais-APU dos estados e dos 2.082 municípios com RPPS ativos entre 2011 e 2016. Em segundo lugar, supôs-se que o tipo de vínculo com mais anos era o correto.

Um resultado particularmente importante da correção dos vínculos é que ela reduz a quantidade de erros do tipo 1 desproporcionalmente em relação aos demais erros – vale dizer, muitas pessoas aparecem na Rais-APU como estatutárias sem terem sido admitidas apenas porque foram admitidas em um vínculo diferente do de servidor estatutário. Isso é interessante porque os demais tratamentos impactam desproporcionalmente os erros do tipo 2, afetando relativamente menos os erros do tipo 1.

4.5 Retornos decrescentes em novos refinamentos?

A soma dos resultados dos tratamentos discutidos acima, tomados individualmente, sobrestima ligeiramente o impacto combinado desses tratamentos. Isso ocorre porque cada tratamento impacta os demais. Ainda assim, a referida soma não ultrapassa os 5 p.p. no caso dos vínculos estatutários da Rais-APU de estados e dos 2.082 municípios com RPPS ativos entre 2011 e 2016. Vale dizer, caso o viés de sobrestimação não existisse, a precisão nos referidos vínculos passaria de 84,6%, após os tratamentos básicos (tabela 4), para 89,8%, após os refinamentos discutidos nesta seção.

Haveria muito ainda o que melhorar nos refinamentos discutidos nesta seção. O potencial dos dados dos portais de transparências municipais, em particular, ainda está quase inexplorado – em grande medida por conta das dificuldades associadas à obtenção e à estruturação dos dados em questão. Muito há a aprender, ainda, com o pareamento dos dados dos servidores ativos de estados e municípios dos portais com os dados da Rais-APU. Por fim, os tratamentos discutidos acima – e os demais que vierem a ser desenvolvidos – devem ser devidamente encadeados para evitar a dupla contagem supracitada.

Mas cabe o registro de que, mesmo no melhor cenário, é difícil imaginar a precisão dos registros discutidos acima passando muito dos 90%. O impacto, mesmo sobrestimado, dos refinamentos discutidos nesta seção (5,2%) é pouco superior à metade do impacto (de 9,1%) dos refinamentos básicos de Santos *et al.* (2018). É justo esperar que o impacto dos novos refinamentos não ultrapasse alguns poucos – ainda que valiosos – pontos percentuais.

5 UMA METODOLOGIA GERAL DE UTILIZAÇÃO DOS MICRODADOS DA RAIS PARA ANÁLISE DO EMPREGO PÚBLICO POR ENTE FEDERADO?

A discussão das seções acima procurou explicitar os méritos e as limitações do esforço de tratamento da Rais-APU.

Registre-se, entretanto, que a verificação da consistência lógica dos microdados da Rais-APU dos vários entes federados é apenas uma das partes da metodologia ora proposta. A rigor, a referida consistência – embora significativa – não é condição necessária ou suficiente para a obtenção de microdados de boa qualidade dos servidores ativos de um dado ente federado em um dado ano qualquer. Não é condição suficiente

porque é possível ter dados perfeitamente consistentes apenas de uma subamostra dos dados relevantes. A consistência lógica dos dados não implica boa cobertura, em resumo. Não é condição necessária porque é perfeitamente possível que um ente federado tenha dados muito bons em um ano X qualquer e tenha dados muito ruins em anos anteriores e posteriores a X . Inconsistências intertemporais dos dados não eliminam a possibilidade de que “anos bons” existam, em resumo.

Daí que a segunda parte da metodologia ora proposta é a verificação da consistência “externa” dos dados da Rais-APU, vale dizer a comparação dos dados da Rais-APU com os dados, em geral, agregados – ou de médias das distribuições relevantes – divulgados em pesquisas (como a Pesquisa de Informações Básicas Municipais – Munic ou a Pesquisa de Informações Básicas Estaduais – Estadiv do IBGE) e/ou registros administrativos (como os Demonstrativos de Resultados de Avaliações Atuariais – DRAAs) e/ou censos (como o Censo Escolar) que têm como objetivo específico a mensuração de variáveis da Rais-APU. Naturalmente, as bases de comparação externa variam, a depender do objetivo do usuário da Rais-APU, e podem, inclusive, não existir.

Os autores desta nota estão envolvidos em um esforço de utilização da Rais-APU para a identificação dos microdados dos servidores estatutários ativos dos municípios com RPPS ativos. Em estudos previdenciários faz grande diferença se há cem ou mil servidores em um dado ente federado. Faz grande diferença também se um dado servidor é homem ou mulher, se essa pessoa tem 20 ou 50 anos, se foi contratada há um ou dez anos, e se é professor ou professora. Daí que, no caso dessa pesquisa, em particular, é importante identificar os entes federados com microdados: *i*) que, somados, cheguem próximos ao quantitativo efetivo de vínculos do ente; *ii*) com o gênero, a idade e o tempo de serviço de cada servidor bem repostados; e, por fim, *iii*) com os servidores estatutários professores da educação básica devidamente identificados.

Idealmente, a identificação dos entes com microdados de boa qualidade, no caso da pesquisa acima, deveria ser feita em várias etapas. Primeiramente, testar-se-ia a consistência lógica intertemporal dos dados de cada ente no período 2011-2016 nos moldes de Santos *et al.* (2018). A seguir, submeter-se-ia os referidos microdados aos tratamentos discutidos nas seções 3 e 4 acima. O próximo passo seria construir uma versão tratada dos microdados em questão. Esses microdados tratados seriam, então, submetidos a testes de consistência externa nas variáveis mencionadas no parágrafo anterior. Os municípios/estados com dados tratados tidos como internamente consistentes e que, ademais, passam nos testes de consistência externa supracitados seriam, então, selecionados.

A pesquisa em questão ainda está em andamento. O objetivo no restante desta seção é apresentar – a título de exemplo e com dados ainda subótimos – os resultados de testes de consistência externa feitos para verificar a cobertura dos microdados da Rais-APU no caso dos servidores estatutários ativos dos municípios com RPPS ativos.

As bases de consistência externa, nesse caso, foram a Munic/IBGE e os DRAAs do Sistema de Informações dos Regimes Públicos de Previdência Social (Cadprev) do Ministério da Fazenda. Em ambos os casos a variável de interesse é o número de servidores

estatutários ativos. A Munic/IBGE é uma pesquisa feita por meio de questionário enviado aos municípios. Os DRAAs são registros administrativos produzidos com o objetivo de prestação de contas dos municípios para a SPREV. Ambas as bases também têm problemas, de modo que não se pode identificar com certeza qual base tem o dado correto em caso de discrepância entre elas e entre elas e a Rais-APU. A tabela 9 mostra alguns desses casos.

TABELA 9

Discrepâncias entre os quantitativos de servidores estatutários ativos na Rais-APU, no DRAA e na Munic/IBGE em municípios selecionados (2015)

Município	Rais-APU	DRAA	Munic/IBGE	Consistência interna Rais-APU (%)
Itaúna-MG	1.874	1.469	1.301	86,20
Catalão-GO	2.449	1.559	448	85,70

Fontes: Rais-APU, DRAA e Munic/IBGE.
Elaboração dos autores.

Em linha com a metodologia exposta acima, a operacionalização da verificação da cobertura dos dados dos servidores estatutários ativos da Rais-APU dos municípios com RPPS ativos envolveu vários passos. Inicialmente foram selecionados os municípios com dados brutos com mais de 70% de consistência interna entre 2004 e 2016. Ainda não foi possível construir uma base de dados tratados para cada ente federado no período em questão para o período 2011-2016, o que seria o ideal. Em segundo lugar, os municípios com mais de 70% de consistência interna e com diferenças entre os quantitativos da Rais-APU e do DRAA menores do que 20% no ano de 2015 foram tidos como consistentes interna e externamente. Os municípios selecionados no primeiro passo (isto é, com mais de 70% de consistência interna) e barrados no segundo (isto é, com diferenças entre os quantitativos da Rais-APU e do DRAA maiores que 20%) foram, então, novamente submetidos ao teste de consistência externa, agora usando a Munic/IBGE como base de comparação. Os resultados desses procedimentos são resumidos na tabela 10. Ao fim e ao cabo, 1.134 municípios com RPPS ativos tiveram microdados de servidores estatutários ativos tidos como consistentes na Rais-APU de 2015. As tabelas 11 e 12 apresentam exemplos concretos de municípios com dados com cobertura tida como satisfatória e dados com cobertura tida como insatisfatória.

TABELA 10

Resultados dos testes de consistências externa e interna para verificação da cobertura dos cadastros de servidores estatutários ativos da Rais-APU dos municípios com RPPS ativos

	Inconsistentes	Consistências interna e externa	Consistência interna	Consistência externa	Consistência externa – Rais DRAA	Consistência externa – Rais Munic
Número de municípios	948	1.134	1.331	1.605	1.219	386
% do total	45,50	54,50	63,90	77,10	58,50	18,50

Elaboração dos autores.

TABELA 11

Exemplos de municípios cuja cobertura dos servidores estatutários ativos na Rais-APU foi tida como satisfatória após os dois primeiros testes de consistência externa

Município	Rais-APU	DRAA	Munic	Consistência interna (%)	Rais-APU e DRAA	Rais-APU e Munic
Ariquemes-RO	2.219	2.188	2.136	91,6	<i>True</i>	-
Juiz de Fora-MG	7.048	6.772	5.941	87,6	<i>True</i>	
Abaetetuba-PA	3.720	0	3.632	70,1	<i>False</i>	<i>True</i>
Santo Antônio do Descoberto-GO	1.742	0	1.636	73,6	<i>False</i>	<i>True</i>
Arceburgo-MG	289	281	285	92,6	<i>True</i>	-

Fontes: Rais, DRAA e Munic/IBGE.
Elaboração dos autores.

TABELA 12

Exemplos de municípios cuja cobertura dos servidores estatutários ativos na Rais-APU foi tida como insatisfatória após os dois primeiros testes de consistência externa

Município	Rais-APU	DRAA	Munic	Consistência interna (%)	Rais-APU e DRAA	Rais-APU e Munic
Poço Fundo-MG	301	253	260	31,0	<i>True</i>	-
Três Passos-RS	44	484	474	32,0	<i>False</i>	<i>False</i>
Barreirinhas-MA	1.406	0	1.374	5,0	<i>False</i>	<i>True</i>
Paranaguá-PR	4.178	4.041	3.956	42,9	<i>True</i>	-
Duque de Caxias-RJ	10.588	10.264	10.496	26,7	<i>True</i>	-

Fontes: Rais, DRAA e Munic/IBGE.
Elaboração dos autores.

Uma vez mais, é possível argumentar que os testes de consistência impostos são demasiadamente duros com a Rais-APU. Como visto acima, não há, em particular, por que supor que as bases utilizadas na comparação externa sejam modelos de perfeição. Daí que os testes em questão podem levar o analista a jogar fora dados consistentes da Rais-APU apenas e tão somente por conta de problemas na Munic/IBGE ou nos DRAAs (tabela 9).

A fim de dar aos 197 municípios com dados internamente consistentes da Rais-APU que não passaram nos dois primeiros testes de consistência externa uma última instância de apelação, optou-se por utilizar os municípios tidos como consistentes em “grupos de treinamento” e “teste” (James *et al.*, 2015, p. 176) na estimação de uma suíte de modelos de previsão da quantidade de servidores estatutários ativos municipais. A ideia é usar as previsões dos modelos como juízes das discrepâncias entre os dados da Rais-APU e os dados das bases utilizadas nos testes de consistência externa.

Os dados dos municípios consistentes foram, então, segregados em dois grupos: treinamento (80% ou 907 municípios) e teste (20% ou 227 municípios). As variáveis utilizadas nos vários modelos foram: *i*) tamanho da população do município; *ii*) receita corrente líquida do município; *iii*) receita própria do município; *iv*) receita de transferências recebidas pelo município; *v*) despesa total do município; *vi*) despesa do município com pessoal e encargos sociais; e *vii*) despesas do município com os vencimentos do pessoal civil do município. A suíte de modelos consistiu em: a) modelo de regressão

linear (James *et al.*, 2015, cap. 3); b) modelo de regressão de Poisson (StataCorp, 2013, p. 377); c) modelo de árvore de regressão (James *et al.*, 2015, cap. 8); d) modelo de *boosting* (*ibidem*); e) modelo de *random forest* (*ibidem*). As previsões do modelo de *random forest* foram a de menor erro quadrado médio entre todos da suíte.

As previsões do modelo de *random forest* foram então utilizadas para retestar a consistência externa dos dados dos referidos 197 municípios com dados internamente consistentes e que foram reprovados nos testes de consistência anteriores. Esse último teste salva 61 dos referidos 197 municípios. As tabelas 13 e 14 dão exemplos de municípios salvos e não salvos.

TABELA 13

Exemplos de municípios cuja cobertura dos servidores estatutários ativos na Rais-APU foi tida como satisfatória após o terceiro teste de consistência externa

Município	Rais-APU	DRAA	Munic	<i>Random forest</i>	Consistência interna (%)	Rais-APU e DRAA	Rais-APU e Munic
Catalão-GO	2.449	1.559	448	2.803	85,74	<i>False</i>	<i>False</i>
Parei Novo-RS	190	153	157	191	94,04	<i>False</i>	<i>False</i>
Itaúna-MG	1.874	1.469	1.301	1.992	86,15	<i>False</i>	<i>False</i>
Marianópolis do Tocantins-TO	213	268	295	210	80,14	<i>False</i>	<i>False</i>
Jequié-BA	3.004	0	2.223	2.986	78,41	<i>False</i>	<i>False</i>

Fontes: Rais, DRAA e Munic/IBGE.
Elaboração dos autores.

TABELA 14

Exemplos de municípios cuja cobertura dos servidores estatutários ativos na Rais-APU foi tida como insatisfatória após o terceiro teste de consistência externa

Município	Rais-APU	DRAA	Munic	<i>Random forest</i>	Consistência interna (%)	Rais-APU e DRAA	Rais-APU e Munic
Juazeiro do Norte-CE	11.328	4.796	4.747	3.335	73,9	<i>False</i>	<i>False</i>
Bom Jesus-PB	142	0	110	186	87,5	<i>False</i>	<i>False</i>
Santa Maria de Jetibá-ES	975	778	796	1.287	92,4	<i>False</i>	<i>False</i>
Querência-MT	887	418	398	708	80,1	<i>False</i>	<i>False</i>

Fontes: Rais, DRAA e Munic/IBGE.
Elaboração dos autores.

6 CONSIDERAÇÕES FINAIS

Esta nota aplicou a metodologia de Santos *et al.* (2018) para os dados brutos da Rais-APU de todos os estados e municípios brasileiros no período 2004-2016. Os resultados obtidos (tabela 2) sugerem que o grau de precisão desses dados é pouco inferior a 68% e pode ser aumentado – com os tratamentos propostos pelos autores em questão – para cerca de 81%. A aplicação da mesma metodologia apenas para os dados dos servidores estatutários dos estados e municípios com RPPS ativos no mesmo período levou a resultados inferiores – isto é, precisão inicial de 56% aumentada com os tratamentos para 76% (tabela 3).

Foram propostos, ademais, quatro tratamentos adicionais aos sugeridos por Santos *et al.* (2018), quais sejam: *i*) o cruzamento dos vínculos da Rais-APU com os dados do Sisobi da Dataprev; *ii*) o cruzamento dos dados da Rais-APU com os dados dos servidores inativos dos entes federados com RPPS divulgados nos respectivos portais de transparência; *iii*) a análise da robustez intertemporal das declarações da “natureza jurídica” dos estabelecimentos que se declaram órgãos das administrações públicas estaduais ou municipais na Rais em algum ano do período analisado; e *iv*) a análise da robustez intertemporal das informações sobre a natureza dos vínculos empregatícios na Rais-APU. Os tratamentos (*i*) e (*iii*) tiveram impactos bem menores do que o tratamento (*ii*) e principalmente o tratamento (*iv*). No caso dos vínculos dos servidores estatutários dos estados e dos municípios com RPPS ativos, o efeito dos quatro tratamentos adicionais é bem inferior, entretanto, ao efeito dos tratamentos originais propostos por Santos *et al.* (2018), sugerindo retornos decrescentes em novos tratamentos.

Por fim, uma metodologia de validação dos microdados da Rais-APU foi proposta e exemplificada, com dados ainda subótimos, para a análise da cobertura dos servidores estatutários ativos da Rais-APU dos municípios com RPPS ativos. A conclusão preliminar é que essa cobertura foi satisfatória em 2015 para 1.195 municípios, de um total de 2.082 com RPPS ativos. Estudos em andamento deverão elevar esse número para a casa dos 1.500-1.600 no já próximo futuro.

REFERÊNCIAS

BRASIL. Ministério do Trabalho e Emprego. **Nota Técnica MTE (09/09/2015)**: base de dados Rais/2014. Brasília: MTE, 2015.

_____. **Manual de Orientação Rais ano-base 2017**. Brasília: MTE, 2018.

CARDOSO JR., J. C. **Burocracia e ocupação no setor público brasileiro**. Brasília: Ipea, 2011. (Série Diálogos para o Desenvolvimento, v. 5).

JAMES, G. *et al.* **An introduction to statistical learning with applications in R**. 6th ed. New York: Springer, 2015.

SANTOS C. H. *et al.* O quão compatíveis são os dados da Rais e os dos demonstrativos previdenciários brasileiros: estimativas preliminares para o período 2006-2015. *In*: BRASIL. Ministério da Fazenda. **Regimes previdenciários e situação atuarial**: estudos. Brasília: MF, 2018. (Coleção Previdência Social, v. 36).

STATA CORP. **Stata Users Guide Release 13**. College Station: Stata Press, 2013.