

# O Impacto da Remuneração Variável no Desempenho do Servidor Público

## *The Impact of Variable Pay on The Performance of Public Servants*

**Silvio Campos Guimarães**

Doutorado em Administração Pública pela Fundação Getúlio Vargas. Professor da Fundação Getúlio Vargas.  
silvio.guimaraes@fgv.br  
<http://lattes.cnpq.br/0934254500662680>

**Nelson Marconi**

Doutorado em Economia de Empresas pela Fundação Getúlio Vargas. Professor da Fundação Getúlio Vargas.  
nelson.marconi@fgv.br  
<http://lattes.cnpq.br/3941333920095061>

**Resumo:** O objetivo do presente estudo é analisar o efeito da remuneração variável sobre o desempenho do servidor público. Notou-se que a literatura é divergente quanto à associação entre os incentivos e a melhoria do desempenho no setor público. Para contribuir com as pesquisas e com o esclarecimento da divergência observada, desenvolveu-se um modelo empírico visando mensurar os efeitos dos programas no desempenho do servidor. Na sequência, foram realizados testes em um estudo de caso sobre o programa de remuneração variável aplicado aos servidores das escolas públicas do Estado de São Paulo objetivando identificar seus impactos no desempenho dos diretores de escola e dos professores. Considerando que estes servidores fazem parte do sistema de ensino, também foi testado se os efeitos do programa sobre o desempenho dos diretores e professores impactam o desempenho dos seus alunos.

**Palavras-chave:** Bonificação; Produtividade; Educação; Desempenho.

**Abstract:** The aim of this study is to analyze the effects of performance-related pay (PRP) on the public servants' activity. It was observed that the literature is divergent when it comes to the association between benefits and performance improvement in the public sector. In order to endow our research and to clarify this mentioned divergence, we developed an empirical model in pursuance of measuring the effects of PRP programs on servers' activity. Afterwards, tests were taken in a case study about a performance-related pay (PRP) program that was tested on public servants in public schools in the state of São Paulo. The aim was to identify the program's impact on the performance of school principals and teachers. Since these servers are a part of the educational system, it was also tested whether the effects of this program have any impact on the performance of these teachers' and principals' students.

**Key-Words:** Performance-related. Systems subsidies. Productivity. Education.

Texto completo em português: <http://www.apgs.ufv.br>  
Full text in Portuguese: <http://www.apgs.ufv.br>

### 1. Introdução

O presente estudo objetiva analisar se programas de remuneração variável exercem algum impacto no desempenho dos servidores públicos. Como objetivo secundário, com base nos efeitos observados no desempenho dos servidores, procurou-se identificar se há reflexos no desempenho dos alunos, que é o escopo final da educação.

A relevância do tema é significativa, visto que em 2014, segundo dados do Ministério do Trabalho e Emprego, o Brasil possuía 49.571.510 de empregos formais. Deste total, 9.355.833 eram trabalhadores no setor público nos governos municipais, estaduais e federal, incluso a administração direta e indireta. Aproximadamente 20% dos empregos formais do país estão na administração pública, sendo que muitos deles podem ser objetos de um modelo de remuneração variável, visando melhorar o desempenho do servidor.

Ainda tratando da relevância do tema, vale lembrar que os serviços oferecidos pelo Estado ao cidadão têm sido crescentes, desde que se confirmou que o mesmo pode ser a melhor alternativa para atenuar diferenças sociais, atuar nas falhas de mercado, agir na defesa dos interesses coletivos, entre outras funções importantes. Portanto, aparelhar o Estado com instrumentos de

gestão adequados de seus recursos é importante para o atingimento dos propósitos sociais, sendo que o principal recurso utilizado nesta relação Estado-Cidadão é o servidor público, aqui abrangendo qualquer outra denominação jurídica ou gerencial que defina alguém que seja prestador de serviço com vínculo empregatício com o Estado.

Por outro lado, a administração pública tem encontrado dificuldade em adotar estruturas de incentivo que favoreçam de maneira clara a melhoria do desempenho do servidor público. Como o servidor é um elemento crucial no processo de prestação de serviços, qualquer tentativa de progresso nessa área deve levar em consideração as políticas de recursos humanos envolvidas no processo.

Nas experiências desenvolvidas em diversos países, e no Brasil, ainda não se obteve respostas convincentes sobre como constatar e mensurar o impacto da remuneração variável no desempenho do servidor. Já na iniciativa privada os indicadores atrelados ao resultado econômico são utilizados como referencial para premiação, entretanto, o mesmo não é o objeto principal do interesse do Estado, mas sim o bem-estar da população. Diante disso, a avaliação do desempenho dos servidores envolvidos na produção de bens e serviços, principalmente os ligados à



administração direta, há de ser feita por indicadores não econômicos.

Além disso, pode ser notada a carência de estudos em países em desenvolvimento sobre os efeitos da remuneração variável, seja no setor público ou na área privada. Nesse sentido, pode-se citar o estudo de Hasnain, Manning e Pierskalla (2012), que analisou 153 estudos sobre o tema, sendo que apenas 26 foram realizados em países em desenvolvimento. Todos os demais foram em países membros da Organização para a Cooperação e Desenvolvimento Econômico - OCDE. Para tanto, a revisão da literatura aqui utilizada encontrou poucos estudos na América Latina, sendo que no Brasil são escassos os estudos sobre o tema. Nota-se uma janela de oportunidade para se desenvolver estudos nesta área, considerando a realidade da cultura local, e que esta pesquisa intenciona contribuir na procura por meios que possam aferir adequadamente os efeitos da remuneração variável.

Assim, dado a importância do tema, tanto para governos quanto para servidores, a carência de estudos sobre o tema, principalmente na América Latina, e em específico no Brasil, além das divergências de resultados encontrados nos efeitos do PRV, o presente trabalho pretende contribuir na identificação dos impactos dos programas de remuneração variável no desempenho do servidor público.

## 2. Fundamentos Teóricos

A remuneração variável é considerada um instrumento de gestão de pessoas que visa reduzir a assimetria de informações entre o trabalhador e o empregador. Para entender melhor essa assimetria e como lidar com ela, é preciso conhecer os fundamentos que lhe dão origem. A Teoria de Agência é considerada a principal abordagem neste campo de estudo. Seus conceitos podem ser observados no artigo seminal de Jensen e Meckling (1976). Os autores definem o relacionamento de agência como "um contrato no qual uma ou mais pessoas, o(s) principal(is), engajam outra pessoa – o agente – para desempenhar alguma tarefa em seu favor, envolvendo a delegação de autoridade para a tomada de decisão pelo agente". Dessa forma, se ambas as partes agem tendo em vista a maximização de suas utilidades pessoais, a Teoria de Agência afirma que existe uma boa razão para acreditar que o agente não agirá sempre no melhor interesse do principal.

Os problemas de agência, ou conflitos de agência, que identificam as dificuldades nas relações agente/principal, são comuns. Tais problemas podem surgir quando duas condições são notadas: (1) os objetivos do principal e do agente são diferentes; e (2) as ações realizadas pelo agente, ou as informações de posse desse agente, são difíceis de serem observadas. O motivo para a existência desta segunda condição é que, se as ações e informações forem facilmente observáveis, então seria fácil para o governante/gestor estruturar a relação de agência, de modo a alinhar os interesses do servidor aos seus próprios. Isto é, as diferenças de objetivos podem ser superadas se as ações ou informações do servidor puderem ser facilmente observadas e, sendo observadas, podem ser controladas.

Mais importante do que conhecer os problemas e conflitos de agência, é saber como lidar com eles, de forma a solucioná-los, ou amenizar seus efeitos. Para Besanko, Dranove, Shanley e Schaefer (2012), utilizar incentivos de pagamento por desempenho, pode ajudar a alinhar os interesses do servidor com os do governante, mitigando, assim, os conflitos de agência. Para tanto, o governante oferece atrelar o pagamento do servidor ao resultado que o governante espera receber em decorrência da ação do servidor, ou seja, o servidor ganha mais quando uma ação de governo tem bom desempenho, e menos quando uma ação de governo tem um desempenho ruim, ficando, portanto, disposto a realizar ações que beneficiem o governo. Entretanto, assim como nos dois grupos anteriores, há complicações no ambiente real das organizações pelas seguintes razões:

Primeiramente, o servidor tem controle da sua dedicação para produzir melhores resultados, porém fatores exógenos ao seu ambiente de trabalho podem interferir no seu desempenho, tal como carência ou baixa qualidade de materiais e insumos com os quais ele lida. Até mesmo crises políticas e econômicas podem lhe prejudicar sem ter algo que possa individualmente fazer.

Além disso, fornecer incentivo para um servidor que exerce tarefas múltiplas parece ser uma dificuldade. Portanto, é de extrema importância equilibrar os incentivos entre as diferentes tarefas, no entanto, muitas dessas tarefas são difíceis de serem avaliadas. Por consequência, o servidor pode receber incentivo somente para algumas e, neste caso, ele irá canalizar seus esforços às tarefas mensuráveis e que lhe proporcionem retorno financeiro.

Ademais, para uma tarefa como a de elaborar um programa educacional complexo, não é possível atrelar o resultado esperado pelo governante às ações de um servidor individual. Isso complica o problema de fornecer incentivos baseados no desempenho, porque não é possível dar a cada servidor a propriedade sobre o produto extra por ele criado.

Saber como lidar com estas dificuldades no ambiente real das organizações é um desafio permanente para gestores e pesquisadores.

## 3. Modelo Empírico

Para averiguar o efeito do programa de remuneração variável, adotado na Secretaria de Educação do Estado de São Paulo para os diretores e professores da rede estadual, a partir de 2008, deveríamos comparar o desempenho em duas situações, com e sem a presença do programa. Porém, como o programa é aplicado a todos os servidores da educação deste Estado indistintamente, não é possível observar as duas situações ao mesmo tempo.

Angrist e Pischke (2009) sugerem como alternativa encontrar grupos com características semelhantes e que não tenham participado do programa, ou de outro com características similares. Desta forma, pode-se utilizar o grupo que participou do programa como grupo de tratamento, e outro que não participou como grupo de controle, considerando as características observáveis. Seguindo esta recomendação, o grupo de controle poderá ser os diretores e professores das escolas municipais do Estado de São Paulo

naqueles municípios que não adotaram qualquer programa de remuneração variável.

A fim de manter as características semelhantes, foram selecionados os municípios paulistas com mais de 200 mil habitantes, baseado nos dados populacionais do censo de 2010, do Instituto Brasileiro de Geografia e Estatística - IBGE.

Desta forma, encontrou-se 39 municípios paulistas com mais de 200 mil habitantes. Nestes municípios se localiza, baseado nos dados do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira - INEP, autarquia ligada ao Ministério da Educação, 1/3 do total de escolas paulistas, o que equivale a mais de 2.700 escolas.

Neste grupo de municípios, originalmente selecionados, foi realizada uma pesquisa, diretamente nas suas Secretarias de Educação, para saber se algum deles poderia ter em funcionamento um programa de remuneração variável. Se afirmativo, o município teria que ser excluído do grupo de controle, pois estaria sendo comparado o desempenho de duas instituições que receberam tratamento, mesmo que fossem de programas diferentes.

Somente no município de São Paulo foi encontrado programa de remuneração variável associado ao desempenho do servidor. Os municípios que informaram ter programa, mas que não era associado ao desempenho, e que caracterizava um mero complemento de salário, foram mantidos como parte do grupo de controle.

Após o tratamento dos dados brutos obtidos no INEP, com exclusão das observações que continham erros de preenchimento, ou estavam incompletas, nestes municípios foram encontradas, baseado nos dados de 2007, os seguintes quantitativos de observações em condições de serem avaliadas:

Tabela 1- Dados disponíveis em 2007

	Estaduais (paulista)	Municipais	Total
Diretores	1.748	953	2.701
Professores	15.955	8.693	24.648
Alunos	257.042	127.461	384.503

Uma vez identificado o universo a ser pesquisado, foi preciso definir como serão trabalhados os dados encontrados. O *status* de Tratamento de um indivíduo  $i$  foi dado por meio de uma variável binária  $D$ , que assume o valor 1 quando participa do programa, e 0 se não participa. O valor da variável de interesse para o indivíduo  $i$  será  $Y_{i1}$ , caso ele seja submetido ao tratamento, e  $Y_{i0}$  caso contrário. O efeito do tratamento para o indivíduo  $i$  ( $\tau_i$ ) é dado por  $\tau_i = Y_{i1} - Y_{i0}$ . No entanto, o contrafactual de um indivíduo não tratado não é observável, dado que o mesmo indivíduo não tem como estar no grupo de tratamento e controle simultaneamente.

Neste caso, segundo Cameron e Trivedi (2010), é possível estimar os resultados comparando indivíduos do grupo de tratamento com os do grupo de controle que tenham características semelhantes. A metodologia do *matching* (pareamento) identifica indivíduos semelhantes em ambos os grupos, dado um conjunto de características observáveis  $X$ , de modo que os resultados desses possam ser comparados.

A análise dos dados disponíveis para a realização desta pesquisa indica que o possível número de características comparáveis relevantes, qualquer que seja a situação e ano que será analisado, é superior a 10, o que é considerado grande ( $2^{10} = 1024$ ), apresentando o problema da multidimensionalidade. Este problema pode ser resolvido com a adoção do método proposto por Resenbaun e Rubin (1983), denominado *propensity score* ou escore de propensão, que consiste em estimar uma probabilidade,  $P(X)$ , de o indivíduo receber tratamento, dadas suas covariadas, que é o conjunto de características observáveis nos diretores, professores, ou alunos, e que serão controladas no experimento.

O método gera uma área de suporte comum, que consiste na seleção apenas de indivíduos tratados com características semelhantes a indivíduos do grupo de controle. Sendo assim, os indivíduos que não encontram no outro grupo par semelhante são excluídos da análise.

Dentro da área de suporte comum, os indivíduos selecionados pelo escore de propensão podem ser pareados por vários métodos. O mais simples deles, e amplamente utilizado, consiste no pareamento por vizinho mais próximo (*nearest neighbour*), no qual o indivíduo do grupo de tratamento, com probabilidade mais próxima de cada indivíduo do grupo de controle, é a ele pareado.

Pode-se também pré-definir a diferença máxima entre as probabilidades dos indivíduos tratados e os controles (*caliper*), porém, nesta pesquisa, a definição de qual a distância máxima adequada obedeceria a critérios arbitrários, sem sustentação teórica. Optou-se, então, por deixar livre o pareamento para todos os indivíduos tratados pertencentes ao suporte comum.

Em adição, o pareamento pode ser feito com ou sem reposição do indivíduo pareado. Nesta pesquisa, dada a ampla base de dados utilizada, optou-se pelo pareamento sem reposição, ou seja, uma vez que foi encontrado o par para o indivíduo tratado, não haverá mais replicação. O recurso com reposição é mais indicado quando o número de indivíduos da amostra é pequeno, o que não é o caso.

Com os dados disponíveis será possível testar o desempenho dos diretores de escola, professores e alunos, antes e depois do programa de bonificação da Secretaria de Educação do Estado de São Paulo, empregando a metodologia descrita anteriormente.

Porém, a metodologia descrita anteriormente não captura os efeitos das características não observáveis e que não variam no tempo, como por exemplo o *background* ou herança cultural familiar. É possível que diretores, professores e alunos, que forem filhos de pais mais estudados, possam ter desempenhos diferenciados. Esta característica não observável (ser filho de pais estudados) não tem seus efeitos capturados pelo *propensity score*.

Segundo Angrist e Pischke (2009), pode-se eliminar tais efeitos utilizando um estimador de Diferenças em Diferenças (DD) e comparando o resultado de antes e depois do tratamento. O método DD consiste na comparação da diferença dos resultados ( $Y$ ) do grupo de tratamento e controle ao longo do tempo (diferença no tempo).

Desta forma o DD pode ser escrito como a diferença do ATT (*average treatment effect on treated*) antes e depois da adoção do programa por:

$$\beta_3 = \tau_{ATT}^D - \tau_{ATT}^A \quad (1)$$

O estimador DD pode ser calculado pelo método de Mínimos Quadrados Ordinários (MQO) a partir do seguinte modelo:

$$\text{Desempenho médio} = \beta_0 + \beta_1 EESP_i + \beta_2 Ano_t + \beta_3 EESP_i * Ano_t + \beta_4 X + \varepsilon_{it} \quad (2)$$

Onde:

$EESP_i = 1$ , se é uma escola estadual paulista (ou seja, se recebeu tratamento) e 0 caso contrário;

$Ano_t = 1$ , se é referente a 2009, 2011 ou a 2013 (depois do programa) e 0 se é referente a 2007 (antes do programa).

O  $\beta_3$  é o coeficiente calculado para a variável de interesse estimada pela interação entre as duas variáveis anteriores, de modo a identificar o efeito do programa em 2009, 2011 ou 2013 nas escolas estaduais paulistas em relação às escolas municipais.

X = vetor de covariadas.

$\varepsilon_{it}$  = Resíduo da regressão

#### 4. Testes

Para a análise do impacto do programa estadual paulista, que foi implantado em dezembro de 2008, foram utilizados os dados dos questionários aplicados aos diretores, professores e aos alunos de ambas as séries, bem como os dados sobre as condições da escola, obtidos pelo aplicador do INEP da Prova Brasil nos anos de 2007 (antes da implantação do programa) e de 2009, 2011 e 2013 (depois da implantação). Os questionários são distintos para cada um dos três elementos avaliados (diretor, professor e aluno).

Dado que o diretor não possui avaliação objetiva, tal como o aluno, seu desempenho foi analisado pela adoção de um conjunto de práticas gerenciais que podem beneficiar tanto o desempenho do professor quanto do aluno. O mesmo acontece com o professor, que teve seu desempenho avaliado pelo emprego de um conjunto de práticas pedagógicas, que podem beneficiar o aprendizado do aluno. O desempenho dos alunos, por sua vez, foi analisado por sua proficiência em português e matemática nas respectivas séries.

Além do desempenho de diretores, professores e alunos, foram selecionadas algumas questões para identificar dados sobre o perfil socioeconômico de cada um deles. Estas questões (características observáveis) foram utilizadas como variáveis de controle e compuseram os vetores de covariadas, que serão melhor explicados nos capítulos seguintes desta pesquisa.

As características observáveis das escolas avaliadas também foram utilizadas para controlar os experimentos do professor e aluno. Para tanto, coletou-se informações disponíveis no questionário preenchido pelo aplicador da prova Brasil sobre as condições de uso da escola, segurança proporcionada aos alunos, e disponibilidade de computadores para professores. Este vetor de covariadas não foi utilizado para controle no experimento dos diretores, em vista que estes têm condições de interferir nas condições de uso desses itens nas escolas, por meio de atitudes proativas, podendo gerar resultados endógenos.

Desta forma, foi possível avaliar o efeito do programa sobre o diretor da escola, professor e aluno, controlando as variáveis

observáveis que podem afetar o desempenho destes grupos. As análises dos três grupos foram feitas de forma distinta.

Já as variáveis observáveis controladas foram segmentadas em um vetor de controles comuns aos três elementos avaliados (diretores, professores e alunos), e em controles comuns apenas aos professores e alunos. Além destes controles comuns, há variáveis observáveis específicas de cada um dos elementos avaliados, que comporão um vetor próprio.

O grupo do controle comum aos três elementos inclui sete variáveis, sendo uma delas o produto interno bruto (PIB) per capita, a preços constantes dos municípios analisados em seus respectivos anos, e outras seis variáveis, obtidas do questionário aplicado ao aluno, que visam capturar o ambiente (entorno) de sua convivência.

Estas outras 6 variáveis indicativas do perfil do aluno, pertencentes ao grupo de controle comum dos três elementos, são características socioeconômicas obtidas do questionário especificamente aplicado ao aluno. Este perfil poderá indicar preferências de diretores e professores com desempenho diferenciado. Em outras palavras, diretores e professores com alto desempenho podem, por exemplo, preferir escolas com maior proporção de alunos brancos –na idade adequada – e com renda familiar melhor, que por sua vez vai indicar áreas da cidade com melhores condições de moradia. Com este controle será neutralizado esse efeito.

Estas variáveis são as médias aritméticas simples de cada escola e não de um aluno em específico, ou seja, refletem a imagem da escola vista pelo diretor e o professor, condizente com a metodologia utilizada para construção do Índice de Desenvolvimento da Educação do Estado de São Paulo - IDESP, que é o índice utilizado para pagamento do bônus para a educação estadual paulista.

No grupo de controle comum aos professores e alunos, foram selecionadas 11 variáveis que indicam as condições de uso da escola, o que inclui a conservação e manutenção predial, disponibilidade de computadores com acesso à internet, e a segurança do entorno. Como informado anteriormente, estes controles não foram utilizados nos testes do diretor para evitar a endogenia dos resultados.

Neste mesmo grupo de controles comuns aos professores e alunos, também foram adicionados os resultados do desempenho do diretor, pois variações positivas ou negativas neste último poderão impactar o desempenho do professor e o aluno. As variáveis que mensuram o desempenho do diretor estão detalhadas na seção seguinte deste trabalho.

Foi controlado o efeito fixo por município. Porém, os testes realizados a posteriori, utilizando estes controles, não geraram diferenças significativas a ponto de justificar a substituição dos resultados apresentados.

Auxiliado pelas informações apresentadas nesta seção, as avaliações do desempenho dos diretores, professores e alunos serão apresentadas nas seções seguintes.

#### 4.1. Testes do efeito do programa sobre o desempenho dos diretores

O desempenho dos diretores foi avaliado segundo a adoção, por parte dos mesmos, de práticas gerenciais na gestão das escolas. E tais práticas, utilizadas para avaliar o desempenho dos diretores, podem ser identificadas a partir dos estudos apresentados por Sammons, Hillman e Mortimore (1995), Coleman, Collinge e Scifert (1993) e em Hanushek (2003), que tratam da efetividade das escolas.

O primeiro estudo cita também que entre 12% e 18% dos resultados dos alunos podem ser explicados pelos fatores inerentes às escolas, espectro de atuação do diretor. Também foi observado que a escola consegue maior efeito no ensino da matemática e ciências do que no ensino da língua nativa (português), que sofre maior influência dos hábitos familiares.

Baseado nestes estudos pode-se identificar características associadas à efetividade da escola, independentemente de sua condição e localização, como sendo<sup>1</sup>:

- 22. Promove atividades de formação continuada
- 24. Reúne o conselho de escola mais de 3 vezes no ano
- 29. Reúne o conselho de classe mais de 3 vezes no ano
- 30. Trabalhou no projeto pedagógico da escola pessoalmente
- 43. Desenvolve reforço de aprendizagem para os alunos
- 94. Escolha conjunta dos livros didáticos

Com base nos dados, notou-se que em todas as seis questões há respondentes que indicam não adotar boas práticas gerenciais em 2007, e que possibilitam testar se a implementação de algum estímulo ou programa de incentivo alterou a adoção de tais práticas nas avaliações seguintes. Porém, as questões 24, 29 e 43 indicam que a grande maioria já adota essas práticas e não se pode esperar grandes variações nos seus resultados, exceto se houver piora. As questões 22, 30 e 94 têm mais liberdade para indicar avanços nos resultados.

As características representadas pelas covariadas se mostraram em quantidade expressiva, sendo identificadas 43 para os diretores, estando entre essas as variáveis controladas em comum e as específicas. Já para grandes amostras, a literatura consultada indica fazer o pareamento com todas as características observáveis, conforme recomenda Stuart (2010).

Após o pareamento, do universo de 47.596 observações possíveis, 5.280 foram selecionadas pelo método do *propensity score*. E, mesmo tendo demonstrado que o pareamento foi feito com rigor, os resultados entre grupo de tratamento e controle, no período de Antes e Depois do tratamento, se mostraram estatisticamente distintos.

Em cada variável dependente se fez o teste de diferenças entre médias, antes e depois do pareamento. As diferenças antes do pareamento são esperadas, pois indivíduos com características distintas são comparados, entretanto não deveria acontecer depois de pareados, porém aconteceram. Isto apontou para a existência de variáveis não observáveis que afetaram os resultados, ou seja, características conhecidas do diretor que puderam ser pareadas no experimento proporcionam a formação de pares não tão perfeitos.

Como descrito no modelo empírico apresentado anteriormente, este viés pode ser corrigido se for aplicado o método do estimador de Diferenças em Diferenças (DD), que permite eliminar a influência das características não observadas com efeitos fixos no tempo. O método avalia se a diferença entre as características não observáveis permanecem constante após a implantação do PRV.

Por conta disso, o estimador DD será calculado pelo método dos Mínimos Quadrados Ordinários (MQO) de forma individualizada para cada variável dependente.

Dado a quantidade de variáveis de controle, a tabela a seguir apresenta apenas os resultados da variável de interesse. As questões estão numeradas conforme constam no questionário que lhes deram origem.

Tabela 2- Coeficientes estimados – Diretores

Variáveis Dependentes	Coeficientes
22. Promove atividades de formação continuada	0,0423 (0,0283) ***
24. Reúne o conselho de escola mais de 3 vezes no ano	-0,0804 (0,0203)
29. Reúne o conselho de classe mais de 3 vezes no ano	0,0135 (0,0130)
30. Trabalhou no projeto pedagógico da escola pessoalmente	-0,0827 (0,0282) ***
43. Desenvolve reforço de aprendizagem para os alunos	-0,0281 (0,0124) **
94. Escolha conjunta dos livros didáticos	0,1178 (0,0259) ***

(1) Desvio padrão entre parênteses

(2) \*\*\*p<0,01; \*\*p<0,05; \*p<0,1

O resultado apresentado indica efeito negativo e significativo no desempenho do diretor de escola estadual, na promoção de reuniões dos conselhos de escola (questão 24), na sua participação na elaboração do projeto pedagógico (questão 30), bem como no desenvolvimento do reforço de aprendizagem para os alunos (questão 43). O efeito indicado pela questão 30 é mais relevante

que das outras duas variáveis, pois, a base de dados indica que havia uma quantidade expressiva de diretores em 2007 que não participava da elaboração do projeto pedagógico, mas poderia fazê-lo nos anos seguintes. Porém, isto não aconteceu, ao contrário, houve redução de 8,27% na participação dos diretores nesta atividade. Não se pode afirmar que o PRV desmotivou os diretores a exercer esta atividade, pois isto pode ter origem em outras

variáveis omitidas. Porém, o programa não foi por si só suficiente para reverter esses outros efeitos negativos. Ou seja, sem o PRV poderia ter sido muito pior.

Outra possibilidade do impacto neutro ou negativo na maioria das variáveis é que o propósito do bônus tenha sido distorcido pelo efeito *crowding-out*, descrito por Weibel, Rost e Osterloh (2009), que é a apropriação do benefício como um direito permanente e não relacionado à melhoria do desempenho. Ou seja, o servidor passa a tratar o bônus como um direito adquirido permanente, que não está associado ao seu desempenho.

Também deve ser considerado que a autonomia da escola possa não existir na prática, cabendo aos diretores ser apenas um ponto de distribuição de determinações vindas do poder central. Por exemplo: o projeto pedagógico já pode vir pré-estabelecido pelo gabinete central e, mesmo que se permita alterá-lo, tal alteração implica em riscos de insucesso com responsabilização do diretor. Caso seja verdadeiro, o diretor irá optar cautelosamente pelo modelo sugerido, que implica em menos riscos para ele.

O efeito positivo significativo é notado apenas na participação do diretor na escolha dos livros didáticos (questão 94). Este resultado é relevante, pois havia uma quantidade expressiva de diretores em 2007 que não participava da elaboração do projeto pedagógico, mas poderiam fazê-lo nos anos seguintes. Nesta atividade houve aumento de 11,78% na participação dos diretores na escolha dos livros didáticos.

No capítulo final deste estudo serão complementadas essas conclusões considerando também os resultados sobre professores e alunos.

#### 4.2. Testes do efeito do programa sobre o desempenho dos professores

Dado que o professor não possui avaliação específica e objetiva, para identificar seu desempenho foram utilizadas questões que mensuram a dedicação nos aspectos pedagógicos vinculados aos conteúdos ministrados pelos docentes no ensino de português e de matemática, no 5º e no 9º ano. As avaliações foram feitas de forma individualizada, por série e disciplina.

Identificou-se, no questionário aplicado ao professor, um grupo de 6 questões sobre as atividades visando a aplicação do conteúdo programático de português, e outro grupo de 6 questões sobre o conteúdo de matemática. Em ambas as situações, as atividades se baseiam na Taxinomia de Bloom, que segue os conceitos desenvolvidos por um grupo de especialistas liderados por Benjamin S. Bloom, nos Estados Unidos, publicados em 1956, sob o título "*Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain (Bloom, Engenhart, Furst Hill & Kathwohl, 1956)*".

Segundo Krathwohl (2002), a taxonomia é estruturada em níveis de complexidade crescente – do mais simples ao mais complexo. Isso significa que, para ensinar uma nova habilidade pertencente ao próximo nível, o professor deve ter ensinado a habilidade do nível anterior.

Conforme demonstram Ferraz e Belhot (2010), embora essa sequência seja a mais conhecida (6 categorias), ela é um pouco

mais complexa, pois é dividida em subcategorias com o objetivo de melhor direcionar a definição dos objetivos instrucionais assim como esclarecer os limites entre eles.

Para fins deste trabalho, será utilizada a versão que foi revisada em 2001, tendo David R. Krathwohl como participante do grupo revisor, um dos criadores da versão original publicada no ano de 1.956. A categorização atual da Taxonomia de Bloom, em níveis de complexidade crescente das habilidades, é a seguinte: 1) Lembrar; 2) Entender; 3) Aplicar; 4) Analisar; 5) Avaliar e 6) Criar.

As alternativas de respostas para as questões selecionadas, com base nos conceitos apresentados, foram padronizadas no questionário do INEP, devendo o respondente selecionar a frequência com que cada atividade correspondente ocorreu. A estrutura se manteve inalterada para todas as questões, sendo que a frequência de respostas variava desde "nunca" até "uma vez por bimestre".

Cada variável dependente foi transformada em uma *dummy*, tratando como satisfatório as respostas na alternativa A e B, de maior frequência, e insatisfatória para as alternativas C, D e E. Esta classificação foi feita com base em consultas realizadas a professores com mais de 20 anos de experiência em sala de aula. Esta consulta não se caracterizou como uma *survey*, pois não foi feita com base em questionários estruturados, nem foi aplicada numa amostra significativa, mas foi o suficiente para a construção de um critério de seleção.

Além dos dois grupos de questões apresentadas, foi utilizada também a questão 55, que não se refere à técnica pedagógica, mas sim ao percentual do conteúdo pedagógico dado em sala. As respostas variam de menos de 40% até mais de 80%, nesta ordem, considerando que quanto mais conteúdo foi repassado, melhor para o aluno. Esta questão foi utilizada tanto para avaliar o professor de português quanto de matemática.

O critério de classificação utilizado também se baseou nas consultas realizadas aos mesmos professores acima citados. Para eles, 60% do conteúdo pedagógico são o mínimo aceitável.

Com base nas questões selecionadas foi avaliado o desempenho dos professores de português e matemática do 5º ano e 9º ano. As avaliações em separado se justificam porque tais professores, independentemente se leciona português ou matemática, são, quase sempre, pessoas distintas. Esta diferenciação é menos intensa no 5º ano, na qual é mais comum o mesmo professor ministrar as duas disciplinas, porém a base de dados, mesmo que com frequência bem mais reduzida, comprova que há distinção.

A avaliação do desempenho em cada atividade foi feita pela média aritmética, por escola, das respostas dos professores nos quatro segmentos apresentados: português 5º e 9º ano e matemática 5º e 9º ano (variável dependente). A média agrupa o resultado dos professores por escola, sendo esta a unidade utilizada nesta pesquisa, mantendo-se coerente com o IDESP, apurado por escola, que é utilizado para pagamento do PRV paulista.

Após o pareamento, do universo de 292.828 observações possíveis, foram selecionadas, pelo método do *propensity score*, o seguinte quantitativo com características similares (tabela 3):

O pareamento se mostrou mais rigoroso para todos os grupos avaliados do que para os diretores. Esse aumento no rigor se deve a um maior número de covariadas, sendo 60 para os professores, enquanto para os diretores foram 47. Tais covariadas representam as diversas características individuais utilizadas para se procurar o par “quase perfeito”.

Tabela 3- Quantitativo de professores selecionados para avaliação

Disciplina e Série	Quantidade
Professor de Português do 5º ano	3.075
Professor de Português do 9º ano	2.524
Professor de Matemática do 5º ano	3.024
Professor de Matemática do 9º ano	2.670

Para eliminar qualquer viés produzido pelo pareamento, foi aplicado o método do estimador de Diferenças em Diferenças (DD).

No resultado, o coeficiente da variável de interesse indicou o efeito no desempenho dos indivíduos tratados, ou seja, professor participante do PRV das escolas estaduais paulista após a implantação do programa em 2008. Este procedimento foi replicado para cada uma das 7 questões referentes às atividades desenvolvidas em língua portuguesa, tanto no 5º quanto no 9º ano, que são as variáveis dependentes, obtendo, com isto, 7 coeficientes distintos por série. O mesmo procedimento foi realizado para as 7 questões referentes às atividades de matemática.

Além das covariadas específicas do professor, controlou-se o experimento também pelas covariadas do ambiente (entorno) em que vive o aluno e o PIB per capita, que são comuns aos três elementos desta pesquisa (diretor, professor e aluno), e também pelo desempenho do diretor, conforme apresentado anteriormente.

A tabela a seguir apresenta os coeficientes da variável de interesse, considerando-a como representativa do critério de avaliação do desempenho do professor. As questões estão numeradas conforme consta no questionário que lhe deu origem.

Tabela 4- Coeficientes estimados – Professor de Português

Variáveis Dependentes	5º ano	9º ano
24. Copiar textos extensos do livro didático ou do quadro	-0, 0870 *** (0, 0271)	-0, 1430*** (0, 0481)
25. Conversar sobre textos de jornais e revistas	0, 0227 * (0, 0121)	0, 0772** (0, 0396)
26. Fazer exercícios de gramática relacionados com textos de jornais ou revistas	-0, 0266 (0, 0178)	-0, 0241 (0, 0467)
29. Ler contos, crônicas, poesias ou romances	-0, 0218 ** (0, 0103)	-0, 1083*** (0, 0438)
31. Usar contos, crônicas, poesias ou romances para exercitar aspectos da gramática	-0, 0531 ** (0, 0165)	-0, 0432 (0, 0492)
32. Fixar nomes de conceitos gramaticais e linguísticos	-0, 1132 *** (0, 0276)	-0, 1138** (0, 0559)
55. Ministrou menos de 60% do conteúdo programático	0, 0347 * (0, 0196)	-0, 0324 (0, 0388)

(1) Desvio padrão entre parênteses

(2) \*\*\*p<0,01; \*\*p<0,05; \*p<0,1

Nota-se, com base na tabela anterior, que os professores de português do 5º ano tiveram efeito negativo, e estatisticamente significativo, no desempenho mensurado pelas questões 24, 25, 29, 31, 32 e 55. Somente a questão 26 para avaliar o desempenho do professor do 5º ano não se constatou significância estatística.

Apenas nas atividades previstas das questões 25 e 55 os professores de português do 5º ano melhoraram seu desempenho e somente a 10% de nível de significância. E, quanto aos professores do 9º ano de português, o efeito é positivo e significativo apenas para a questão 25 – conversar sobre textos de jornais e

revistas, que é uma atividade quase que passiva para o aluno, podendo contribuir pouco para o seu aprendizado. As questões 24, 29 e 32 tiveram efeito negativo, entretanto, as demais 3 questões não apresentaram resultados significantes.

Os resultados dos professores de português são pouco expressivos, indicando que o PRV, por si só, não gerou influência relevante no seu desempenho. Os resultados não são iguais aos do diretor, que foram avaliados na seção anterior, porém, os comentários sobre os resultados dos diretores se aplicam a estes professores aqui avaliados. Em comum têm quase todos os resultados negativos ou sem significância estatísticas.

Tabela 5- Coeficientes estimados – Professor de Matemática

Variáveis Dependentes	5º ano	9º ano
34.Fazer exercícios para automatizar procedimentos	-0,0765*** (0,0238)	-0,0569* (0,0291)
37.Gravar as regras que permitem obter as respostas certas dos cálculos e problemas	-0,0716*** (0,0275)	-0,0503 (0,0445)
39.Interpretar resultados numéricos obtidos para dar uma resposta adequada ao problema.	-0,0088 (0,0102)	0,0377* (0,0245)
40.Lidar com situações que lhes sejam familiares e que apresentem temas do interesse dos alunos.	0,0202*** (0,0079)	-0,0421 (0,0360)
41.Experimentar diferentes modos de resolver um problema ou de efetuar um cálculo.	-0,0057 (0,0058)	0,0087 (0,0270)
43.Experimentar diferentes ações (coletar informações recortar, analisar, explorar, discutir, manipular etc.) para resolver problemas.	-0,0071 (0,0154)	0,0241 (0,0511)
55. Ministrou menos de 60% do conteúdo programático	0,0435** (0,0202)	-0,0457 (0,0386)

(1) Desvio padrão entre parênteses

(2) \*\*\* $p < 0,01$ ; \*\* $p < 0,05$ ; \* $p < 0,1$

Pode-se notar que também nesse grupo de professores os resultados são pouco expressivos, mantendo a tendência dos dois grupos anteriores, avaliados de que o PRV não produz efeitos positivos no desempenho do professor, ou, numa visão mais otimista, o PRV não foi forte o suficiente para reverter uma tendência negativa proporcionada por outros fatores.

Para o 9º ano os testes não capturam praticamente efeito nenhum sobre os professores. É como se o PRV não existisse para eles, já para o 5º ano há 3 questões com forte significância, porém as questões 34 e 37 tem efeito negativo, e a que tem efeito positivo, questão 40, tem coeficiente relativamente baixo.

Parece que os resultados pouco se diferenciam nos grupos até aqui avaliados. No capítulo final deste estudo serão complementadas essas conclusões.

#### 4.3. Testes do efeito do programa sobre o desempenho dos alunos

A avaliação do efeito do programa sobre os alunos não é diretamente o objeto de interesse desta pesquisa, porém como a função do diretor e do professor decorre da existência do aluno, e o bom desempenho do diretor e do professor, poderá, então, refletir no desempenho do aluno, torna-se necessário avaliar o efeito finalístico do comportamento destes sobre a nota do aluno. Essa será uma forma possível, dentre outras, de avaliar o impacto do PRV sobre o serviço prestado pela Secretaria de Educação do Estado de São Paulo.

Para avaliar o desempenho do aluno da rede estadual paulista foi utilizado os dados da Prova Brasil, que é um teste padronizado. Conforme demonstra Hanushek e Rivkin (2010), Hanushek e Woessmann (2008), e Moriconi (2012), as notas em testes padronizados para medir habilidades cognitivas são altamente relacionadas a fatores como os rendimentos escolares.

A avaliação do desempenho em cada disciplina foi feita pela média aritmética por escola das notas dos alunos nos quatro segmentos apresentados: português 5º e 9º ano e matemática 5º e 9º ano. A média agrupou o resultado dos alunos por escola, sendo esta a unidade utilizada nesta pesquisa, mantendo-se coerente com o Índice de Desenvolvimento da Educação do Estado de São Paulo (IDESP), apurado por escola, que é utilizado para pagamento do PRV paulista.

O efeito do programa sobre o aluno foi testado controlando os efeitos do diretor e dos professores de forma separada, ou seja, apurou-se o efeito do aluno controlado pelo desempenho do diretor, mensurado por 6 características observáveis, e depois refez-se a operação controlando pelos desempenhos dos professores, mensurados por 7 características em português e 7 em matemática, sem o diretor. Tais desempenhos são os mesmos que foram apurados nas seções anteriores a esta.

Após o pareamento, do universo de 4.109.265 alunos possíveis, foram selecionados pelo método do *propensity score* o seguinte quantitativo com características similares:

Tabela 6- Quantitativo de Alunos selecionados para avaliação

Controle	Disciplina e Série	Quantidade
Diretor	Português e Matemática - 5º ano	223.183
	Português e Matemática - 9º ano	294.720
Professor	Português do 5º ano	222.360
	Português do 9º ano	266.984
	Matemática do 5º ano	294.720
	Matemática do 9º ano	266.047

Como as características do diretor não dependem da disciplina, há variação apenas nos quantitativos dos alunos por série quando controlado pelo desempenho do diretor. Ou seja, o diretor de escola do 5º ano será o mesmo, não importando se a disciplina é de Língua Portuguesa ou matemática. No entanto, há escolas que oferecem

apenas um dos ciclos do ensino fundamental (5º ou 9º ano), o que justifica ter quantitativo de alunos, e também de diretores, diferentes por série. No caso do controle pelo professor, esse entendimento não prevalece, pois o professor se distingue por disciplina e série.

Mesmo com aumento de 7 covariadas em relação ao professor, o que aumenta o rigor do pareamento, a grande quantidade de observações favoreceu a seleção de uma amostra bem representativa da população.

O coeficiente da variável de interesse indicou o efeito no desempenho dos indivíduos tratados, ou seja, aluno de escola participante do programa estadual paulista de remuneração variável, após a implantação do programa em 2008.

Este procedimento foi aplicado individualmente nas disciplinas de português e matemática, que são as variáveis dependentes, tanto no 5º quanto no 9º ano, obtendo, com isto, 4 coeficientes distintos.

A escala de desempenho do SAEB é medida de 0 a 500, portanto, diferente da escala utilizada para avaliar o desempenho dos diretores e professores, que era uma *dummy* variando entre 0 e 1.

Além das características observáveis apresentadas na seção anterior, específicas do aluno, controla-se o experimento também pelo PIB per capita, comum aos três elementos desta pesquisa (diretor, professor e aluno) e também, de forma alternada, pelo desempenho do diretor e do professor, conforme apresentado anteriormente nas seções próprias que tratam do desempenho destes últimos.

A tabela a seguir apresenta apenas os resultados da variável de interesse.

Tabela 7- Coeficientes estimados – Aluno/Variáveis Dependentes

	Controle	5o Ano	9o Ano
Proficiência Língua Portuguesa	Diretor	0,3822 (0,3447)	-4,2632*** (0,7823)
	Professor	-0,6797 * (0,3502)	-4,5259*** (0,9112)
Proficiência Matemática	Diretor	4,1805 *** (0,3640)	-5,3116*** (0,7871)
	Professor	3,0933 *** (0,3702)	-6,3848*** (0,9236)

(1) Desvio padrão entre parênteses

(2) \*\*\*p<0,01; \*\*p<0,05; \*p<0,1

Os efeitos positivos e significantes foram notados na proficiência em matemática 5º ano, indicando que o aluno pode aumentar sua nota em 4,18 pontos, se controlado pelo desempenho do diretor, e, 3,09 pontos, se controlado pelo desempenho do professor. Tal resultado está alinhado com o que foi observado por Hanushek (2003), o qual afirma que a escola consegue maior efeito no ensino da matemática e ciências do que no ensino da língua nativa (português), que sofre maior influência dos hábitos familiares.

Apesar de significante estatisticamente, considerando a escala SAEB de 0 a 500, parece ser pouco para um programa que custa cerca de 500 milhões de reais ao ano para os cofres públicos. Porém, em educação não se encontrou até agora soluções individuais que gerem grandes impactos. Somente um conjunto de soluções que podem produzir efeitos relevantes.

Os resultados mostram também que controlar pelo resultado do diretor ou professor não gerou alteração significativa no coeficiente. Testes subsequentes, retirando ambos da regressão, também confirmaram a pouca influência destes sobre o desempenho do aluno. Importante salientar que isso não deve ser entendido como se diretores e professores tenham pequena influência sobre o desempenho dos alunos, pois, o que se avaliou neste quesito, é se o efeito do PRV nos diretores e professores influencia os alunos e, neste caso, a resposta é negativa.

No 9º ano todos os coeficientes se apresentaram negativos e significantes, seja em matemática ou português, indicando que o PRV paulista não proporcionou melhoria do desempenho destes alunos. Este resultado contraria a expectativa do PRV, que é o de melhorar o desempenho. Uma possível explicação é a existência de variáveis omitidas que podem gerar desmotivação, tais como

greves de professores e baixa qualidade do material didático. Como não existem dados sobre estas variáveis, elas não são controladas e, se isto for verdadeiro, o PRV fez com que o resultado não fosse mais negativo ainda.

Outro ponto que vale mencionar se refere à robustez dos resultados. Todos os resultados significantes dos alunos foram confirmados quando os municípios foram substituídos por outros estados da federação. Apenas os resultados do 9º ano foram menos negativos, mas continuaram significantes ao nível de confiança de 1%.

No capítulo seguinte deste estudo serão complementadas essas conclusões.

## 5. Conclusões e Recomendações

A hipótese de pesquisa definida foi que a remuneração variável tem efeitos positivos no desempenho do servidor público, e, por consequência, na melhoria dos serviços oferecidos à população.

Pelos resultados encontrados nos testes do diretor, não foi possível afirmar que o PRV serviu para estimular a melhora do desempenho do mesmo nos quesitos avaliados, contrariando o propósito do programa. Pode ser que outras variáveis omitidas sejam a causa do efeito negativo e, neste caso, o programa não causou estímulo suficiente para atenuar os outros efeitos negativos, nem sequer foi um motivador significativo.

Na análise das possíveis justificativas dos resultados encontrados, um ponto a ser observado refere-se à idade do diretor. Pelos dados do INEP, 68% dos diretores tem mais de 40 anos de idade, e cerca de 85% são mulheres. Como afirma Ryan e Deci (2000), os jovens são mais preocupados em obter altos salários,

mas, na medida em que a idade avança, a preocupação migra para os benefícios, principalmente os relacionados à aposentadoria. Se considerarmos que, pela legislação brasileira atual, a aposentadoria poderá ser solicitada a partir dos 60 anos de idade, se mulher, e que os bônus do PRV paulista não são pagos aos aposentados, pode-se supor que o PRV, para estes servidores, seja visto como desestímulo, pois o mesmo não é incorporado à aposentadoria, sendo um indicador de que sua terceira idade terá forte restrição financeira.

Outro ponto que deve ser considerado diz respeito ao bônus pago de acordo com o desempenho médio dos alunos da escola, que está ligado indiretamente ao desempenho do diretor. Pode ser que o PRV não esteja servindo como fator de motivação, pois o diretor sozinho pode estar se sentindo limitado para influenciar o desempenho do aluno, alinhado com Souza *et al.* (2005), que diz que a ênfase nos resultados corporativos pode trazer maior cooperação, mas também um certo acomodamento pela despersonalização do resultado. O pagamento máximo de até 2,4 salários pode não compensar o possível desgaste que os diretores teriam com envolvimento pessoal em cobranças aos professores e alunos pela melhora nos desempenhos.

Ainda tratando do valor do bônus pago, o diretor também pode não se sentir premiado o suficiente para ser motivado, dado que estudos como da OCDE (2005) apontam que cargos de gestão dentro das organizações recebem até o dobro das funções não gerenciais. Neste caso, equipará-lo ao mesmo valor do benefício pago aos professores pode estimulá-los a atuarem como um carona (*free-riding*), deixando ao professor a atribuição de melhorar o desempenho dos alunos e minimizando a importância do seu empenho nas práticas gerenciais que poderiam beneficiar o aluno.

O segundo objeto de avaliação desta pesquisa foram os professores. Seu desempenho foi mensurado com base nas atividades pedagógicas que deveriam ser desenvolvidas em sala de aula.

Notoriamente, os efeitos para os professores de português do 5º ano não foram bons. Em apenas duas das 7 atividades avaliadas houve efeitos positivos e ainda ao nível de confiança de 10%. Já no 9º ano, apenas a atividade de “conversar sobre textos de jornais e revistas” teve evolução de 7,72% dos professores. Pela taxonomia de Bloom, esta atividade está relacionada à capacidade de reconhecer e reproduzir ideias e conteúdos. Tal atividade atinge o nível 1 de dificuldade, o mais baixo pela referida taxonomia.

Os resultados obtidos pelos professores de matemática foram ainda menos promissores do que em português. Dos 14 itens avaliados (7 no 5º ano, e 7 no 9º), apenas 6 foram significantes, sendo que 3 deles foram negativos.

No 9º ano foram identificados apenas dois itens com significância estatística a 10% de nível de confiança.

Numa análise conjunta das duas disciplinas, para o 5º e 9º ano, os resultados apurados por esta pesquisa indicaram que os professores estaduais, seja em português ou matemática, não foram motivados pelo PRV paulista a desenvolver as atividades avaliadas com maior frequência nos anos subsequentes à implantação do programa. Com estes resultados não se pode

afirmar que o PRV gera contribuição relevante para o desempenho do professor.

Entre as justificativas possíveis para os resultados encontrados, notou-se que após a implantação do programa houve redução da participação dos professores em atividade de formação continuada de 83% em 2007, para 76% em 2013. Caso os recursos para financiar o programa estejam subtraindo investimentos na formação do professor, pode-se estar trabalhando na contramão do pretendido pelo PRV, assim como advertido por Weibel *et al.* (2009).

Outra possibilidade se refere à dificuldade que os professores possam ter em identificar meios de alavancar o desempenho dos alunos. Eles sabem quais são as metas e indicadores, mas dado a complexidade das iniciativas, e a demora em perceber seus resultados, falta-lhes confiança na seleção das práticas pedagógicas que darão os melhores resultados. Na incerteza do quanto devem se esforçar, e em qual direção e, ainda, diante do risco de escolhas ineficazes, opta-se por não fazer nada. Para Fryer (2011), se os professores tiverem somente uma vaga ideia de como eles poderão melhorar o desempenho dos alunos, então terão pouco incentivo para aumentar sua dedicação.

Mesmo não sendo considerado parte das justificativas, vale comentar que o valor do bônus foi avaliado nesta pesquisa a fim de identificar se foram seguidas as várias recomendações encontradas na literatura para se evitar os pagamentos simbólicos. Com base nos dados do INEP foi possível identificar que 74% dos professores da rede estadual paulista tem remuneração mensal de até 3 salários mínimos, equivalente em 2013 a R\$ 2.034,00 reais. Os dados gerais divulgados em 2014 pela Secretaria de Educação do Estado de São Paulo foram os seguintes<sup>4</sup>:

Em 2013, 205.869 servidores de 4.183 escolas estaduais que atingiram a meta receberam R\$ 590,2 milhões em bônus. Cerca de 114 mil profissionais ganharam até R\$ 2.500,00. Mais de 52 mil receberam até R\$ 5.000,00. Outros 28.978 educadores tiveram até R\$ 8.000,00 em bônus. Valores superiores a R\$ 8.000,00 foram pagos a 9.894 servidores. Dos 163.824 integrantes do magistério que foram contemplados, cerca de 158 mil são professores.

Os dados disponíveis não são suficientes para afirmar o que se segue, mas pode se inferir que os professores com renda de até 3 salários mínimos mensais, o que dá uma renda de anual de R\$ 26.442,00 (12 salários mais o 13º salário), receberam até R\$ 2.500,00 de bônus. Isto equivale a 9,45% da renda anual, bem acima do 4,1% proposto pelo programa da cidade de Nova Iorque. No programa indiano relatado por Muralidharan e Sundararaman (2009) pagou-se 3% do salário anual e no Quênia apresentado por Glewwe, Ilias e Kremer (2010) foram 2%.

Ou seja, se o programa não incentivou como esperado, o valor do bônus não pode ser considerado uma justificativa para o baixo desempenho.

Baseado nos fracos resultados presume-se haver indicativos de que o PRV paulista está sofrendo com o efeito *crowding-out*, que é a apropriação do benefício como um direito permanente e não relacionado à melhoria do desempenho. Isto vale tanto para diretores como para professores. Neste caso, a interrupção do PRV com a incorporação da premiação ao salário trará melhores resultados.

Por fim, foi analisado o desempenho dos alunos na proficiência em português e matemática, tanto do 5º quanto do 9º ano, visando identificar se as alterações no desempenho dos diretores e professores impactam o desempenho do aluno. Das 8 avaliações realizadas, em 7 foram encontrados resultados significativos. Os benefícios mais animadores foram notados nos alunos do 5º ano, na disciplina de matemática. Alinhado com estes resultados, Oshiro e Scorzafave (2011) apurou ganhos de 6,35 pontos na nota do aluno.

Todos os 4 resultados do 9º ano, indistintamente, sejam em português ou matemática, com ou sem diretores ou professores, são negativos.

De forma ampla, pode-se concluir que o programa de remuneração variável proporcionou melhora na proficiência dos alunos do 5º ano em matemática, mas que esta melhoria não pode ser associada ao desempenho dos diretores e professores. Para os diretores essa associação é naturalmente difícil, pois ele está vinculado ao conjunto de atividades da escola e não a uma disciplina ou série específica. Quanto aos professores do 5º ano de matemática, nenhum dos seus coeficientes se apresentou positivo e significativo nos 7 itens utilizados para avaliar o seu desempenho.

Apesar desses resultados pouco favoráveis aos alunos, os mesmos devem ser tratados com menor relevância do que os resultados dos diretores e professores, pois o impacto para aqueles é indireto. Este estudo visa avaliar se o PRV impactou o diretor e o professor e, por consequência, se esse impacto beneficiou o aluno. Outros muitos fatores, tais como: família, renda etc. podem ter mais relevância no desempenho do aluno. Aqui ficou apenas demonstrado que o PRV indiretamente não produz efeitos relevantes aos alunos.

Baseado na teoria da agência, é possível concluir que o PRV contribuiu de forma tênue na redução de assimetrias entre agente e principal nos itens avaliados. Complementarmente, também foi possível notar, baseado na teoria dos incentivos, que o uso do programa como instrumento de motivação extrínseca não pode ser percebido de forma evidente.

Encerrando este estudo, sugere-se a realização de pesquisa qualitativa com os diretores e professores, bem como com outros representantes da categoria, para identificar sua percepção quanto aos propósitos e benefícios deste programa. A visão de quem está participando do PRV irá contribuir para melhor entender os resultados desta pesquisa. Pode ser que a Prova Brasil não esteja apresentando em seus questionários itens importantes no

desempenho do servidor, tais como assiduidade e pontualidade ou outros valores que o servidor associe mais fortemente ao PRV.

#### Referências Bibliográficas

- Angrist, J. D., & Pischke, J. (2009). *Mostly Harmless Econometrics*. An Empiricist's Companion. Princeton University Press.
- Besanko, David, Dranove, D., Shanley, M., & Schaefer, S. (2012). *A Economia da Estratégia*. (5ª ed.) Porto Alegre: Bookman.
- Cameron, A. C., & Trivedi, P. K. (2010). *Microeconometrics Using Stata*. Texas-USA: Stata Press
- Coleman, P., Collinge, J., & Scifert, T. (1993) *Seeking the Levers of Change: Participant attitudes and school improvement*. (Vol. 4, pp. 59-83). School Effectiveness and School Improvement, International Journal of Research, Policy and Practice.
- Fryer, R. G. *Teacher incentives and student achievement: Evidence from New York City Public Schools*. (2011). National Bureau of Economic Research. Cambridge, MA.
- Glewwe, P., Ilias, N., & Kremer, M. (2010). *Teacher Incentives*. (pp. 205 – 227). American Economic Journal: Applied Economics 2.
- Hanushek, E. (2003) *The Failure of Input-Based Schooling Policies*. (pp. 64-98) The Economic Journal.
- Hanushek, & Rivkin, S. G. (2010). *Generalizations about Using Value-Added Measures of Teacher Quality*. (100(2), pp. 267–271) American Economic Review.
- Hanushek, & Woessmann, L. (2008) *The Role of Cognitive Skills in Economic Development*. (46(3), pp. 607–68) Journal of Economic Literature.
- Hasnain, Z., Manning, N., & Pierskalla, J. H. (2012). *Performance-related Pay in the Public Sector*. A Review of Theory and Evidence. World Bank.
- Jensen, M. & Meckling, W. (1976). *Theory of the Firm: Managerial Behavior, Agency Costs And Ownership Structure*. V. 3, Issue 4, October 1976, Pages 305-360. Journal of Financial Economics.
- Moriconi, G. M. (2012). *Medindo a eficácia dos professores: o uso de modelos de valor agregado para estimar o efeito do professor sobre o desempenho dos alunos*. Tese de doutorado, Fundação Getúlio Vargas/EAESP.
- MTE – Ministério do Trabalho e Emprego. Disponível em: < <http://acesso.mte.gov.br/portal-mte/rais/#2> > Acesso em: 23 de nov. de 2015.
- Muralidharan, K., & Sundararaman, V. (2009). *Teacher performance pay: experimental evidence from India*. National Bureau of Economic Research. Cambridge, MA.
- OECD- Organisation for Economic Co-Operation and Development. *Performance-related pay: policies for government employees*. Paris: OECD Publishing, 2005a.
- Oshiro, C. H., & Scorzafave, L. G. (2011). *Efeitos do Pagamento de Bônus aos Professores sobre a Proficiência Escolar no Estado de São Paulo*. 38º Encontro Nacional de Economia. Anais do ANPEC. Foz do Iguaçu/PR.
- Ryan, R. M., Deci, E. L. (2000). *Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions*. (pp. 54-67) Contemporary Educational Psychology 25.
- Rosenbaum, Paul R., Rubin, Donald B. (1983). *The Central Role of the Propensity Score in Observational Studies for Causal Effects*. Vol. 70, No. 1, pp. 41-55. Biometrika.
- Sammons, P., Hillman, J., & Mortimore, P. (1995). *Key Characteristics of Effective Schools*. School Effectiveness & School Improvement Journal.
- Souza, M Z A. et al. *Cargos, Carreiras e Remuneração*. (2005) Rio de Janeiro: FGV.
- Stuart, E. A. (2010) *Matching methods for causal inference: A review and a look forward*. National Institute of Health.
- Weibel, A., Rost, K., & Osterloh, M. (2009). *Pay for Performance in the Public Sector – Benefits and (Hidden) Costs*. (Vol. 20, pp. 387-412) Oxford: Journal of Public Administration Research and Theory.

<sup>1</sup>O número que identifica a questão equivale ao formulário do INEP.

<sup>2</sup>Consulta realizada em 05/03/2014 em <http://www.educacao.sp.gov.br/noticias/secretaria-da-educacao-paga- hoje-r-590-2-milhoes-em-bonus-para-206-mil-servidores>.