

FUNDAÇÃO GETULIO VARGAS  
ESCOLA DE ADMINISTRAÇÃO DE EMPRESAS DE SÃO PAULO

SILVIO CAMPOS GUIMARÃES

**REMUNERAÇÃO VARIÁVEL NO SERVIÇO PÚBLICO COMO FATOR DE  
INDUÇÃO NA MELHORIA DO DESEMPENHO DO SERVIDOR.**

SÃO PAULO  
2014

SILVIO CAMPO GUIMARÃES

**Remuneração variável no serviço público como fator de indução na melhoria do desempenho do servidor.**

Tese apresentada a Escola de Administração de Empresas de São Paulo da Fundação Getúlio Vargas para obtenção do título de Doutor em Administração Pública e Governo

Área de Concentração:

Política e Economia do Setor Público

Orientador: Prof. Dr. Nelson Marconi

SÃO PAULO  
2014

## FICHA CATALOGRÁFICA

Guimarães, Silvio Campos.

Remuneração variável no serviço público como fator de indução na melhoria do desempenho do servidor / Silvio Campos Guimarães. - 2014.

143 f.

Orientador: Nelson Marconi

Tese (CDAPG) - Escola de Administração de Empresas de São Paulo.

1. Sistemas de bonificações. 2. Sistemas de remunerações salariais. 3. Salários e produtividade no trabalho. 4. Servidores públicos - São Paulo (Estado). 5. Educação pública - Desempenho. I. Marconi, Nelson. II. Tese (CDAPG) - Escola de Administração de Empresas de São Paulo. III. Título.

CDU 658.323(816.1)

SILVIO CAMPO GUIMARÃES

**REMUNERAÇÃO VARIÁVEL NO SERVIÇO PÚBLICO COMO FATOR DE  
INDUÇÃO NA MELHORIA DO DESEMPENHO DO SERVIDOR.**

Tese apresentada a Escola de Administração de  
Empresas de São Paulo da Fundação Getulio Vargas  
para obtenção do título de Doutor em Administração  
Pública e Governo

Área de Concentração:

Política e Economia do Setor Público

**Data de aprovação:**

\_\_\_/\_\_\_/\_\_\_\_\_

**Banca examinadora:**

---

Prof. Dr. Nelson Marconi (Orientador)  
FGV-EAESP

---

Prof. Dr. Ciro Biderman  
FGV-EAESP

---

Prof<sup>a</sup>. Dr<sup>a</sup>. Regina Silvia Pacheco  
FGV-EAESP

---

Prof. Dr. Francisco Caldas Pinto  
UCAM-Universidade Candido Mendes

---

Prof<sup>a</sup>. Dr<sup>a</sup>. Paula Baptista Jorge Louzano  
Faculdade de Educação – Universidade de São Paulo

## **Agradecimentos**

A Deus, que me proporcionou saúde e paz, fatores importantes na longa e dura jornada de conciliação dos afazeres próprios do trabalho, fonte de sustento, e do doutorado, fonte de conhecimento, que ao longo desses quatro anos me consumiram todo o tempo disponível.

Aos meus familiares, que são a base de minha sustentação, agradeço o carinho e a paciência, indispensáveis na trajetória deste trabalho. Em especial, agradeço aos meus filhos Raphael e Luciana, que muito me incentivaram e, ainda, pela compreensão da minha ausência.

A GVPesquisa pela concessão das bolsa que possibilitou a realização deste curso.

A Secretaria da Educação do Estado de São Paulo, em especial ao amigo Willian Massei da Coordenadoria de Informação, Monitoramento e Avaliação (CIMA).

Ao Professor Doutor Nelson Marconi, que soube encontrar tempo para a orientação necessária, com dedicação, atenção e paciência, mesmo envolvido com inúmeras responsabilidades e afazeres.

Aos demais professores do Programa, agradeço os valiosos ensinamentos em sala de aula e as contribuições para a formação profissional e pessoal.

Aos colegas de turma, em especial Leonardo Baroni e Frederico Ramos, que se tornaram bons amigos ao longo dessa jornada, sempre presentes e dispostos a estender a mão com desprendimento pessoal.

À companheira e doutoranda Marne Melo pelas discussões construtivas que ajudaram a balizar este trabalho.

Aos amigos e professores da FGV Sérgio Amad Costa e Enrique Saravia pela ajuda construtiva deste trabalho.

Aos colegas de trabalho Roberto Meireles Pinheiro e Carlos Augusto Costa pela compreensão e apoio nesta empreitada.

Aos companheiros Bruno Oliva, Roberto Aragão e Anderson Cardozo de Souza agradeço-lhes as diversas contribuições dadas na elaboração desta tese.

A toda a minha família, em especial aos meus pais, Dario e Ivanilda, pelo carinho, apoio e incentivo.

A todos vocês, o meu mais sincero agradecimento.

## **Resumo**

O objetivo do presente estudo é analisar os efeitos dos incentivos vinculados a programas de remuneração variável sobre o desempenho do servidor público. Para isso se procurou identificar estudos que apresentassem métodos capazes de correlacionar o desempenho com a adoção dos programas de incentivo. Notou-se que a literatura é divergente quanto à associação entre os benefícios proporcionados e a melhoria do desempenho no setor público. Para contribuir com as pesquisas e com o esclarecimento da divergência observada, desenvolveu-se 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: Desempenho do Servidor Público. Programas de Remuneração Variável. Sistemas de bonificações. Sistemas de remunerações salariais. Salários e produtividade no trabalho. Servidores públicos - São Paulo (Estado). Educação pública - Desempenho.

## **Abstract**

The aim of this study is to analyze the effects of incentives linked to variable compensation programs on the performance of public servants. In this sense, we tried to identify studies that presented methods to correlate the performance with the adoption of incentive programs. It was noted that the literature is divergent as the association between the proportionate benefits and improving performance in the public sector. To contribute to the research and clarification of divergence, we developed an empirical model in order to measure the effects of programs on server performance. Further, tests were performed in a case study on the variable remuneration program applied to public servants in public schools in the state of São Paulo aiming to identify their impact on the performance of school principals and teachers. Since these servers are part of the education system, it was also tested whether the effects of the program directors and teachers have impact on the performance of their students.

Keywords: Performance of Public Servants. Performance-related Pay. Systems subsidies. Systems of wage rates. Wages and productivity in the workplace. Civil servants - São Paulo (state). Public Education – Performance.

## Sumário

1.	Introdução .....	1
2.	Objetivo e Justificativas .....	4
3.	Fundamentos Teóricos .....	6
3.1.	A Teoria Elementar dos Incentivos.....	8
3.2.	Modelos de Recompensas .....	14
3.3.	Conceitos e Características dos PRV's .....	18
4.	Revisão da Bibliografia.....	24
4.1.	Programas de remuneração variável no setor público.....	24
4.2.	Avaliação do PRV nas empresas privadas .....	36
4.3.	A experiência do setor público brasileiro com PRV .....	38
5.	O Programa Paulista dos Bônus na Educação.....	41
6.	Hipótese de pesquisa .....	45
7.	Modelo Empírico .....	47
8.	Dados.....	56
9.	Testes.....	58
9.1.	Testes do efeito do programa sobre o desempenho dos diretores.....	62
9.1.1.	Resultados dos diretores.....	71
9.2.	Testes do efeito do programa sobre o desempenho dos professores.....	72
9.2.1.	Resultados dos professores .....	85
9.3.	Testes do efeito do programa sobre o desempenho dos alunos .....	88
9.3.1.	Resultados dos alunos .....	95
10.	Análise de Robustez.....	98
10.1.	Testes sobre o desempenho dos diretores .....	100
10.1.1.	Resultados dos diretores.....	102
10.2.	Testes sobre o desempenho dos professores .....	103
10.2.1.	Resultados dos professores .....	108
10.3.	Testes sobre o desempenho dos alunos.....	111
10.3.1.	Resultados dos Alunos .....	114
11.	Conclusões e Recomendações.....	117
11.1.	Desempenho dos diretores .....	119
11.2.	Desempenho dos professores .....	120



11.3. Desempenho dos alunos .....	123
11.4. Recomendações .....	125
Referências Bibliográficas .....	126
Apêndice A .....	132

## LISTA DE TABELAS

Tabela 1- Características mais importantes no trabalho por setor .....	12
Tabela 2- Valores de referência para níveis de proficiência .....	42
Tabela 3- Municípios do grupo de controle .....	48
Tabela 4- Dados disponíveis em 2007 .....	49
Tabela 5- PIB per capita municipal.....	59
Tabela 6- Controles comuns ao diretor, professor e aluno.....	60
Tabela 7- Controles comuns ao professor e aluno – Condições de uso da escola .....	61
Tabela 8- Variáveis Dependentes do Diretor .....	64
Tabela 9- Descrição das Covariadas do Diretor.....	65
Tabela 10- Área de Suporte Comum - Diretores. ....	69
Tabela 11- Resultado do Pareamento no Desempenho do Diretor - 2007 .....	69
Tabela 12- Coeficientes estimados - Diretores .....	71
Tabela 13- Atividades Desenvolvidas em Língua Portuguesa.....	74
Tabela 14- Atividades Desenvolvidas em Matemática.....	74
Tabela 15- Conteúdo Programático Desenvolvido .....	77
Tabela 16- Descrição das Covariadas do Professor .....	77
Tabela 17- Quantitativo de professores selecionados para avaliação .....	80
Tabela 18- Área de Suporte Comum - Professores.....	80
Tabela 19- Resultado do Pareamento do Professor de Português da 4ª série .....	81
Tabela 20- Resultado do Pareamento do Professor de Português da 8ª série .....	82
Tabela 21- Resultado do Pareamento do Professor de Matemática da 4ª série .....	82
Tabela 22- Resultado do Pareamento do Professor de Matemática da 8ª série .....	83
Tabela 23- Coeficientes estimados – Professor de Português.....	86
Tabela 24- Coeficientes estimados – Professor de Matemática.....	87
Tabela 25- Proficiência comparativa do Aluno 4ª série.....	89

Tabela 26- Proficiência comparativa do Aluno 8ª série.....	89
Tabela 27- Variáveis dependentes – Aluno 4 série.....	90
Tabela 28- Variáveis dependentes – Aluno 8 série.....	90
Tabela 29- Covariadas selecionadas dos alunos .....	91
Tabela 30- Quantitativo de Alunos selecionados para avaliação.....	92
Tabela 31- Área de Suporte Comum – Alunos .....	93
Tabela 32- Resultado Pareamento do Aluno – 2007 – controlado pelo diretor.....	94
Tabela 33- Coeficientes estimados – Aluno .....	96
Tabela 34- Estados do grupo de controle.....	98
Tabela 35- Dados disponíveis em 2007 .....	99
Tabela 36- Variáveis Dependentes do Diretor .....	100
Tabela 37- Área de Suporte Comum - Diretores. ....	101
Tabela 38- Coeficientes estimados - Diretores .....	102
Tabela 39- Atividades Desenvolvidas em Língua Portuguesa.....	103
Tabela 40- Atividades Desenvolvidas em Matemática.....	104
Tabela 41- Conteúdo Programático Desenvolvido .....	105
Tabela 42- Área de Suporte Comum - Professores.....	106
Tabela 43- Coeficientes estimados - Professores Português.....	108
Tabela 44- Coeficientes estimados - Professores Matemática.....	109
Tabela 45- Proficiência comparativa do Aluno 4ª série entre Estados .....	111
Tabela 46- Variáveis dependentes – Aluno ano de 2007.....	112
Tabela 47- Área de Suporte Comum - Alunos.....	113
Tabela 48- Coeficientes estimados – Aluno .....	115

### **LISTA DE FIGURAS**

Figura 1- Projeto de PRV.....	20
Figura 2- Área suporte comum .....	52
Figura 3- Efeito Médio nos tratados e não tratados .....	53

## 1. Introdução

Os sistemas de gestão de recursos humanos estão atentando cada vez mais para a significativa influência das políticas de remuneração sobre o comportamento das pessoas dentro das organizações. Estes processos procuram obter o comprometimento dos trabalhadores por meio da criação de novas metodologias salariais num ambiente globalizado e que resultem em benefícios para todos nesta “moderna economia do trabalho”.

Para a organização, tais benefícios podem ser traduzidos na otimização da produtividade com menor custo, ou seja, aumento dos resultados. Para quem trabalha, corresponde à motivação para trabalhar, à satisfação em produzir e em se sentir recompensado pelo seu desempenho. Para o cidadão-usuário, espera-se que haja melhoria da qualidade dos serviços oferecidos.

Atualmente as organizações estão sendo incentivadas a seguir algumas linhas de gestão com práticas administrativas que favorecem a emergência do capital humano como fator principal para serem bem sucedidas. De acordo com Ehrenberg (2000) e Flannery (1997), essas transformações geram mudança de valores, de cultura e reconfiguram as estruturas organizacionais. Estas, por sua vez, criam novas formas de administrar as pessoas por meio de políticas de recursos humanos que integram todos os setores da organização e dimensionam suas estratégias para um único foco, isto é, o da valorização das pessoas como condição para obter comprometimento de todos e atingir os objetivos determinados.

As políticas de cargos e salários, que cuidam das promoções e aumentos de salários, e os planos de carreira, que possibilitam ao servidor almejar uma escalada vertical, estão perdendo espaço dentro do ambiente organizacional. Para Karawejczyk (2005), há uma tendência na gestão de pessoas em contratar trabalhadores para múltiplas tarefas e numa organização mais horizontal do trabalho. Com isso, segundo Flannery (1997), faz-se cada vez mais necessário estudar alternativas de gestão que gerem um sentimento de crescimento pessoal, que estimulem o desenvolvimento dentro das organizações, através de políticas que possam construir planos de remuneração vinculados ao desenvolvimento pessoal, ao comprometimento com os resultados, ao interesse em inovar e ao trabalho com qualidade.

Em ambientes caracterizados por formas modernas de organização do trabalho – como celularização, multifuncionalidade e poucos níveis hierárquicos – não faz sentido recompensar a contribuição individual ou coletiva apenas com base em descrições de atividades

e definição de responsabilidades. Para Wood (1996), a questão central é transformar a visão da remuneração como fator de custo para uma visão da remuneração como fator de aperfeiçoamento da organização, como impulsionador de processos de melhoria e aumento da competitividade.

Partindo desses pressupostos, estabelecer uma gestão baseada na remuneração vinculada à produtividade e ao desenvolvimento das pessoas seria uma estratégia possível para impactar positivamente o comprometimento dos trabalhadores, aumentar o interesse em praticar a criatividade e alavancar os níveis de produtividade, todos indispensáveis para o sucesso da organização.

Os modelos tradicionais de remuneração fixa para o servidor exercem mais a função de disciplinar a relação entre contratante e contratado e não se propõem a servir como forma permanente de incentivo à melhoria do desempenho do servidor. Contrapondo este modelo, algumas experiências têm sido tentadas com o propósito de atrelar o cumprimento de metas à premiação em dinheiro ou equivalente, que são denominados como Programas de Remuneração Variável (PRV).

Tais programas de Remuneração Variável dos tempos modernos têm em sua maioria origem na segunda metade do século XX e com maior intensidade a partir dos anos oitenta. Por sua vez, os estudos desenvolvidos visando avaliar os impactos desses programas tem encontrado resultados que se contradizem, ora identificando efeitos positivos, ora negativos, ora neutros. Um dos estudos que talvez melhor apresenta essa dubiedade é o de Marsden (2009), em que o autor questiona o porquê de tantos insistirem na adoção dos programas de remuneração variável se há evidências que eles falham em motivar o trabalhador. Neste mesmo sentido segue Weibel et alli (2009) na análise de 46 programas estudados. Resultados contraditórios também foram encontrados por Hasnain, Manning e Pierskalla (2012) ao analisar 153 programas distintos, dos quais 40% foram considerados neutros ou tiveram efeitos negativos sobre a motivação do trabalhador.

Parece que os argumentos utilizados para avaliar os efeitos dos programas não desmotivaram os elaboradores destes programas a insistir na sua implantação ou mesmo desincentivaram novos gestores a adota-los. Esta contradição sugere que há necessidade de revisar os conceitos em que se baseiam a criação destes programas ou que as métricas utilizadas para avaliar seus impactos não estejam capturando os seus reais efeitos. É neste último ponto que este trabalho pretende oferecer contribuição, identificando e analisando as teorias aplicáveis ao tema, bem como avançando nas discussões do ponto de vista empírico com base nos dados disponíveis.

Na sequência deste trabalho, serão apresentados o objetivo e justificativa do estudo, além dos fundamentos teóricos seguidos da revisão bibliográfica para se conhecer os avanços metodológicos conseguidos até o momento. Em seguida serão apresentados os detalhes do programa paulista de incentivos na educação e a hipótese de pesquisa. Por fim, será apresentado o modelo empírico, seguidos dos testes e apresentação de resultados.

## 2. Objetivo e Justificativas

O objetivo do presente estudo é 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 nos serviços prestados à população por estes servidores.

A relevância do tema é significativa. Em 2011 o Brasil possuía 46.310.631 de empregos formais<sup>1</sup>. Deste total, 9.103.601 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 objeto de um modelo de remuneração variável visando melhorar o desempenho do servidor.

Ainda tratando da relevância do tema, vale lembrar que a oferta de serviços pelo Estado ao cidadão tem sido crescente 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. Na iniciativa privada os indicadores atrelados ao resultado econômico são utilizados como referencial para premiação, porém, resultado econômico não é o objeto principal do interesse do Estado, mas sim o bem estar da população. Diante disso, a avaliação

---

<sup>1</sup> Consulta em 29/07/13 em <http://portal.mte.gov.br/rais/resultados-definitivos.htm>

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.

Também 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. 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 e também das divergências de resultados encontrados nos efeitos do PRV, o presente trabalho pretende contribuir para identificar os impactos dos programas de remuneração variável no desempenho do servidor público.

### 3. 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.

Conforme Dixit, 2002, Gibbons, 1996 e Prendergast, 1999, o modelo do agente/principal é amplamente aplicável, inclusive nas instituições estatais. Todos os servidores de um governo podem ser considerados agentes, pois todos realizam ações ou tomam decisões que podem causar um impacto sobre resultados esperados dos governantes, os principais. Neste mesmo meio, outras camadas nas relações de agência podem existir, tais como um diretor, gerente, chefe de departamento ou qualquer outra forma de gestor público, na condição de principal, e seus respectivos comandados, na condição de agente.

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 satisfeitas: (1) os objetivos do principal e do agente são diferentes, e (2) as ações realizadas pelo agente ou as informações de posse do 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. Sendo observadas, podem ser controladas.

Baseado nestes mesmos autores, pode se entender que em relações de agência, o objetivo do governante/gestor é tipicamente maximizar a diferença entre o valor que recebe em decorrência das ações do servidor e qualquer pagamento que ele faça ao servidor. Na ausência de algum mecanismo que alinhe os interesses das duas partes, o servidor provavelmente não se



importará com o valor gerado para o governante. Ao contrário, o servidor estará interessado no valor que ele recebe com sua participação na relação, menos quaisquer outros custos incorridos em fazê-lo. O servidor valoriza qualquer pagamento realizado pelo governante e pode receber outros benefícios indiretos, como oportunidades de subir na carreira ou incentivos financeiros para um melhor desempenho, ambos provenientes das atitudes que ele empreende em seu trabalho.

Mais importante do que conhecer os problemas e conflitos de agência, é saber como lidar com eles em busca de solucioná-los ou amenizar seus efeitos. Para Besanko (2012), três linhas gerais podem agrupar a maioria das estratégias utilizadas: Monitoramento, Burocracia e Incentivos baseados no desempenho.

O Monitoramento é a maneira de mitigar os problemas de ações e informações ocultas nas relações de agência empregando recursos para observar os servidores ou levantando as informações que os mesmo utilizam para tomar decisões. Isto, como consequência, afeta as escolhas feitas pelos servidores, que tendem a ser mais próximas dos interesses do principal. Porém, o monitoramento sofre algumas limitações significativas. Em primeiro lugar, frequentemente o monitoramento é imperfeito. Dada a complexidade das organizações, saber o que monitorar, a que tempo e com qual ferramenta é um desafio a parte. Em segundo lugar, contratar monitoramento pode ser bem dispendioso. Seja contratando gerentes de monitoramento ou empresas de auditorias, estes custos podem ser substanciais e não resolvem o problema em sua plenitude. Em terceiro lugar, contratar um monitor geralmente introduz uma outra camada à relação de agência, gerando novos problemas.

A Burocracia é a maneira de limitar o conjunto de ações que os servidores podem realizar através de regras e procedimentos pré-estabelecidos de forma que os resultados esperados sejam uniformes entre os servidores. Este é o modelo defendido por Max Weber e que sustentou a organização do Estado ao longo do século XX. Mesmo sendo reconhecida como uma maneira eficaz para atingir resultados esperados, quando adotada isoladamente produz excessos de procedimentos para os servidores, Isto, por sua vez, afeta a produtividade, limita a liberdade de ação e desestimula a criatividade.

O terceiro grupamento de formas para se lidar com o problema de agência é dos Incentivos Baseados no Desempenho. Para fins deste trabalho, este é o que mais interessa, dado que os programas de remuneração variável nele se enquadram. 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 fazer isso, o governante oferece atrelar o pagamento do servidor ao resultado que o governante espera receber em decorrência

da ação do servidor. O servidor ganha mais quando uma ação de governo tem bom desempenho, e menos quando uma ação de governo tem um desempenho ruim, e então fica disposto a realizar ações que beneficiem o governo. Porém, assim como nos dois grupos anteriores, há complicações no ambiente real das organizações.

Primeiro: 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.

Segundo: fornecer incentivo para um servidor que exerce tarefas múltiplas parece ser uma dificuldade. É importante equilibrar os incentivos entre as diferentes tarefas, porém muitas dessas tarefas são difíceis de serem avaliadas. Por consequência, o servidor pode receber incentivo somente para algumas. Neste caso, ele irá canalizar seus esforços às tarefas mensuráveis e que lhe proporcionem retorno financeiro.

Terceiro: 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.

Para saber como lidar com estas dificuldades no ambiente real das organizações, há de se compreender melhor a teoria dos incentivos.

### **3.1. A Teoria Elementar dos Incentivos**

Os problemas de assimetria de informações entre agente e principal, denominados também como problemas de agência, tem na Teoria dos Incentivos, segundo Ryan & Deci (2000) e Gibbons (1996), contribuições importantes sobre a forma de lidar com tais situações.

Para estes autores, que se fundamentam na teoria da autodeterminação, há diferentes tipos de motivações baseadas em distintas razões e objetivos que dão origem a uma ação. A distinção mais básica é entre a motivação intrínseca, que se refere a fazer algo porque é inerentemente interessante ou agradável, e motivação extrínseca, que refere-se a fazer alguma coisa porque isso leva a um resultado diferenciado.

A motivação intrínseca é definida pelas ações que geram a satisfação pessoal. Uma pessoa intrinsecamente motivada age por diversão ou desafio, sem estímulos externos, pressões ou recompensas. Um professor pode ter outras motivações além do salário para exercer sua profissão. Para Vandenabeele (2008), variáveis culturais como o sentimento de estar contribuindo com a própria comunidade, a satisfação em ver seus alunos com bom desempenho escolar, entre outras motivações, podem amenizar substancialmente a insatisfação com o baixo salário do professor. São os motivadores intrínsecos que estariam predominando sobre o desestímulo salarial.

Para o autor, nos seres humanos, a motivação intrínseca não é a única forma de motivação, ou de ação voluntária, mas é importante. Desde o nascimento, os seres humanos, em seus estados normais, são criaturas ativas, curiosas e brincalhonas, mostrando disposição constante para aprender e explorar, e não precisam de incentivos externos para fazê-lo. Esta tendência motivacional natural é um elemento crítico no desenvolvimento cognitivo, social e físico, pois é agindo em busca dos seus próprios interesses que o indivíduo cresce em conhecimento e habilidades. A tendência a ter interesse pela novidade, para assimilar ativamente e aplicar criativamente as competências não se limita à infância e é uma característica importante da natureza humana que afeta o desempenho, persistência e bem-estar através de épocas da vida.

Embora, em certo sentido, a motivação intrínseca existe dentro dos indivíduos, em outro sentido, a motivação intrínseca existe na relação entre os indivíduos e atividades. As pessoas são intrinsecamente motivadas para algumas atividades e não outras, e nem todo mundo está intrinsecamente motivado para qualquer tarefa particular.

A motivação intrínseca é um tipo importante de motivação, porém a maioria das atividades que as pessoas fazem não são, estritamente falando, intrinsecamente motivadora. Este é especialmente o caso após a infância, quando a liberdade de estar intrinsecamente motivado torna-se cada vez mais cerceada por demandas sociais e pelas funções que exigem dos indivíduos assumir responsabilidades por tarefas não intrinsecamente interessantes. Nas escolas, por exemplo, parece que a motivação intrínseca torna-se mais fraca a cada grau de avanço na escolaridade, abrindo espaço para adoção de motivadores extrínsecos.

Quanto à motivação extrínseca, Vandenabeele (2008) diz que a mesma é uma estrutura que envolve as atividades executadas para atingir algum resultado diferenciado e contrasta com a motivação intrínseca no que se refere a fazer uma atividade simplesmente pelo prazer da atividade em si, ao invés de seu valor instrumental. No entanto, ao contrário de algumas perspectivas que veem o comportamento extrinsecamente motivado como

invariavelmente não-autônomo, a teoria da autodeterminação apresenta argumentos distintos e diz que a motivação extrínseca pode variar muito no grau em que ela é autônoma. Por exemplo, um trabalhador que executa sua obrigação só porque teme as sanções do chefe ao não fazê-lo é extrinsecamente motivado porque está fazendo o trabalho a fim de alcançar o resultado diferenciado para evitar sanções. Da mesma forma, um trabalhador que faz a obrigação porque ele pessoalmente acredita que é importante para a carreira escolhida é também extrinsecamente motivado porque ele está fazendo isso pelo seu valor instrumental e não porque considera interessante. Ambos os exemplos referem-se à instrumentos, mas o último caso implica envolvimento pessoal e um sentimento de escolha, enquanto que o primeiro envolve o mero cumprimento de um controle externo. Ambos representam um comportamento intencional, mas os dois tipos de motivação extrínseca variam em sua autonomia relativa.

Dado que muitas das atividades no trabalho não são projetadas para serem intrinsecamente interessante, a questão central diz respeito à forma de motivar os trabalhadores a valorizar e autoregular tais atividades e sem pressão externa, para realizá-las por conta própria. Esse problema é descrito por Ryan & Deci (2000) dentro da teoria da autodeterminação em termos de promover a internalização e integração de valores e normas de comportamento. Neste contexto, internalização é o processo de absorver um valor ou regulamento, e integração é o processo pelo qual o indivíduo transforma um regulamento internalizado como sendo seu próprio. De forma contínua, o conceito de internalização é descrito como uma motivação que varia do comportamento de falta da vontade, do cumprimento passivo, ao compromisso pessoal ativo. Com o aumento da internalização (e seu sentido associado de compromisso pessoal) vem maior persistência.

Os mesmo autores apresentam quatro classificações para a motivação extrínseca associadas às suas origens. A primeira é a motivação por regulação externa, quando o indivíduo executa sua tarefa de forma plena por orientação de normas e regras pré-definidas.

A segunda classificação se refere à regulação injetada (*introjected regulation*), quando o trabalhador aceita as regras e normas elaboradas externamente sem questioná-las. Neste caso, pouco controle é necessário para acompanhar o seu trabalho.

A terceira forma de motivação extrínseca é denominada identificação, quando o trabalhador assume que as normas e regras externas são coincidentes com os seus valores, sem admitir que são exatamente iguais aos seus.

A última forma é denominada regulação integrada, quando o indivíduo assume plenamente os conteúdos da regulação como sendo iguais aos seus e, desta forma, passa a se autorregular.

Esta última forma é considerada o ideal para minimizar os problemas de agência. Porém, nada impede que um mesmo incentivo exerça influências distintas em diferentes indivíduos de maneira que num mesmo grupo haja pessoas enquadradas em qualquer uma das quatro formas anteriormente apresentadas.

Dar incentivo não garante, por si só, bons resultados. É preciso que os incentivos sirvam para reduzir a assimetria entre agente e principal. Porém, saber qual incentivo resulta em motivação para o trabalhador e, dentre os identificados, quais produzem os maiores impactos tem sido objeto de várias pesquisas e seus resultados nem sempre coincidem. Para o trabalhador privado, não é surpreendente saber que o salário atraente é fonte de motivação para o trabalho. Todavia, mesmo entre estes trabalhadores pode haver diferença entre o grau de motivação dependendo se este trabalhador está no início ou no fim de carreira. Por exemplo, se estiver no início, normalmente com boa saúde, valoriza mais os incentivos financeiros que vão lhe proporcionar garantias de bem estar agora e no futuro. Se estiver em fim de carreira, com saúde mais fragilizada, tende a reduzir o valor dado aos incentivos financeiros e dar mais atenção aquilo que lhe proporciona bem estar nesta fase da vida, tais como plano de saúde e benefícios após sua aposentadoria.

Se há diferenças entre os próprios trabalhadores privados, então entre estes e os servidores públicos espera-se que as preferências sejam mais distintas ainda. Alguns pesquisadores, como os apresentados a seguir, deram contribuições através de testes empíricos para esclarecer a questão.

O trabalho de Lewis & Frank (2002) apresenta os resultados de forma comparativa entre aqueles que pretendem trabalhar no governo e aqueles que já estão estabelecidos no governo e utiliza motivadores extrínsecos, como salário atraente, estabilidade no emprego, e motivadores intrínsecos agregados numa variável denominada “motivação para o serviço público”, que inclui a prioridade para ajudar outras pessoas e sentimento de ser útil à sociedade. O resultado foi apresentado comparando os percentuais de respostas. Enquanto 55% dos que querem trabalhar para o governos consideram uma alta remuneração como atrativo para ser um servidor público, apenas 29% dos que já são servidores dão a mesma importância para o salário.

No estudo, a estabilidade no trabalho atrai 47% dos pretendentes contra 33% dos já estabelecidos. Mesmo se apresentando de forma declinante entre os pretendentes e estabelecidos, os motivadores intrínsecos (ajudar outras pessoas e ser útil à sociedade) se mantiveram altos proporcionalmente para os dois grupos (55% contra 36%).

Os estudos de Houston (2000) apresentam resultados segmentados entre motivadores intrínsecos e extrínsecos, sendo que neste último quesito os resultados são

semelhantes aos apresentados por Crewson (1997). Destaca-se os resultados extrínsecos apresentados por este último autor, por sua forma compreensiva, por fazer parte das recompensas financeiras, em específico a remuneração variável, que é o objeto principal deste estudo. A classificação pela importância do que mais atrai ou retém o trabalhador se apresentou na tabela seguinte:

Tabela 1- Características mais importantes no trabalho por setor

Características do trabalho	Setor de emprego			
	Emprego Público		Emprego Privado	
	%	Classificação	%	Classificação
Relevância do trabalho	57,40	1	51,50	1
Oportunidade de promoção	17,80	2	14,50	3
Estabilidade no emprego	11,90	3	7,40	4
Remuneração elevada	10,90	4	22,90	2
Jornada reduzida	2,00	5	3,80	5

Fonte: Houston (2000)

A coluna com o percentual indica a concentração de resposta naquela características em relação ao total de respondentes.

Apesar de não ter sido mencionado pelo autor, que utilizou dados de servidores públicos federais, é possível que, ao se refazer a pesquisa para outros níveis de governo, tais como estados ou municípios, a ordem dessas prioridades possam se alterar. Há necessidade de mais pesquisas para que se possa estender ou não os resultados apresentados na tabela anterior.

Houston também analisa os dados utilizando a regressão logística, que mostra coeficientes negativos para o servidor público nos motivadores: salário atraente, jornada reduzida e estabilidade no emprego. Ou seja, se comparado com o empregado do setor privado, o servidor público não dá a estes motivadores a mesma importância. Por outro lado, o servidor público apresenta coeficiente positivo e estatisticamente significativo para os motivadores: oportunidade de promoção e relevância do trabalho.

Mesmo não sendo considerado pelo servidor público o incentivo mais importante entre os incentivos extrínsecos mais utilizados para minimizar os problemas de agência, os incentivos financeiros baseados no salário são explorados com mais dedicação neste trabalho por sua associação direta com o PRV. Nesta linha, segundo Besanko (2012), um agente irá claramente preferir um salário prefixado a um salário vinculado ao desempenho se não houver vantagens que o motivem a optar pelo segundo. Uma pessoa avessa ao risco prefere um resultado seguro a um resultado arriscado com o mesmo valor esperado. O grau de aversão ao

risco depende de cada indivíduo. Alguns vão exigir grandes compensações para correr riscos, outros nem tanto e haverá aqueles que nem com vantagens expressivas vão optar por correr riscos.

Numa situação hipotética apresentada pelo autor, havendo uma oferta de emprego seguro com salário de R\$ 10.000,00 e outra vaga de emprego arriscado que paga R\$ 4.000,00 fixos e mais uma parcela variável por desempenho até o teto de R\$ 16.000,00, a preferência será pelo emprego seguro. No caso do emprego arriscado o valor esperado do salário é de R\$ 10.000,00 com variância de R\$ 6.000,00 para mais ou menos. Nesta configuração, o agente irá optar pelo salário seguro, pois o risco de receber somente R\$ 4.000,00 pelo um fraco desempenho (neste caso ele perde R\$ 6.000,00 em relação ao salário esperado) não compensa o prêmio que poderá receber se atingir o teto da remuneração, com ganhos de R\$ 6.000,00 acima do esperado. Resumidamente pode se dizer que os bens consumidos com a parcela inicial são mais importantes para o empregado, pois entre estes poderá haver bens essenciais, do que aqueles que compraria com a parcela que for superior aos R\$ 10.000,00 esperados.

Neste caso, para que o funcionário se sinta estimulado a optar pelo salário com incentivo deverá ser oferecido um prêmio maior que as possíveis perdas que poderá ter. Assim, o piso deverá ser elevado para algum ponto entre os R\$ 4.000,00 inicialmente proposto e R\$ 10.000,00, mantendo o teto em R\$ 16.000,00. A dificuldade agora é identificar qual é este ponto em que a preferência do tomador de decisões se altera (ponto de indiferença ou equivalente de certeza), pois tal equilíbrio depende da propensão ao risco que cada indivíduo se sujeita e também do grau de risco que o contrato submete ao agente.

Para o autor há um dilema neste modelo. Ao vincular o salário ao desempenho, o principal oferece mais incentivo. Isto leva a maiores esforços por parte do agente e, assim, a mais resultado. Por outro lado, como a mensuração do desempenho está sujeita a fatores aleatórios, vincular o salário mais intimamente ao desempenho avaliado também aumenta a variabilidade da remuneração do agente. Isso torna o emprego menos atraente para o agente e significa que o principal tem que pagar salários mais altos a fim de atrair o agente. Isso leva a custos mais altos para o principal. A quantidade ótima de incentivos é determinada por um equilíbrio entre essas duas forças. Porém, as organizações raramente tem informações detalhadas sobre as preferências de risco e custo de esforços dos funcionários e, portanto, não podem resolver a equação precisamente com um valor de incentivo ótimo.

O aumento de custos para o principal também pode advir quando a medida de desempenho não capta todos os aspectos do desempenho desejado. Considere o problema de oferecer incentivos ao esforço de professores dividido em atividades que (1) melhorem a

capacidade do aluno em fazer provas e (2) melhorem a capacidade dos alunos de raciocínio em alto nível. Para atividades como multiplicar e interpretar textos é fácil criar uma prova padronizada para avaliar o progresso do estudante. Porém, encontrar medidas claras que meçam a evolução do raciocínio do aluno são mais difíceis. Desta forma, seria previsível que o modelo de avaliação elege-se a capacidade do aluno em fazer prova como medida de desempenho do professor. Isto, por sua vez, deslocará as escolhas dos professores em direção a atividade (1) possivelmente à custa da atividade (2).

Se medidas de desempenho são de baixa qualidade, a organização pode se sair melhor pagando salários fixos e instruindo seus funcionários sobre como dirigir seus esforços para várias atividades. Apesar dessa abordagem não motivar os funcionários a fazer esforço extra no trabalho, por outro lado ela não motiva a ignorar tarefas que são importantes, mas difíceis de serem mensuradas. Os problemas de identificar boas medidas de desempenho podem ser tão severos que a melhor alternativa pode ser confiar na preocupação inerente do professor com o progresso educacional dos alunos.

Para entender como lidar com as questões inerentes aos incentivos financeiros, tais como medidas de desempenho e definição do valor do incentivo, seja ele fixo ou variável, há de se conhecer melhor os modelos de recompensa utilizados, objeto da próxima seção.

### **3.2. Modelos de Recompensas**

As formas estruturadas utilizadas para remunerar o trabalhador pelo seu esforço são denominadas de modelos de recompensas. Seguindo a cronologia do seu desenvolvimento, os três principais sistemas de remuneração podem atrelar a remuneração de um trabalhador à sua função, às suas habilidades ou às suas competências, não sendo incomum encontrar modelos híbridos com os dois últimos. Essa tipologia clássica dos modelos foi descrita por diversos autores (Wood Jr., 2004; Hipólito, 2001; Souza, 2005; Plohow, 2008; Fernandes, 2008)

O primeiro modelo, denominado funcional, é aquele que considera a função do trabalhador como referência para sua remuneração, foi utilizado amplamente até a década de 70. Tal modelo é centrado em cargos e teve origem na necessidade de introduzir um método sistematizado para diferenciar a remuneração dos profissionais à medida que as organizações iam ganhando complexidade no período pós-revolução industrial. A expectativa era de que as



pessoas restringissem seu trabalho àquelas atividades previamente descritas no cargo que ocupavam, situação típica de um ambiente de pouca mudança e competição. A dimensão do cargo traduz a contribuição do profissional e constitui o ponto de partida para a definição salarial.

Neste modelo a atribuição de valor monetário ao cargo e, por consequência, ao trabalhador, a técnica mais comumente utilizada é a definição de pontos seguindo um conjunto de fatores previamente definidos e caracterizados de acordo com níveis crescentes de complexidade, tais como o nível de autonomia requerida, grau de responsabilidade, grau de estruturação das atividades desempenhadas, entre outros.

Para Plothow (2008), as críticas mais contundentes ao modelo funcional se referem à noção de cargo ser um referencial estático/rígido, por impossibilitar o reconhecimento do ocupante e de suas qualificações/contribuições e por desestimular o aprendizado e o desenvolvimento de novas habilidades e competências, quando estas não são diretamente exigidas para o exercício do cargo. Apesar das críticas, o modelo continua sendo aplicado, principalmente por conta da sua simplicidade e objetividade em relação a outros modelos.

O segundo modelo clássico é o da remuneração por habilidades, que abandona o foco no cargo ou função e direciona para a pessoa, mais especificamente, para as habilidades da pessoa aplicáveis ao trabalho. Para Wood Jr. (2004), uma habilidade por ser definida como a capacidade de realizar uma tarefa, ou seu conjunto, em conformidade com determinados padrões exigidos pela organização. Caracteriza-se por agregar conhecimentos e aptidões pessoais aplicados na execução das tarefas.

O processo de construção do modelo parte do mapeamento das habilidades a serem valorizadas na organização para execução das tarefas necessárias para o atingimento dos objetivos organizacionais. A partir da análise das tarefas criam-se os cargos com as características profissionais que obtêm os melhores desempenhos na execução desses cargos. As características que diferenciam estes profissionais dos demais é que devem ser valorizadas e estimuladas, na medida em que justificam um desempenho superior. Alternativamente sugere-se que se faça uma reflexão dos objetivos estratégicos da organização para identificar as capacidades universais necessárias para atingi-los. São essas capacidades que subsidiarão a construção das habilidades individuais a serem almejadas.

Na remuneração por habilidades, os aumentos de salários são diretamente vinculados a um processo de avaliação, no qual o empregado tem de demonstrar a aplicação da habilidade na prática com autonomia. Este mesmo raciocínio pode ser aplicado ao conceito de remuneração variável, principalmente no formato de pagamento por peça produzida, muito

normal na linha de produção, em que o trabalhador mais hábil tende a receber uma remuneração extra pelo que faz.

Algumas críticas são creditadas ao modelo de remuneração por habilidades. As mais contundentes referem-se à dificuldade de definir habilidades para o pessoal administrativo e gerencial, o que dificulta o processo de avaliação; necessidade de disponibilização de programas de treinamento contínuo para permitir a aquisição de novas habilidades e dificuldade em acompanhar mudanças tecnológicas ou organizacionais que implicam obsolescência de determinada habilidade.

Apesar das críticas o modelo continua sendo usado, principalmente para pessoal que trabalha em atividades operacionais.

O terceiro e último modelo de recompensas é o que utiliza as competências para valorizar pessoal, que pode ser considerado uma evolução sobre os dois modelos anteriores. Sua essência consiste na identificação, avaliação, desenvolvimento e reconhecimento de habilidades e conhecimentos que os profissionais devem ter para desempenhar atribuições inerentes a suas posições. Em relação ao modelo de remuneração por habilidades, a diferença refere-se ao público que se pretende atingir. Enquanto a remuneração por habilidades foca nos profissionais de posições operacionais, a competência tem como alvo profissionais de posições técnicas e gerenciais ou que envolvam a aplicação de conhecimentos (Hipólito, 2001).

A definição de competências mais aceita pelos autores (Fernandes, 2004; Souza, 2005; Hipólito, 2008; Wood Jr., 2004b) corresponde ao conjunto de conhecimentos, habilidades e atitudes que justificam um alto desempenho, acreditando que os melhores desempenhos estão lastreados na inteligência e personalidade das pessoas. Trata-se de uma característica interna, que envolve comportamentos observáveis e mensuráveis que se relacionam com o trabalho. A valorização passa a ser feita pelo que as pessoas podem fazer ao invés da valorização simplesmente pelo que fazem.

Segundo Souza *et. all* (2005) e Hipólito (2008), a elaboração de um plano de remuneração por competências inicia-se pelo levantamento das informações relacionadas à estratégia da organização, como: visão de futuro, missão e direcionamento estratégico. Em seguida identifica-se as competências essenciais necessárias para o que uma organização seja bem sucedida, que depois são desdobradas em competências genéricas, válidas para todos os cargos, que por sua vez serão desdobradas em competências específicas, válidas apenas para o cargo.

Na sequência são construídas bandas salariais amplas, alocando os cargos de acordo com sua contribuição relativa para a organização. Esse processo é parecido com o modelo de

remuneração por habilidades, porém, a amplitude das bandas salariais permite mais flexibilidade ao modelo de competências. É possível assim oferecer aumentos maiores para quem tem bom desempenho e se manter dentro da banda salarial tanto no modelo de recompensas por habilidades quanto por competências. Os termos júnior, pleno e sênior caracterizam níveis dentro dos cargos e são muito utilizados para permitir essa evolução salarial.

O principal risco de se adotar o conceito de competências consiste na possibilidade de se produzir listas irrelevantes de competências ultrapassadas, que não estão mais associadas às necessidades organizacionais. Outra dificuldade encontrada no modelo se refere à ausência de relação clara entre as competências disponíveis na organização e os resultados por ela gerado, dado que tais resultados são obtidos em momento posterior à disponibilidade das competências.

Assim como a remuneração por habilidades, o sistema de remuneração e carreira por competências também não privilegia, nem remunera, a senioridade ou o tempo de experiência, como ocorre no sistema funcional.

Souza *et. all* (2005) menciona duas correntes teóricas marcantes nos sistema de remuneração por competências. A primeira, surgida na década de 1970 nos EUA, toma por base o conjunto de conhecimentos, habilidades e atitudes, no modelo denominado de *inputs*, pois se trata do que o trabalhador deve trazer para produzir um determinado resultado (requisitos de acesso). A outra corrente, denominada *outputs*, corresponde à linha de pensamento surgida na Inglaterra e França e entende como competências o conjunto de contribuições disponibilizadas pelas pessoas para a organização (entregas).

Mais recentemente, a literatura relacionada ao serviço público tem defendido que as competências tem que estar associadas ao conceito de efetividade e impactos, também chamados de *outcomes*, que são as mudanças efetivas desejadas pela sociedade em decorrência das políticas públicas, como por exemplo, a redução das desigualdades sociais ou a melhoria do desempenho dos alunos nas escolas públicas.

Se as competências podem ser associadas aos *outcomes*, então podem ser mensurados os impactos que as mesmas proporcionam à organização. Se estes impactos forem positivos, o servidor poderá se beneficiar de parte destes resultados. Para isto, é necessário que um sistema faça o vínculo entre competências e *outcomes* e estabeleça métricas que avaliam os impactos, papel que pode ser desempenhado por um PRV.

### 3.3. Conceitos e Características dos PRV's

Para Wood & Picarelli (2004), a remuneração variável é um incentivo relacionado ao atingimento de resultados em um determinado período, associado ao alcance de objetivos que garantam a continuidade do negócio, a valorização da instituição e a retenção de profissionais através de maior comprometimento dos empregados na busca pelos resultados da organização.

De forma ampla, pode-se definir a remuneração variável como uma forma de recompensa vinculada às metas de desempenho dos indivíduos, das equipes ou da organização.

A definição de Schuster & Zingheim (1995) segmenta em dez modalidades o que pode ser considerado remuneração variável:

- Prêmios especiais de reconhecimento
- Remuneração variável individual
- Remuneração variável em grupo
- Prêmios fixos (“lump sum”)
- Participação nos lucros
- Participação em ganhos
- Participação em metas (“winsharing”)
- Remuneração variável de longo prazo
- Bônus discricionários
- Opções de compra de ações

Qualquer que seja a terminologia empregada, neste trabalho as diversas formas de premiação serão tratadas sob a ótica da remuneração variável.

A remuneração variável pode ser utilizada para valorizar habilidades e/ou competências. A remuneração por habilidades tem como objetivo justamente valorizar indivíduos e grupos pelo uso de suas capacidades, buscando ainda um aperfeiçoamento contínuo destas. Nesta forma o que se foca é o indivíduo, e não o cargo por ele ocupado; e o fato de sua remuneração estar relacionada ao uso e desenvolvimento de suas habilidades tende a promover a motivação para o trabalho.

A remuneração por competências cobre a área que a remuneração por habilidades não atinge. Enquanto a remuneração por habilidades trata do trabalho técnico, funcional,

caracterizado pela reprodutibilidade e previsibilidade, a remuneração por competências abrange o trabalho administrativo, que se caracteriza pela incerteza, abstração e criatividade, características estas que são próprias da atividade do servidor público.

Para a implantação de um programa de remuneração, baseado em tais conceitos, é necessário primeiramente descobrir que competências são essenciais para a organização. Identificá-las requer um trabalho sistematizado de análise da estratégia e competências da organização, para então se determinar quais são necessárias nos indivíduos ou grupos. O resultado mostrará o caminho para remunerá-los adequadamente, promovendo a motivação e o compromisso com as metas organizacionais.

Por sua vez, os programas de remuneração variável (PRV's) são variações dos sistemas de remuneração por competências, sem a perda da sua essência, acrescido de métricas que mensuram os impactos das competências dos indivíduos ou grupos nos resultados (*outcomes*) das organizações. Para se implantar um PRV é necessário ter percorrido todos os passos necessários à implantação de um programa por competências/habilidades. A organização pode adotar o PRV a partir da sua constituição, porém dada sua complexidade e riscos de provocar insatisfação nos empregados, mais comumente se implanta inicialmente o modelo com remuneração fixa e, assim que o modelo se sedimenta na cultura da organização, poder-se-á avançar para a remuneração variável.

A literatura que retrata a remuneração por competências é ampla, enquanto que a que trata adicionalmente da remuneração variável de forma conjunta é um pouco mais escassa. Para Souza, 2005 e Wood Jr & Picarelli, 2004b, a organização deve ter clareza que resultados espera alcançar com cada tipo de incentivo presente no seu programa de remuneração variável e como ele se integra com a sua estratégia de remuneração. Se um dos objetivos da organização for integrar os esforços das diversas áreas para a realização dos objetivos, o PRV deve considerar tanto indicadores de resultados globais da organização quanto valorizar os resultados das áreas. Caso o objetivo seja alinhar os interesses dos servidores com os dos governantes, então deve contemplar resultados de curto e de longo prazo.

Desta forma, a definição do público-alvo, o projeto de PRV, a escolha dos tipos de incentivos, a formas de aferição de resultados e de pagamento dependem em primeiro lugar dos objetivos que se pretende atingir. Na sequência, deve-se entender o contexto da organização. Além dos objetivos do programa, uma série de variáveis influencia as escolhas e decisões relacionadas às características que devem ser contempladas no PRV. Entre as mais importantes, destaca-se:

✓ Variáveis estratégicas – missão, visão de futuro, objetivos estratégicos, fatores críticos de sucesso, valores e crenças da organização, estratégia de gestão do capital humano.

✓ Variáveis organizacionais – configuração da estrutura organizacional, processos internos, características culturais e políticas, demais políticas de remuneração, sistema de planejamento e controle de resultados, sistemas de informação, mecanismos de gerenciamento do desempenho.

✓ Variáveis externas – cenário econômico, tendências de mercado, legislação, organização sindical.

Uma vez definidas as diretrizes que orientam o programa, foca-se nas características específicas. Três categorias de decisões são fundamentais: que tipo de desempenho se pretende recompensar; como este desempenho será avaliado e como se dará o pagamento da recompensa.

A figura 1 sintetiza as decisões:

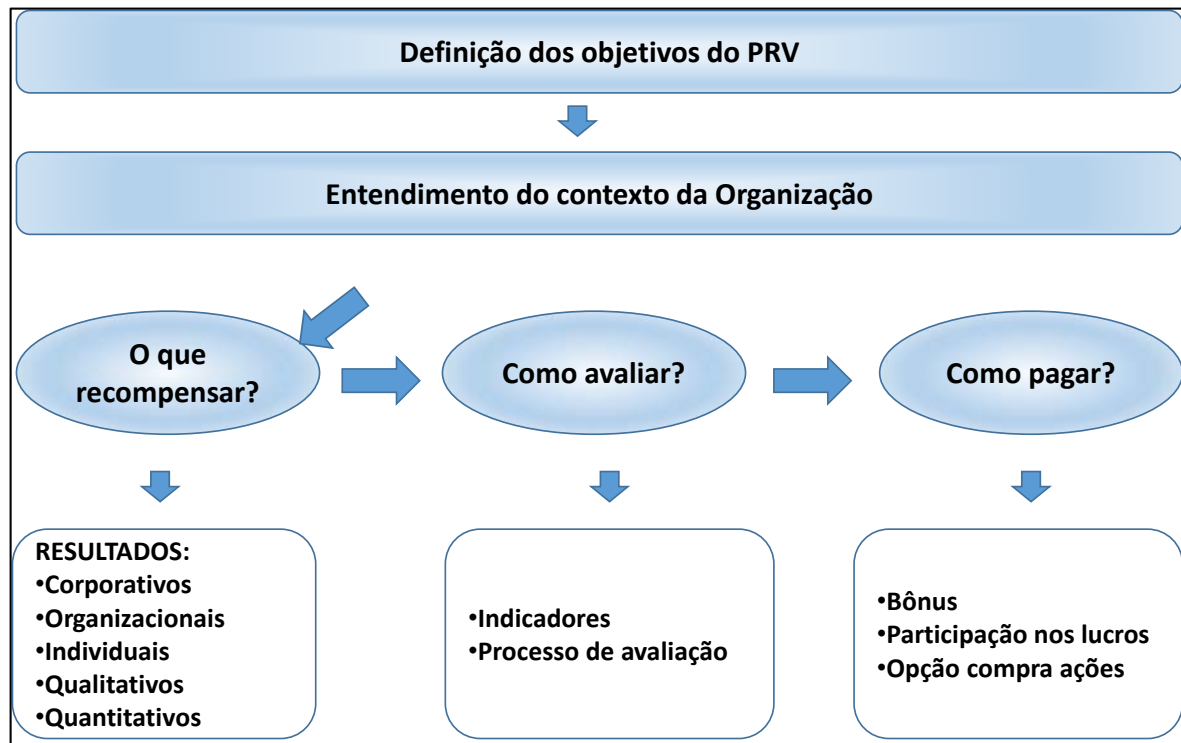


Figura 1- Projeto de PRV  
Fonte: Souza (2005)

A primeira questão sobre o que recompensar absorve a maior parte do esforço de elaboração do PRV. Esta análise deve ser feita sob três perspectivas:

- Perspectiva de abrangência – orientado pelos objetivos do PRV, define se a recompensa será sobre resultados corporativos, por área ou por indivíduo ou ainda uma combinação destes três. Em tese, se o objetivo for restrito, como por exemplo, melhorar o desempenho das notas dos alunos, o PRV poderia se restringir apenas aos professores, porém, uma vez que os resultados da organização dependem da relação sistêmica entre as várias áreas e processos, o programa deve abranger, no mínimo, toda a área de educação.
- Perspectiva da natureza – separa os resultados em qualitativos e quantitativos, sendo que o primeiro deve ser a preferência por ser mais objetivo e de fácil compreensão, porém, nem sempre consegue traduzir corretamente os resultados, como por exemplo, na melhoria na assistência social de um município.
- Perspectiva de tempo – divide-se os resultados em curto, médio e longo prazo de acordo com o ciclo da atividade. Por exemplo: nas empresas o ciclo de atividade normalmente é de um ano, quando se pode avaliar se as ações para melhorar o lucro surtiram efeitos, porém, em políticas públicas, os ganhos na área de educação e saúde são medidos em períodos maiores, podendo ser em décadas.

A segunda característica que um PRV deve ter é o conjunto de indicadores de desempenho que seja efetivo para mensurar os objetivos quantitativos e mecanismos de avaliação para verificar a realização dos objetivos qualitativos.

Outro componente importante na concepção do PRV é o processo de avaliação de resultados, que define quem faz, com que periodicidade e grau de flexibilidade no cumprimento das metas. Também deve conter os instrumentos de comunicação dos resultados aos envolvidos, para que sirvam de referência para melhoria das partes envolvidas.

A terceira e última característica relacionada ao PRV é a forma de recompensa dos resultados bem sucedidos. No caso do serviço público, a premiação quase sempre se dá pelo pagamento em dinheiro como um percentual do salário. Participação acionária ou outra forma que leve à transferência de propriedade para o servidor público não são aceitas pela legislação brasileira.

Há também alguns fatores críticos de sucesso, que a sua não observação poderá minar o sucesso do PRV. A lista não é exaustiva, porém cita as situações mais comumente encontradas. São os seguintes:

- Alinhamento estratégico – os resultados reconhecidos devem ser considerados importantes para área ou setores e também para toda a organização. Se o PRV for implantado apenas na área de saúde, deve se observar se os resultados estão alinhados com as estratégias do governo.
- Atendimento das expectativas de agente e principal – os resultados devem atender de forma equilibrada as expectativas tanto dos servidores públicos quanto dos governantes, para estimular o comprometimento adequado das partes envolvidas no processo de geração de resultados da organização.
- Equilíbrio dos resultados a serem recompensados – a ênfase exagerada em resultados individuais pode levar a um ambiente de competição interna pernicioso e foco apenas no curto prazo. Por outro lado, a ênfase exagerada nos resultados corporativos pode trazer maior cooperação, mas também um certo acomodamento pela despersonalização do resultado.
- Simplicidade - o modelo de PRV deve ser facilmente compreendido pelos participantes. Se for complexo, os funcionários tendem a suspeitar da sua confiabilidade. É importante que o modelo seja o mais objetivo possível para que o funcionário associe o resultado do seu esforço com a recompensa recebida.
- Indicadores de desempenho e processo de avaliação consistente – a lógica do PRV exige que os indicadores de desempenho precisam ser adequados para medir resultados que se busca recompensar, bem como produzir avaliações imparciais. Sem isto, o PRV perde confiabilidade e passa a sofrer resistência.
- Definição de metas e objetivos factíveis – se tais elementos forem definidos de forma que não possam ser atingidos ou que possam ser atingidos muito facilmente, o impacto do incentivo tende a ser reduzido ou neutro.
- Linha direta de comunicação com os envolvidos – os funcionários precisam tomar conhecimento com a antecedência necessária de todas as informações sobre o PRV, bem como ter um canal livre de contato para esclarecimentos.
- Preparação dos gestores e da organização – Por exigir que o PRV seja construído de forma participativa para que tenha sucesso, é necessário que



os gestores e organização estejam preparados para atuar com base em princípios como responsabilização, transparência e meritocracia.

É bastante provável que em algum momento da implantação do PRV apareça algumas das ameaças anteriormente citadas, porém, a existência de gestores do PRV preparados para lidar com situações de ameaça e com habilidade em negociação e comunicação fará que o PRV se consolide como instrumento de incentivo para os envolvidos.

Vale ressaltar que os modelos de PRV precisam ser dinâmicos, adaptando continuamente a realidade e propondo novos desafios.

## **4. Revisão da Bibliografia**

Visando identificar estudos acadêmicos que pudessem referenciar o presente trabalho, realizou-se pesquisa na literatura que trata do PRV, com ênfase nos programas destinados ao setor público por ser este diretamente relacionado ao objetivo deste trabalho. Ao final deste capítulo são relatadas as experiências brasileiras neste tema.

### **4.1. Programas de remuneração variável no setor público**

Há diversos trabalhos publicados que objetivaram analisar modelos de remuneração variável no setor público. Um amplo estudo foi publicado pela Organização para Cooperação e Desenvolvimento Econômico – OCDE, que em 2005 divulgou o resultado de uma pesquisa sobre programas de remuneração variável implantados em seus países membros. Dos 30 países associados, 14 apresentaram modelos estruturados que puderam ser avaliados em profundidade.

Segundo este estudo, a primeira onda de programas de remuneração variável (PRV) foi posta em prática nos países da OCDE em 1980. Os governos do Canadá, Dinamarca, Holanda, Nova Zelândia, Espanha, Suécia, Reino Unido e os Estados Unidos foram os primeiros a adotar PRV. A segunda fase começou em 1990, com a adoção de políticas de PRV na Austrália, Finlândia, Irlanda e Itália. Mais recentemente, países como Alemanha, Coreia e Suíça, bem como alguns da Europa Oriental (República Tcheca, Hungria, Polônia e República Eslovaca), colocaram em prática mecanismos de PRV. Em 2004, a França começou a experimentar o PRV para o nível de altos funcionários (diretores) em seis ministérios. Mais de dois terços dos países da OCDE já adotaram pagamentos relacionados com o desempenho para pelo menos parte de seus servidores.

Nos países estudados, há grandes variações no grau em que o PRV é efetivamente aplicado. Em muitos casos o PRV se refere apenas ao pessoal de direção ou ministérios específicos. Apenas em alguns países membros pode ser considerado como um modelo amplo e formalizado (Dinamarca, Finlândia, Coreia, Nova Zelândia, Suíça, Reino Unido). Noutros países, a cobertura do PRV varia muito. No Canadá, Irlanda, Itália, Noruega, este tipo de programa é aplicado apenas ao nível de gestão; Irlanda, Noruega e França aplicam em

ministérios piloto e para os funcionários mais graduados. Às vezes, porém, os gestores de topo são excluídos do PRV (por exemplo, Finlândia). Um dos motivos é evitar a avaliação do desempenho daqueles que são nomeados através de indicações políticas.

Mesmo diante da diversidade de modelos encontrados, o estudo identificou algumas tendências comuns, como segue:

- Políticas de remuneração variável se espalharam a partir do nível de gestão para cobrir diversas categorias de pessoal nos últimos dez anos (1995-2005);
- Entre as políticas de remuneração variável, tem havido certo aumento na utilização de sistemas de desempenho coletivo, ao nível de equipe ou organizacional;
- Amplos esquemas padronizados de remuneração variável evoluíram para sistemas mais descentralizados, o que facilita a delegação de funções gerenciais;
- Há um aumento da diversidade dos critérios para avaliação de desempenho. Tendem a convergir na atenção dada aos resultados, mas consideram também as competências e habilidades sociais em geral;
- Sistemas de avaliação de desempenho são menos padronizados, menos formalizados e menos detalhados que há dez anos. As avaliações de desempenho confiam mais na avaliação dos objetivos pré-identificados e em diálogo com a gestão da linha do que estritamente em indicadores quantificáveis;
- O tamanho dos pagamentos por desempenho é geralmente uma percentagem relativamente modesta do salário base, especialmente entre os trabalhadores não-gerenciais. Incrementos por mérito tendem a ser menores do que os bônus e são muitas vezes inferiores a 5% do vencimento base. Bônus, que tendem a complementar ou substituir incrementos de mérito, são mais elevados - mas em geral, prêmios máximos representam menos de 10% do salário base dos servidores. Ao nível da gestão, os pagamentos de desempenho são geralmente mais elevados, em torno de 20% do salário base.

O mesmo estudo afirma que, enquanto a remuneração variável parece motivar uma minoria do pessoal no setor público, a grande maioria simplesmente não vê isso como um incentivo para trabalhar melhor. Isso ocorre porque as recompensas são muitas vezes limitadas no setor público, e também porque o conteúdo do trabalho e as perspectivas de progressão na carreira têm sido considerados os mais fortes incentivos para os funcionários públicos, confirmando os resultados dos estudos de Houston (2000) e Crewson (1997).

Apesar dessas limitações importantes, o PRV se mostrou como uma janela de oportunidade para melhoria da gestão e mudanças organizacionais. Em outras palavras, a

introdução do PRV não pode, por si só, melhorar a motivação e o desempenho, mas as alterações necessárias para introduzir o PRV podem. Isto é especialmente verdadeiro para prestação de serviços públicos - arrecadação, segurança, saúde - quando a organização pode ser mais diretamente ligada às demandas dos cidadãos do que as da administração central. O PRV pode facilitar uma ampla mudança organizacional por vincular o pagamento dos bônus aos novos objetivos no âmbito individual e no âmbito departamental.

O estudo indica também que o PRV proporciona aos gestores um incentivo adicional para gerir de forma eficaz e os estimula a adotar uma abordagem baseada em metas. O estabelecimento de metas e a avaliação motivam o contato direto entre os funcionários e seus gerentes de linha no qual a forma de trabalho pode ser discutida e explicada. O estabelecimento de objetivos funciona bem quando acompanhado por uma maior interação entre o gestor e o agente de modo que qualquer redução de controles formais é substituída pelo controle informal. Outro nível em que a avaliação do desempenho aparece como positiva reside no âmbito de relacionar objetivos mais amplos da organização aos dos servidores.

A pesquisa da OCDE finaliza afirmando que a implantação da remuneração por desempenho tem evidenciado problemas de concepção e execução (por exemplo, a falta de dotação orçamentária e formação dos gestores). Os mais importantes são aqueles relacionados ao acompanhamento e avaliação de desempenho, especialmente quanto à forma como os gestores conduzem as avaliações. Uma estrutura inadequada para gestão de desempenho (por exemplo, a falta de transparência, confiança ou a delegação de gestão dentro de uma organização) é um sério obstáculo ao PRV. O impacto dos incentivos financeiros do PRV parece ser limitado e outros incentivos são mais influentes na motivação dos funcionários, tais como a satisfação no trabalho e a progressão na carreira. Para a OCDE a importância do PRV tem sido superestimada.

Apesar de ser uma pesquisa ampla da questão, a avaliação se restringiu aos aspectos qualitativos dos PRV's, sem abordar à avaliação dos impactos causados no desempenho do servidor.

Em outro estudo, Perry *et al.* (2009) em busca da eficácia e efetividade dos programas de PRV identificaram 68 estudos publicados entre 1977-2008 em diferentes países que aplicaram a remuneração variável. Após estabelecer critérios, os autores selecionaram 57 deles para uma análise comparativa. Os mesmos ainda segmentaram os estudos encontrados em dois períodos, 1977-1993 e deste ano até 2008. Os autores fazem uma análise qualitativa dos estudos e concluíram o seguinte:

Nossa análise confirma as conclusões sobre a eficácia limitada do pagamento contingente no setor público. O somatório dos resultados não parece ser diferente para os períodos anteriores a 1993 e de 1993 até o presente, e isolar os estudos pós-1993 levou a conclusões idênticas às do pré-1993. Assim, no âmbito global, a nossa análise considera que pagamentos relacionados com o desempenho no setor público sempre deixam de cumprir sua promessa.

No estudo de Marsden (2009) foi apresentado o seguinte paradoxo dos programas de remuneração variável: se há diversos estudos acadêmicos indicando que os resultados ficam aquém do esperado e que as variáveis envolvidas no seu sucesso são muitas e complexas, por que continuam insistindo em adotar os PRV nas organizações públicas?

O autor relata inicialmente a recomendação dada pela *Priestley Commission* em 1956 como uma prática a não ser adotada como instrumento de motivação do servidor público. O tema não prosperou muito desde então. Porém, na década de oitenta e após a publicação de estudos indicando que poderia haver modelagem de programas de forma a torna-los bem sucedidos, desde que o programa premiasse os bons desempenhos e penalizasse os ruins, o tema voltou à agenda do serviço público. Neste período os governos de Margareth Thatcher e John Major na Inglaterra adotaram diversos programas de remuneração variável, atingindo uma parcela significativa do serviço público inglês.

Outra questão sobre a qual advertiram os estudos da época foi da existência do problema de agência, com múltiplos principais no governo, ou seja, há dificuldade no relacionamento agente-principal, pois enquanto o agente é solitário o principal se subdivide em muitos, tais como políticos, usuários, imprensa etc., que por sua vez avaliam o desempenho do servidor através de suas próprias óticas.

Continuando sua avaliação, Marsden identificou estudos que reconhecem ganhos, mesmo que modestos, mas houve indicativos de que a forma de avaliar poderia não estar medindo adequadamente o desempenho do servidor. A evolução maior parece vir do relacionamento entre a linha gerencial, que avalia, e o servidor, o avaliado. Se esta relação for justa e inspiradora de confiança, então o sistema poderá ser considerado bem sucedido.

Outra constatação de Marsden foi que as diversas agências governamentais britânicas adotaram arranjos diferentes em seus programas. Para evitar grandes discrepâncias entre modelos, o autor sugere que avaliações dos indicadores quantitativos sejam feitas com bons tratamentos estatísticos, o que pode ser de vital importância para o sucesso do programa, sem desprezar estratégias de *benchmarking* em outras instituições similares para ajustar seus programas às melhores práticas.

As conclusões do autor sobre o paradoxo apresentado foram de, apesar dos muitos programas falharem em suas principais promessas, há ganhos não desprezíveis para todos. Para a instituição, mesmo que não atinja seus objetivos, há uma revisão em seus valores e propósitos. Em alguns casos, há melhoria de qualidade do servidor pelo fato da remuneração se tornar mais atrativa para candidatos mais capacitados. A linha gerencial torna-se mais dedicada e melhora sua eficiência, porém os próprios têm dúvida da sustentabilidade dos resultados no longo prazo.

Bowman (2010) examinou pesquisas acadêmicas, artigos recentes publicados na mídia, publicações governamentais e dados de 15 entrevistas não estruturadas de 30 a 50 minutos por telefone no ano de 2008 com interesse em avaliar os programas de remuneração variável no governo federal americano. O autor buscou justificativas para as contínuas tentativas de viabilizar novos programas de remuneração variável, apesar dos estudos indicarem resultados consistentemente insatisfatórios.

No mesmo estudo, ao serem analisados dois relatórios governamentais, *The Civil Service Reform Act* em 1978 e *Recognition Act* em 1984, foi apontado que houve constrangimentos dos participantes, atrasos no programa, custou mais do que o previsto e não reconheceu corretamente os melhores servidores, concluindo que se houve efeitos positivos na produtividade, satisfação do trabalhador e no *turnover*, foram limitados ou nenhum. Quaisquer que sejam as virtudes assumidas por estes programas, administradores foram incapazes de usá-los de uma maneira robusta. Em alguns dos trabalhos citados por Bowman, tais programas são piores do que não fazer nada, porque viciam, corrompem os interesses intrínsecos em trabalhar. Mesmo sendo repetidamente provado o seu mau funcionamento, o governo reluta em abandonar esse modelo.

No segundo relatório do *Government Accountability Office* (GAO) em 2007 as críticas não foram menos contundentes, indicando que 81% das pessoas acreditavam que moralmente estavam piores que antes e que houve sindicalização excessiva como resultado. Em outro programa, *Senior Executive Service*, constatou que não houve impacto no desempenho, afetou negativamente a motivação e o moral, antecipou aposentadorias e desencorajou as gerências de nível médio a adotar o programa.

Na contramão destes relatos e apresentando pontos positivos, o relatório do *National Institute of Standards and Technology* constatou que o programa deles contribuiu para eficiência dos altos talentos, reteve os melhores talentos e expandiu a autoridade dos gerentes para contratar e decidir sobre remuneração.

A análise das experiências práticas feitas por Bowman indica a dificuldade em associar políticas de recompensa com os resultados desejados. Decidir quem recebe, quanto recebe e pelo que recebe é um problema complicado de se resolver.

Bowman constata que nas empresas, apesar da pouca documentação, os resultados deixam a desejar. Uma pesquisa empírica demonstrou que apenas 17% das empresas acreditam que seus programas de pagamento por desempenho foram bem sucedidos. No geral, pesquisadores não acusam os programas de não funcionarem sob determinadas circunstâncias, mas que a condição ideal para que funcionem raramente é encontrada na realidade empírica. Eles endossam a visão na qual os altos e baixos dos programas são mais influenciados por mudanças sociais, políticas e modismos econômicos, sem acumular evidências científicas de quão bem os programas funcionam.

Concluindo o trabalho de Bowman, o mesmo faz recomendações importantes além das já apresentadas: planos precisam ser bem desenhados, ajustando as expectativas dos ganhos e ser implantado num ambiente de confiança entre as partes. Recomenda-se também evitar a adoção de programas baseados em modelos pouco sofisticados que assumem uma relação direta entre motivação e remuneração.

Outra contribuição ao presente estudo vem de Christopher Hood (2006), que realizou um estudo comparando os programas de remuneração baseado em desempenho no Reino Unido com o sistema de metas utilizados na antiga União Soviética e em outros países socialistas. Para ele, este sistema leva ao que chamou de *gaming*, que é o uso de artifícios pelos servidores públicos para atingir as metas estabelecidas com dados e informações que, na realidade, estão superavaliados.

O autor constata que as práticas empregadas pelos órgãos britânicos nos programas lançados pelo primeiro ministro inglês Tony Blair em 2001 eram as mesmas utilizadas pelos órgãos soviéticos, especialmente na era Brezhnev (1964-1982). As distorções nas avaliações das metas atingidas foram permeando o sistema até sua ruína completa.

As auditorias efetuadas nos órgãos ingleses constataram dados inflados na maioria deles. Nos órgãos de saúde foram encontradas situações como a retenção de pacientes dentro das ambulâncias para não registrar suas entradas e assim diminuir o tempo de atendimento. Também foi constatado extração das rodas das macas com pacientes para que estas fossem consideradas como leito normal. No geral, tais iniciativas tinham a concordância ou complacência dos gestores e até mesmo de políticos interessados em apresentar aos eleitores os progressos obtidos no serviço público.

O debate entre aqueles que defendem o sistema de metas e os que defendem outras abordagens, tais como o sistema de melhorias contínua, é problema central na gestão por desempenho. Estudiosos do sistema de metas têm identificado ao menos três grandes tipos de *gaming* e estratégias em torno das metas. Um é conhecido como *ratchet effect*, quando os responsáveis por fixar a meta para o próximo ano, o fazem acima da meta do ano anterior, porém abaixo da fronteira de possibilidade de produção.

Um segundo tipo de comportamento é conhecido como *threshold effect*, quando a meta definida para todas as unidades de um sistema não oferece incentivo para a excelência e na verdade encoraja as unidades de alto desempenho a reduzir seu desempenho.

O terceiro tipo de comportamento consiste na distorção das informações ou manipulação dos relatórios com resultados.

Algumas afirmações podem ser feitas com confiança sobre o sistema britânico. Primeiro, os avanços indicados nos relatórios de desempenho não podem ser dedicados exclusivamente ao sistema de metas, pois no mesmo período houve aumento significativo dos gastos públicos. A segunda afirmação é de que os relatórios sobre desempenho ficaram longe de demonstrar a real situação e não eram de confiança. A terceira afirmação é de que há fortes evidências de que foram utilizados os três tipos de *gaming* anteriormente descritos, os mesmos tipos que foram encontrados na literatura sobre o regime de metas soviético.

Por fim, Hood identifica alguns antídotos para combater o *gaming* no modelo de remuneração por desempenho: adotar regras rígidas; utilizar dados precisos; detalhar as metas, realizar auditorias, punir também aqueles que são complacentes com a indisciplina e o erro.

Weibel *et al.* (2009) oferece uma contribuição apresentando um estudo utilizando uma amostra de 46 estudos, nos quais buscou identificar se os incentivos financeiros são relacionados ao desempenho. A preocupação dos autores foi em desvendar os reais efeitos dos PRV, dado que alguns estudos afirmam que há efeitos positivos no desempenho, outros estudos apresentaram justamente o contrário, que claramente a correlação é negativa.

O estudo ajuda a explicar o sucesso modesto de remuneração por desempenho do setor público, baseando-se em cinco justificativas. Em primeiro lugar, parece que no setor público elevar a motivação é uma dificuldade. Assim, a remuneração por desempenho pode potencialmente criar um efeito de *crowding-out* forte (apropriação da premiação pelo servidor como um direito permanente, assemelhando-a à remuneração fixa). Em segundo lugar, o financiamento público é claramente mais limitado do que o financiamento do setor privado. Em terceiro lugar, remuneração por desempenho pode reduzir os investimentos em capacitação e na seleção de servidores contratados. Em quarto lugar, o problema do servidor executar tarefas



múltiplas pode representar uma dificuldade adicional para o pagamento por desempenho no setor público. O pagamento por desempenho exige medição exata do desempenho e conhecimento prévio pelo servidor da sua atribuição para que o esforço individual seja eficaz, no entanto, as instituições públicas muitas vezes têm de entregar produtos e serviços complexos, tais como "boa saúde" ou "boa educação". Em quinto lugar, do ponto de vista político-econômico, a aplicação de recompensas por desempenho também traz o risco de manipulação política.

O esforço de Weibel *et al.* para encontrar um modelo mais objetivo para mensurar os ganhos e perdas do PRV é louvável, porém os resultados não esclarecem em definitivo quais os impactos causados pelo PRV na produtividade ou na qualidade dos serviços públicos.

Nesta mesma linha de raciocínio, Pacheco (2009) afirma que:

Na mensuração de impactos é difícil estabelecer relações de causalidade entre as ações empreendidas e o resultado observado – é difícil isolar, dentre as inúmeras variáveis que afetam a situação, aquelas diretamente ligadas aos serviços prestados por uma determinada organização pública. Por vezes as ações requerem longo tempo de maturação para que os impactos possam ser observáveis. Os impactos desejados podem depender de mudanças substanciais no comportamento dos cidadãos. Por outro lado, mensurar produtos pode levar a um foco excessivo no curto prazo.

Pacheco (2009) afirma em seu trabalho que, apesar de vários autores ressaltarem a importância da gestão por resultados para a melhoria do desempenho do setor público, apontam também polêmicas e problemas frequentes associados a essa nova forma de gestão; mesmo assim, tais problemas não têm justificado o abandono dos modelos; têm, ao contrário, levado ao seu aprofundamento, buscando corrigir rotas e superar os obstáculos identificados.

Hasnain, Manning e Pierskalla (2012) fizeram uma revisão teórica e buscaram evidências dos efeitos do pagamento associado ao desempenho, assim considerada qualquer forma de compensação na qual a remuneração final de um empregado é função de alguma forma de avaliação do desempenho.

Os pesquisadores identificaram 153 estudos relacionados ao tema que objetivaram de alguma forma aferir os resultados obtidos. Em 93 deles foram encontrados efeitos positivos diversos, outros 37 foram considerados neutros e apenas 23 constataram efeitos negativos.

Nos estudos avaliados o desempenho poderia ser obtido por meio de avaliações qualitativas ou medidas quantitativas dos *inputs* (alocação de esforços, comparecimento ao trabalho, contribuição voluntária no ambiente de trabalho e aquisição de novas habilidades/conhecimentos); *outputs* (conclusão de compromisso pré-acordado ou número de

clientes/casos atendidos); ou *outcomes* (teste de avaliação de estudantes, avaliação dos clientes, taxa de utilização dos serviços ou criação de receita).

Do total inicial foram selecionados 110 estudos, que avaliavam atividades semelhantes às do setor público, caracterizadas por serem trabalhos complexos e de resultados difíceis de serem observados. Estes, por sua vez, foram classificados em *craft jobs* (serviços que dependem de habilidades pessoais, mas os resultados são similares para quaisquer indivíduos ou grupos, p.e. professor, médico, auditor) e *coping jobs* (serviços que dependem de habilidades pessoais, mas os resultados não são comparáveis entre indivíduos ou grupos, p.e. gerentes, executivos). Os que foram excluídos (43) focavam em atividades manufatureiras ou outra com movimentos repetitivos. Neste grupo de 110 constatou-se que 65 apresentaram efeitos positivos. Sendo que destes, 60 estavam classificados como *craft jobs* e apenas 5 eram *coping jobs*. A desproporção destes últimos se justifica por ter menor número de estudos e não por diferença de sucesso entre eles. Ao final do estudo constatou-se efeito positivo e significativo nos considerados *craft jobs*, enquanto nos *coping jobs* é muito menor.

Os autores chamam a atenção pela falta de estudos que avaliam os efeitos de longo prazo. A maioria realiza conclusões pontuais, podendo ser modificadas no tempo na medida em que o programa se tornar mais familiar para os participantes, inclusive por comportamentos negativos (p.e. *gaming*).

Assis (2012) realizou estudo na Secretaria de Segurança Pública do Estado de Minas Gerais em busca dos efeitos do PRV nas estruturas da Polícia Civil, Militar e Corpo de Bombeiros. O autor procurou identificar os efeitos organizacionais causados pela adoção de metas e a remuneração variável sobre o desempenho dos indicadores pactuados. Também foram realizadas entrevistas semiestruturadas com altos executivos, policiais de ponta, bombeiros e especialistas em segurança pública, bem como análise das reportagens na imprensa sobre o PRV, além de observação participante com a equipe responsável pelo acompanhamento de metas do governo.

Os resultados foram distintos para cada instituição avaliada. Para o autor, na Polícia Civil notou-se:

Houve um importante aumento de produtividade, que auxiliou a esvaziar as prateleiras de inquéritos antigos. A instituição começou a controlar o desempenho de suas unidades. O conhecimento sobre níveis de desempenho permitiu, por exemplo, um maior pragmatismo na conclusão de inquéritos, o que é um ganho importante para a instituição. Entretanto, não houve um processo de comunicação sobre o propósito das metas, e a ausência de um líder com legitimidade para traduzir e articular o modelo com os objetivos organizacionais cooperou para que

alguns gestores intermediários cobrassem de qualquer forma: “quero número!”. Esse tipo de cobrança contribuiu para a queda na qualidade dos inquéritos.

Na Polícia Militar ficou constatado pelo autor que:

É notável que projetos anteriores ao PRV geraram consequências significativas e deixaram um legado de aprendizado institucional. O Acordo de Resultados, na verdade, intensificou e capilarizou um processo que já ocorria na instituição. A cobrança de resultados, que já era intensa, ficou ainda mais forte. O acordo contribuiu para que a noção de resultado e metas fosse mais difundida, especialmente no interior do Estado.

O efeito mais positivo foi notado no Corpo de Bombeiros, conforme dito pelo autor:

O Corpo de Bombeiros Militar é um caso de referência para os entusiastas do uso de metas e mensuração na administração pública. O impacto do Acordo de Resultados e do Prêmio por Produtividade foi altamente positivo. A feliz escolha dos indicadores pactuados levou a instituição a priorizar o que já era importante. O Acordo estabeleceu parâmetros para aquilo que já era legitimado como prioridade institucional. Os dois indicadores pactuados provocaram uma série de mudanças nos processos organizacionais e geraram efeitos concretos de melhoria para os cidadãos.

O estudo de Assis (2012) oferece uma importante contribuição para identificar os efeitos na organização e na sua cultura. As avaliações qualitativas apresentadas identificam pontos positivos associados ao PRV, tais como o aumento da produtividade do Corpo de Bombeiros, mas também pontos negativos, tais como a queda na qualidade dos inquéritos policiais.

Num outro estudo, Oshiro & Scorzafave (2011) propuseram medir o impacto do PRV implantado em 2008 pelo governo de São Paulo na Secretaria de Educação sobre a proficiência dos alunos das escolas estaduais paulistas. Este é o mesmo programa que esta pesquisa utiliza para avaliar seus impactos no desempenho dos servidores públicos.

Para avaliar o programa os autores compararam o resultado dos alunos antes da implantação do programa com o resultado posterior. Para isto foram utilizados os dados da Prova Brasil edição 2007 e 2009. O grupo de tratamento foram as escolas estaduais paulistas e o grupo de controle foi as escolas dos demais estados e as escolas municipais paulistas. A variável de interesse foi a proficiência em matemática e português dos alunos da quarta e oitava séries apuradas na Prova Brasil.

No estudo foi utilizada a metodologia de *matching* (pareamento) que consiste em encontrar um grupo de controle o mais parecido possível ao grupo de tratamento de modo que os resultados desses possam ser comparados com seus semelhantes. Para resolver o problema da multidimensionalidade, dado que o número de variáveis observadas é muito grande, aplicou-se o método *propensity score* (escore de propensão), que consiste em estimar uma probabilidade do indivíduo pertencer ao grupo de tratamento, dadas suas características observáveis. Em seguida foi aplicado o método de Diferenças em Diferenças (DD), que consiste na comparação da diferença dos resultados do grupo de tratamento e controle ao longo do tempo (diferença no tempo), para depois estimar os resultados pelo método dos mínimos quadrados ordinários.

Os resultados notados foram positivos apenas para a 4ª série. O efeito estimado em matemática foi cerca de 6,4 pontos na escala SAEB (de 0 a 500) e em português foi de 3,7 pontos. Com relação à oitava série os coeficientes de interesse não foram significativos estatisticamente tanto em matemática quanto em português.

Mesmo não demonstrando no estudo que foram excluídos do grupo de controle os municípios e estados que podem ter PRV neste mesmo período, considera-se que a contribuição quanto à metodologia é relevante para os propósitos desta pesquisa.

Glewwe et al. (2010) relatou os resultados de uma pesquisa no Quênia que forneceu incentivos individuais e em grupo aos professores da 4ª e da 8ª séries de escolas selecionadas aleatoriamente para participar do experimento. Com base em resultados de testes identificou-se que o desempenho dos alunos aumentou durante o programa. Porém, isto foi considerado um efeito de curto prazo, pois os alunos não conseguiram reter os ganhos após o fim planejado do programa, 2 anos depois do seu início. Os resultados foram interpretados como sendo consistente com os professores em despende esforço para melhorar o desempenho dos alunos no curto prazo, mas não em relação à aprendizagem de longo prazo.

Muralidharan e Sundararaman (2009) investigaram o efeito dos incentivos individuais e coletivos em 300 escolas em Andhra Pradesh, Índia e descobriu que ambos os incentivos, grupos e individuais, aumentaram o desempenho do aluno por 0,12 desvios-padrão no desempenho de linguagem e 0,16 desvios-padrão em matemática, no primeiro ano, ambos igualmente bem sucedida. No segundo ano, no entanto, os incentivos individuais se mostram mais eficazes com um efeito médio de 0,27 desvios-padrão através de matemática e desempenho de linguagem, enquanto os incentivos do grupo teve um efeito médio de 0,16 desvios-padrão.

Fryer (2011) avaliou o efeito do programa implantado pela cidade de Nova Iorque, que premiava os professores de escolas selecionadas aleatoriamente para participar do

programa. O prêmio de 3 mil dólares por professor era pago à escola e esta, através do seu conselho, poderia decidir se o bônus seria pago igualmente a todos professores ou aos que mais se destacaram, sendo que a escolha predominante foi a primeira opção. Os resultados dos testes no desempenho dos alunos se apresentaram negativos em linguagem e em matemática, tanto para o ensino fundamental quanto para o ensino médio. Para o autor, não há evidências de que os incentivos melhoraram o desempenho dos alunos.

Para Fryer, o fato de Nova Iorque possuir, em média, 60 professores por escola, bem superior à média de 12 professores por escola no Quênia e cerca de 5 professores nas escolas avaliadas na Índia, favorecem a dispersão do comprometimento individual com o aprendizado do aluno, criando um ambiente favorável para o aparecimento da figura do carona (*free-riding*).

Baseado nos estudos apresentados nesta revisão bibliográfica é possível notar que os autores chegaram a diferentes conclusões sobre os efeitos do PRV. Perry *et al.* (2009) e Weibel *et al.* (2009) afirmaram que a eficácia dos programas são limitados e deixam de cumprir suas promessas. Hasnain, Manning e Pierskalla (2012) encontraram efeitos positivos em cerca de 60% dos 153 PRV avaliados, sendo que o restante foi considerado neutro ou negativo.

Por sua vez, Oshiro & Scorzafave (2011) encontraram efeitos positivos no ensino de 4ª série e efeitos neutros no ensino da 8ª série. Assis (2012) encontra efeitos qualitativos positivos e negativos nas instituições ligadas a Segurança Pública de Minas Gerais. Bowman (2010) faz constatações negativas, mas também positivas vindas do governo americano. O mesmo autor realça a necessidade de criação de um ambiente organizacional receptivo antes de se implantar qualquer programa. Marsden, por sua vez, diz que mesmo que tais promessas não sejam cumpridas, os ganhos não são desprezíveis, pois há uma revisão dos valores e propósitos da organização. Hood (2006) realça a necessidade de se criar mecanismos *anti gaming*, a fim de evitar a corrosão dos propósitos do PRV.

Pode-se notar pelos estudos analisados que as avaliações qualitativas dos PRV tendem a identificar pontos positivos com razoável facilidade. Porém, as avaliações quantitativas tendem a encontrar resultados mais modestos, quando não negativos. Neste ponto, recupera-se a afirmação de Marsden (2009), segundo o qual a forma de avaliar pode não estar medindo adequadamente o desempenho do servidor.

## 4.2. Avaliação do PRV nas empresas privadas

Os modelos aplicados à iniciativa privada, mesmo que não estejam associados diretamente ao objeto de pesquisa, não podem ser desconsiderados, pois grande parte das estratégias utilizadas na gestão dos recursos humanos do setor público se baseou em modelos inicialmente desenvolvidos para as empresas. O próprio PRV é um exemplo disso.

Jenkins *et al.* (1998) avaliaram 39 estudos que intencionaram associar os incentivos financeiros à qualidade da produção e à produtividade. A metodologia utilizada buscou correlacionar a qualidade e a produtividade em análises separadas. Os resultados demonstraram que há correlação – 0,34 – positiva com a produtividade, porém a correlação com a qualidade – 0,08 - se mostrou não estatisticamente significativamente diferente de zero.

Outro estudo encontrado foi de Lazear (2000), que tenta mensurar o impacto dos incentivos financeiros na produtividade pelo programa de remuneração variável durante 19 meses, entre 1994 e 1995, em 3.000 empregados de uma empresa fabricante de autopeças americana (Safelite).

A adesão ao programa era opção do empregado, que podia se manter na tradicional forma de pagamento por hora trabalhada (grupo de controle) ou aderir a remuneração por peça produzida, com garantia de remuneração mínima equivalente a hora trabalhada (grupo de tratamento).

Foi aplicada a análise de regressão aos grupos, com utilização de *dummies* para controlar a adesão ou não ao programa e também para controlar o tempo em que o empregado ficou no programa (demissão).

Contatou-se aumento do *turnover*, dado que os menos esforçados e menos habilidosos foram compelidos a deixar a empresa, dando lugar a trabalhadores mais competitivos.

Os lucros da empresa foram melhores, mas não na mesma proporção dos ganhos de produtividade. Algumas medidas tiveram que ser adotadas pela empresa para conter o retrabalho, pois o esforço para aumentar a produtividade implicou em aumento de defeitos na produção.

Os resultados demonstram que os incentivos incrementaram a produção individual entre 22% e 44%.

Noutro estudo, Cadsby, Son e Tapon (2007) avaliaram os resultados comparados do pagamento por desempenho versus salário fixo em ambiente de laboratório.

Foram selecionados 115 estudantes para participar da pesquisa, aos quais eram solicitados que solucionassem anagramas aplicados em 8 fases. Cada fase correspondia a um anagrama e era aplicada em salas individuais sem que os participantes soubessem dos resultados. Os mesmos eram pagos em cada fase. Em algumas etapas o participante poderia escolher entre PRV ou remuneração fixa. Outras etapas somente remuneração fixa e outras somente PRV. Os participantes não foram substituídos durante o experimento. A remuneração não previa valor mínimo, portanto aqueles que optassem por variável e tivessem fraco desempenho saíam perdendo.

Constatou-se que aqueles mais adversos ao risco optaram por fixo, enquanto os mais capazes optaram por variável. Houve aumento de participantes que inicialmente optaram por fixo e depois migraram para remuneração variável, demonstrando que a remuneração variável incentivou os participantes a aumentarem a produtividade.

Concluiu-se que a remuneração variável atrai os participantes de alta qualidade e motiva os demais a serem mais esforçados. Também se notou que a produtividade aumenta mais por conta da seleção dos melhores do que pelo incentivo pago.

Algumas limitações do método foram identificadas, sendo a principal a ausência de risco de perder o emprego, dado que os participantes não eram ameaçados de perder a vaga. Por outro lado, isto se aproxima das condições do servidor público, que, normalmente, possui estabilidade no emprego.

Mesmo com toda dificuldade em avaliar os resultados dos PRV, o mesmo não pode ser considerado um modismo. A disseminação desses programas pelos países e, dentro desses, pelos diversos níveis de governo é uma clara indicação de sua aceitação como instrumento de gestão de recursos humanos que proporciona resultados positivos.

Avaliar os benefícios da remuneração variável exigirá estudos mais aprofundados e o acompanhamento próximo, por dentro das instituições e dos seus programas. Há, notadamente, carência de dados que favoreçam a realização de estudos técnicos, embasados em modelos probabilísticos de análise.

Pelos estudos analisados, a avaliação dos impactos causados pelos programas de remuneração variável não é uma tarefa simples de se executar. A quantidade de variáveis que podem afetar a implantação ainda não foi claramente identificada, muito menos a importância que cada uma delas tem no processo. Os recursos financeiros, cultura organizacional, tempo de implantação, qualidade das metas, são alguns dos componentes que afetam de forma

diferenciada cada instituição. Constatam-se nos trabalhos analisados apenas inferências sobre a relevância que cada uma tem, mas não é raro encontrar opiniões de autores distintos que se contradizem.

Se há contradição nos impactos para o setor privado, que tem como principal objetivo a obtenção de lucro, para o setor público esta contradição é ainda maior, pois o desempenho é medido pelos benefícios proporcionados à população, tais como saúde e educação. Porém, dentro destas demandas há diferentes desempenhos. A educação, por exemplo, vai da pré-escola até o ensino superior, atendendo desde os indivíduos analfabetos até os pesquisadores pós-doutores. Cada um destes grupos possuem medidas próprias de desempenho.

Diante disso, apesar das experiências no setor privado servir de referência, a transposição deles o setor público deve ser vista com cautela, dado a grande diferença entre o que é resultado para empresas e para governos.

### **4.3. A experiência do setor público brasileiro com PRV**

No setor público brasileiro, algumas experiências têm sido tentadas. No Estado de Minas Gerais pode se encontrar o programa mais amplo e mais avançado, se comparado aos demais Estados, e há alguns estudos acadêmicos relatando as experiências. Assis & Neto (2010) divulgaram um estudo que constata a participação de 94% dos órgãos e entidades daquele governo estadual nos programas de remuneração variável, sendo que a participação tem sido crescente desde o início dos programas em 2004.

O programa mineiro partiu do ponto recomendado por Pacheco (2009), a qual considera que tais inovações exigem o recurso prévio do planejamento estratégico. Assis & Neto (2010) afirmam que todo planejamento governamental mineiro se baseia no Plano Mineiro de Desenvolvimento Integrado – PDMI, que foi recentemente revisado com metas de resultado e objetivos estratégicos até 2023. As metas são pactuadas com os órgãos/entidades por meio de acordo de resultados. Os resultados, por sua vez, estão vinculados a metas e cronogramas dos 57 projetos que constam no PMDI.

Em 2007, de acordo com os autores, o programa sofreu uma revisão geral e o acordo de resultados passou a ser estabelecido em duas etapas. Na primeira, o foco foi sintetizar os



objetivos gerais da política e da estratégia por meio dos resultados finalísticos, de taxas de execução de projetos estruturadores, de metas da agenda setorial, de indicadores e de racionalização dos gastos. Os resultados desejados foram pactuados entre o governador do Estado e o dirigente máximo de cada órgão ou entidade. Na segunda etapa, o dirigente fez um desdobramento desses resultados com os servidores das secretarias e das entidades vinculadas.

Com a identificação de metas mais próximas ao dia a dia do servidor, ficou mais evidente a contribuição de cada equipe para a estratégia. No novo modelo, a premiação passou a ser proporcional à nota alcançada no acordo de resultados. Para cada meta pactuada na segunda etapa, foi atribuída uma nota, segundo o que foi realizado, ou seja, cada uma das 4.739 equipes que pactuaram metas específicas obtém uma nota pelo seu desempenho.

Esta estratégia vem ao encontro da constatação de Pacheco (2009), que afirma que os resultados são melhores quando as metas são negociadas (e não impostas unilateralmente) e quando as flexibilidades são concedidas tendo metas a alcançar como contrapartida.

Neto & Assis (2010) constataram também que esta nova lógica de premiação por produtividade proporcionou expansão da assinatura de acordos de resultado relacionada à facilidade de viabilizar o pagamento do prêmio, em caso de resultados positivos. Houve aumento do interesse de boa parte das instituições em fazer o processo de planejamento/alinhamento estratégico prévio necessário à assinatura do acordo de resultados.

Outra iniciativa que constou na reforma do governo mineiro foi a desvinculação entre o pagamento dos prêmios e as economias proporcionadas pelos resultados. As despesas com a remuneração variável são previstas no orçamento do Estado. Isto é especialmente benéfico para as áreas que não têm a economia de recursos como meta, tais como saúde e educação.

A contribuição deste estudo para ajudar a compreender os efeitos dos PRV é significativa, porém, o mesmo não auxilia no desenvolvimento de uma metodologia que mensure adequadamente os impactos dos PRV.

Outros Estados brasileiros possuem programas de remuneração variável em diferentes estágios de implantação e amplitude, tais como Pernambuco, Espírito Santo, Paraíba e Piauí. Porém, constata-se uma carência de estudos acadêmicos que avaliem tais iniciativas. Ao nível de governo local, sabe-se que o município de São Paulo e de Curitiba também possuem PRV. Provavelmente há outros municípios com iniciativas neste sentido entre os 5.565 existentes. O governo do Estado de São Paulo, o maior Estado da federação em termos econômicos, tem iniciativas que podem ser consideradas como tímidas, pois estão concentradas em alguns organismos públicos da administração indireta, tais como na Cia do Metropolitano

de São Paulo (Metrô) e na Cia de Saneamento Básico de São Paulo (Sabesp). O programa mais amplo deste Estado foi encontrado na Secretaria de Educação, no qual esta pesquisa tem interesse e pode ser entendido de forma detalhada na próxima seção deste trabalho.

## 5. O Programa Paulista dos Bônus na Educação

Por ser objeto de avaliação desta pesquisa, descreve-se o programa de remuneração variável na educação paulista, aplicável aos servidores da rede estadual, para que se possa conhecer suas características e objetivos e, com isto, dar suporte ao aprofundamento desta pesquisa em busca dos seus impactos.

As regras do sistema de remuneração foram lançadas pelo Programa de Qualidade da Escola (PQE) em maio de 2008 e o projeto de lei referente ao programa foi aprovado em dezembro do mesmo ano.

Em março de 2009, o governo estadual remunerou cerca de 196 mil servidores baseado em resultados de 2008.

A Secretaria de Educação do Estado de São Paulo estipula metas para a série final de cada ciclo de ensino - quarta série (quinto ano) e oitava série (nono ano) do ensino fundamental e terceiro ano do ensino médio - e remunera todos os professores do ciclo avaliado e os demais funcionários da escola, ou seja, o incentivo é dado ao grupo pertencente à escola avaliada. Pelas regras do programa tenta-se eliminar o problema do absenteísmo, pois funcionários com mais de 1/3 de faltas não têm direito ao bônus, porém, esta regra já era adotada antes de 2008 e vigente há mais de 10 anos para pagamento de prêmio ligado à assiduidade, mas não ao desempenho.

As metas são estabelecidas com base no Índice de Desenvolvimento da Educação do Estado de São Paulo (IDESP) para cada etapa escolar, que por sua vez, é composto por dois outros índices:

$$IDESP_s = ID_s \times IF_s \quad (1)$$

O ID é o índice de Desempenho e IF é o índice de Fluxo e serão detalhados mais a frente. A ideia dessa composição é que uma escola pode ser considerada boa quando a maior parte dos alunos desenvolve as competências requeridas para a sua série num período de tempo ideal - o ano letivo.

O ID é baseado nas notas do SARESP, exame de avaliação externa paulista que é aplicado para a segunda, quarta, sexta e oitava séries (ou terceiro, quinto, sétimo e nono ano) do ensino fundamental, além da terceira série do ensino médio nas escolas administradas pelo governo estadual. O exame avalia a proficiência dos alunos de acordo com as exigências do

Sistema Nacional de Avaliação da Educação Básica, que define o aprendizado desejado para o aluno no final de cada ciclo. São aplicadas provas de Língua Portuguesa com Redação, Matemática e Ciências Humanas (Geografia e História). Para o pagamento de bônus, considera-se apenas a proficiência em Matemática e Português.

O ID representa a quantidade relativa de alunos em cada nível de proficiência. São quatro níveis:

(i) Abaixo do básico, (ii) Básico, (iii) Adequado e (iv) Avançado. Os valores de referência do SARESP para a definição dos níveis de desempenho encontram-se na tabela a seguir:

Tabela 2- Valores de referência para níveis de proficiência

	Quarta série		Oitava série	
	Português	Matemática	Português	Matemática
Abaixo do básico	<150	<175	<200	<225
Básico	entre 150 e 200	entre 175 e 225	entre 200 e 275	entre 225 e 300
Adequado	entre 200 e 250	entre 225 e 275	entre 275 e 325	entre 300 e 350
Avançado	>250	>275	>325	>350

Fonte: Secretaria de Educação do Estado de São Paulo

Assim, o ID para cada disciplina (j) e para cada etapa de escolarização (s) é dado por:

$$ID_{js} = \left(1 - \frac{def_{js}}{3}\right) \times 10 \quad (2)$$

Na fórmula, def é um indicador que mede a defasagem em relação às expectativas de aprendizado.

$$def_{js} = \frac{(3xAB) + (2xB) + (1xAd) + (0xAv)}{100} \quad (3)$$

Em que AB, B, Ad e Av representam respectivamente a proporção de alunos nos níveis abaixo do básico, básico, adequado e avançado em relação ao total de alunos avaliados.

Quando a escola apresenta grande defasagem, ou seja, nível de proficiência baixo, o índice de desempenho é baixo. O ID varia de zero a dez e a média aritmética entre os índices de Língua Portuguesa (LP) e Matemática (Mat) representa o ID para cada etapa de escolarização.

$$ID_s = \frac{ID_{lp} + ID_{Mat}}{2} \quad (4)$$

O indicador de Fluxo (IF) é dado pela taxa média de aprovação em cada etapa de escolarização e é dado por:

$$IF_s = \frac{A_i + A_{i+1} + \dots + A_n}{n} = \frac{\sum_{i=1}^n A_i}{n} \quad (5)$$

Em que A é a taxa de aprovação na série i e n é o número de séries<sup>2</sup> da etapa de escolarização considerada.

$$A_i = \frac{\text{número de alunos aprovados}}{\text{número total de alunos matriculados}} \quad (6)$$

A partir dos dados do IDESP calcula-se a parcela que a escola cumpriu da meta. O Índice de Cumprimento (IC) em determinado ano (t) para cada etapa escolar (s) é dado pela seguinte fórmula:

$$IC_{s,t} = \left( \frac{IDESP_{s,t} - IDESP_{s,t-1}}{META_{s,t} - IDESP_{s,t-1}} \right) \quad (7)$$

O IC é limitado no intervalo de 0% a 120%. Para valores negativos, considera-se o piso (zero). Para valores superiores a 120%, considera-se o cumprimento de 120%.

Um adicional por qualidade passou a ser considerado a partir de 2009. Este índice reflete a posição da escola em relação à média das outras escolas (IDESP global) e à sua meta de longo prazo (META 2030). O Índice de Qualidade (IQ) do ano t para a série s é calculado pela seguinte fórmula:

$$IQ_{s,t} = \left( \frac{IDESP_{s,t} - IDESP_{(global)s,t}}{META_{s,2030} - IDESP_{(global)s,t}} \right) \quad (8)$$

Para efeito de pagamento, considera-se o indicador (IQ+IC), que está limitado entre 0 e 120%.

As metas de longo prazo do IDESP para cada série são iguais para todas as escolas e foram estabelecidas levando-se em consideração níveis de aprendizado comparáveis aos de países da Organização para a Cooperação e o Desenvolvimento Econômico (OCDE). As metas para 2030 para quarta e oitava séries do fundamental são, respectivamente, 7,0 e 6,0.

---

<sup>2</sup> Atualmente, com o ensino fundamental de nove anos, n é igual a 5 para as séries iniciais do EF e 4 para as séries finais

Em 2009, o IDESP médio para a quarta série da rede estadual foi de 3,86. Já para a oitava série foi de 2,84. As escolas têm, portanto, um grande desafio até 2030, pois, na média, estão muito abaixo do nível desejado.

Visando atingir este nível de pontuação, foram estabelecidas metas anuais (que entram no cálculo do IC e IQ) para cada escola de acordo com a distância da meta de longo prazo seguindo uma função de crescimento logística. Desta forma, escolas que partem de uma pontuação baixa têm que apresentar uma melhora percentual maior relativamente àquelas com maiores notas. Embora a variação seja maior, é razoável acreditar que em escolas piores, menores esforços possam gerar maiores mudanças de notas. Portanto, embora escolas melhores tenham que apresentar uma menor variação de notas, como já se encontram em um patamar elevado, os esforços necessários podem ser maiores e o cumprimento da meta pode ser mais difícil. Nesse sentido, torna-se interessante verificar se o desenho do programa é capaz de gerar incentivos tanto para escolas mais fracas quanto para escolas mais fortes, uma vez que elas devem ser muito diferentes em vários aspectos.

Com relação ao pagamento do bônus, o valor varia de acordo com a parcela da meta cumprida e é paga aos professores do ciclo que cumpriu a meta e para funcionários com nível de absenteísmo abaixo de 1/3. Em escolas que cumprem 100% da meta, por exemplo, os funcionários recebem o equivalente a 2,4 salários. Em escolas que superam a meta prevista, o bônus pode chegar a 2,9 salários. Escolas que não cumprem as metas, mas que nos dois anos anteriores foram classificadas entre as 10% melhores, também recebem uma remuneração. Este seria um incentivo para escolas boas, cujas metas são difíceis de cumprir, pois as notas já estão num patamar elevado. Funcionários das diretorias regionais recebem bônus pela média da região em que atuam.

Exceto pelo estudo de Oshiro & Scorzafave (2011), mencionado no capítulo da revisão bibliográfica deste trabalho, que avaliou o efeito no desempenho do aluno, não se identificou estudos que pudessem avaliar a melhora do desempenho dos profissionais incumbidos diretamente do aprendizado do aluno, ou seja, os diretores das escolas e seus professores. Este é o foco do estudo empírico integrante desta pesquisa.

## 6. Hipótese de pesquisa

A revisão teórica indicou que os PRVs intencionam, entre outros objetivos, reduzir a assimetria de informação entre agente e principal, amenizando os problemas de agência. No caso específico do serviço público, o PRV pode auxiliar no alinhamento dos interesses dos governantes e gestores públicos (principal) com os interesses dos servidores públicos. Neste mesmo sentido, a teoria dos incentivos demonstra que incentivos financeiros baseados no desempenho do trabalhador contribui para a melhora do desempenho organizacional, desde que respeitadas as orientações para construção e implantação do modelo de remuneração.

Na revisão bibliográfica, os estudos visando avaliar os efeitos do pagamento por desempenho encontraram resultados diversos. Alguns apontaram no sentido de que os efeitos percebidos são limitados ou mesmo negativos (OCDE, 2005; Perry et al. 2009; Marsden, 2009). Para estes, se há vantagens na adoção destes programas são de caráter qualitativo, restritas ao aprendizado organizacional, o que vai ao encontro do que afirma Souza et. all, 2005; Hipólito, 2008 e Wood & Picarelli, 1999 e 2004b, indicando que os objetivos da adoção do PRV devem ser alinhados com a estratégia organizacional, de forma que as metas associadas ao desempenho do trabalhador, sejam elas de curto, médio ou longo prazo, também estejam associadas às metas esperadas pela organização.

Também se notou na revisão bibliográfica que no Brasil há algumas iniciativas de implantar PRV, quase todas de forma setORIZADA, sendo que o único caso abrangente para um nível de governo é o encontrado no Estado de Minas Gerais, porém, não se identificou estudos que avaliassem seus impactos quantitativamente e de forma ampla.

Os estudos que avaliaram simultaneamente diversos programas encontraram resultados positivos em sua maioria (Bowman, 2010; Weibel et al., 2009; Hasnain, Manning, Perskalla, 2012). Mesmo não havendo consenso dos resultados, o que se nota é uma disseminação dos programas de remuneração variável implantados no setor público. As possíveis justificativas para tal insistência provêm da busca de novas formas de incentivo para o servidor produzir mais e melhor, e também dos indícios de efeitos positivos mais consistentes que os estudos encontrados podem não estar identificando.

Visando contribuir para o esclarecimento dos efeitos do PRV no desempenho do servidor, e diante das evidências positivas apresentadas pelos estudos quantitativos de que os programas de remuneração variável podem contribuir para reduzir a assimetria de informações,

bem como melhor o desempenho do servidor, define-se como hipótese dessa pesquisa o seguinte:

*A remuneração variável tem efeitos positivos no desempenho do servidor público, e, por consequência, na melhoria dos serviços oferecidos à população.*

Para verificação da hipótese de pesquisa será testado o programa implantado em 2008 pela Secretaria de Educação do Estado de São Paulo, que prevê o pagamento de bônus aos diretores de escola e professores.

Os testes intencionarão avaliar se o desempenho dos diretores de escola e dos professores sofreu alteração em virtude da implementação dos programas, bem como, de forma subsidiária, se o desempenho dos alunos também se alterou como possível reflexo das mudanças no desempenho dos dois primeiros.



## 7. 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 & 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 poderão 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 IBGE<sup>3</sup>. Seria possível realizar a pesquisa com todos os 645 municípios paulistas, porém a diversidade socioeconômica entre eles implicaria em adotar controles adicionais para neutralizar esse efeito. Ainda assim, outras e novas variáveis não observáveis presentes em municípios menores poderiam enviesar o resultado.

Desta forma, baseado nos dados do IBGE, encontrou-se 39 municípios paulistas com mais de 200 mil habitantes. Nestes municípios se localizam, baseado nos dados do INEP, 1/3 do total de escolas paulistas, o que equivale a mais de 2.700 escolas.

Neste grupo de municípios originalmente selecionados, foi realizada 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. A pesquisa foi realizada por telefone e e-mail, direcionada ao Chefe/encarregado do departamento de gestão de pessoas ou quem elabora a folha de pagamento das Secretarias de Educação.

A questão apresentada foi a seguinte:

---

<sup>3</sup> Instituto Brasileiro de Geografia e Estatística

*“Há alguma premiação, bônus ou qualquer outra forma de remuneração baseada em meta e indicadores de desempenho aos servidores da educação (diretores e professores)?”*

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.

Os 38 municípios com mais de 200 mil habitantes que se mantiveram no grupo de controle foram os seguintes:

Tabela 3- Municípios do grupo de controle

Código Município	Município	Quantidade de Escolas		
		Estaduais	Municipais	Totais
3501608	Americana	12	10	45
3503208	Araraquara	25	9	34
3505708	Barueri	0	37	37
3506003	Bauru	47	14	61
3509502	Campinas	151	38	189
3510609	Carapicuíba	52	5	57
3513009	Cotia	25	46	71
3513801	Diadema	65	5	70
3515004	Embu	40	15	55
3516200	Franca	50	14	64
3518701	Guarujá	27	19	46
3518800	Guarulhos	159	45	204
3520509	Indaiatuba	23	18	41
3522505	Itapevi	21	30	51
3523107	Itaquaquecetuba	39	28	67
3524402	Jacareí	30	11	41
3525904	Jundiaí	33	37	70
3526902	Limeira	25	30	55
3529005	Marília	33	19	52
3529401	Mauá	61	1	62
3530607	Mogi das Cruzes	68	32	100
3534401	Osasco	50	48	98
3538709	Piracicaba	50	19	69
3541000	Praia Grande	22	24	46
3541406	Presidente Prudente	29	22	51
3543402	Ribeirão Preto	56	26	82
3547809	Santo André	91	32	123
3548500	Santos	22	32	54

Código Município	Município	Quantidade de Escolas		
		Estaduais	Municipais	Totais
3548708	São Bernardo do Campo	68	60	128
3548906	São Carlos	29	8	37
3549805	São José do Rio Preto	31	29	60
3549904	São José dos Campos	73	35	108
3550308	São Paulo	20	35	55
3551009	São Vicente	78	32	110
3552205	Sorocaba	34	11	45
3552403	Sumaré	42	28	70
3552502	Suzano	26	17	43
3552809	Taboão da Serra	18	32	50
3554102	Taubaté	35	10	45
Total		1.748	953	2.701

Fonte: elaboração do autor com base nos dados do IBGE

O município de Barueri não possui nenhuma escola estadual, assim como Mauá possui apenas uma escola municipal. Em menor grau, nota-se também nos outros municípios desta base que não há uma proporção constante ou balanceada entre a quantidade de escolas estaduais e municipais. Esta situação é decorrente da política educacional adotada em cada município, que tem forte correlação com a sua renda.

Segundo os dados do IBGE, Barueri tem a maior renda do grupo. Mais do que o dobro do município em 2º lugar do grupo – Jundiaí - e decidiu assumir integralmente o ensino fundamental, enquanto Mauá, que está quase na base da pirâmide da renda, decidiu pela maior dependência possível do governo estadual. São Vicente, outro exemplo, possui renda inferior à de Mauá, mas optou por uma composição entre escolas estaduais e municipais mais balanceada. Os desequilíbrios notados não irão interferir nos resultados desta pesquisa, principalmente pelo uso da metodologia de pareamento por escore de propensão (*propensity score matching*), que será melhor esclarecido mais adiante no capítulo dos testes.

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 4- Dados disponíveis em 2007

	Estaduais (paulista)	Municipais	Total
Diretores	1.748	953	2.701
Professores	15.955	8.693	24.648

	Estaduais (paulista)	Municipais	Total
Alunos	257.042	127.461	384.503

Fonte: Elaboração do autor com base nos dados do INEP

Estes quantitativos se mantem nos anos seguintes, 2009 e 2011, pois as escolas que entraram em funcionamento após 2007 não fizeram parte das análises por não possuírem o status de “antes” do programa que pudessem ser comparadas com o “depois”. Essa afirmação é válida somente em relação às escolas que foram incluídas na amostra, que é a unidade básica desta pesquisa, critério este alinhado com o praticado no programa SAEB do Estado de São Paulo, que está sendo avaliado nesta pesquisa. Os professores ou alunos novos (ingressaram na amostra após 2007), foram considerados, pois têm seus desempenhos avaliados pelas suas médias obtidas por escola.

Uma vez identificado o universo a ser pesquisado, foi preciso definir como serão trabalhados os dados encontrados. O *status* 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  $E(Y_0^D - Y_0^A | D = 1)$  não é observável, dado que o mesmo indivíduo não tem como estar no grupo de tratamento e controle simultaneamente.

Neste caso, segundo Cameron & 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.

Esse pareamento perfeito somente é possível quando o número de variáveis do vetor  $X$  é pequeno. Por exemplo, se desejarmos comparar o desempenho de dois indivíduos selecionando somente aqueles da mesma cor, teremos facilidade em fazer o pareamento. Se adicionarmos mais um critério, p.e., renda familiar, teremos um pouco mais de dificuldade de encontrar indivíduos que atendam estas duas exigências, ser da mesma cor e ter renda familiar similar. Se adicionarmos um terceiro critério, p.e., nível de escolaridade, será muito mais difícil encontrar dois indivíduos que sejam da mesma cor, com a mesma renda e nível de escolaridade. Para  $n$  características observáveis o número de possíveis valores para o vetor  $X$  será  $2^n$ .

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,  $P(X)$ , do indivíduo receber tratamento, dadas suas covariadas<sup>4</sup>. Seja  $P(X)$  a probabilidade de um indivíduo  $i$  receber tratamento, definido como:

$$P(X) = \Pr(D_i = 1|X) \quad (9)$$

Pelo método do escore de propensão,  $X$  pode ser substituído por  $P(X)$ :

$$\begin{aligned} Y_{i1}, Y_{i0} &\perp D_i | X_i \\ Y_{i1}, Y_{i0} &\perp D_i | P(X_i) \end{aligned}$$

Onde  $\perp$  denota independência

O método gera uma área de suporte comum, que consiste na seleção apenas de indivíduos tratados com escore de propensão superior ao menor escore máximo entre tratados e controles ou inferior ao maior mínimo entre os dois. Os demais serão excluídos da análise e não comporão a área de suporte comum, conforme a seguinte restrição:

$$\begin{cases} \text{Min} (T_{max}; C_{max}) \\ \text{Max} (T_{min}; C_{min}) \end{cases}$$

Onde  $T$  é o grupo de tratados e  $C$  é o grupo de controle.

Isto pode ser melhor visualizado na figura seguinte:

---

<sup>4</sup> Conjunto de características observáveis nos diretores, professores ou alunos e que serão controladas no experimento.

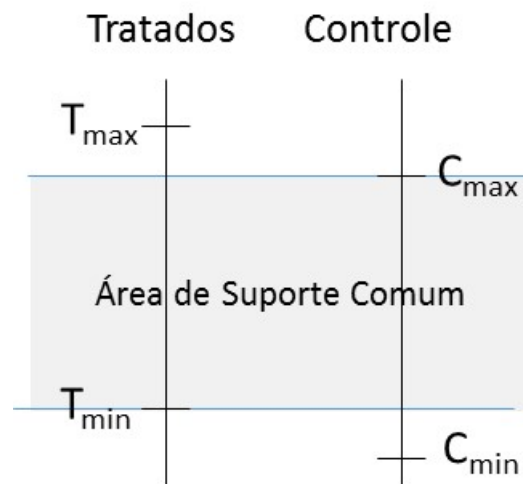


Figura 2- Área suporte comum

Fonte: Elaboração do autor com base nos dados do INEP

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 controle 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 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 & 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). Foi considerado  $Y^A$  como o resultado antes da adoção do programa e  $Y^D$  o resultado depois do programa. Neste caso o DD identifica o efeito da intervenção sobre o resultado de interesse  $Y$ , que pode ser assim escrito:

$$DD = E(Y_1^D - Y_1^A | D = 1) - E(Y_0^D - Y_0^A | D = 1) \quad (10)$$

O efeito médio no grupo de tratamento e controle antes e depois do programa pode ser melhor visualizado na figura a seguir:

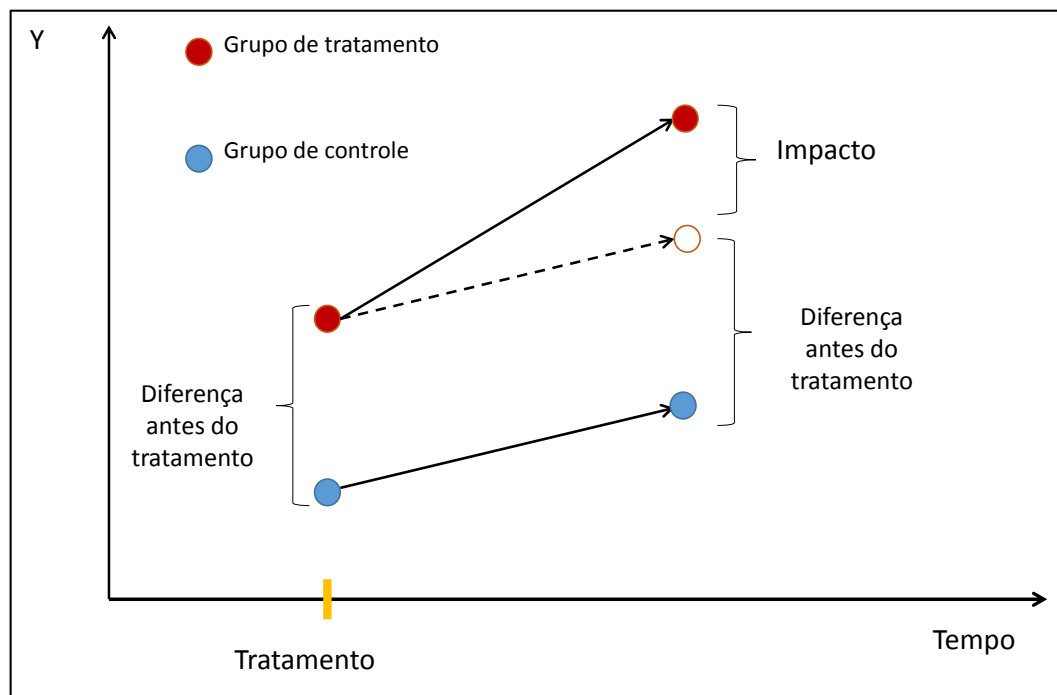


Figura 3- Efeito Médio nos tratados e não tratados  
Fonte: Menezes Filho (2012)

Para que seja possível calcular o estimador de DD é necessário admitir a hipótese de tendência comum que é dada por:

$$E(Y_0^D - Y_0^A | D = 1) = E(Y_0^D - Y_0^A | D = 0).$$

Neste caso o impacto médio é dado por:

$$DD = E(Y_1^D - Y_1^A | D = 1) - E(Y_0^D - Y_0^A | D = 0) \quad (11)$$

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:

$$DD = [E(Y_1^D | D = 1) - E(Y_0^D | D = 0)] - [E(Y_1^A | D = 1) - E(Y_0^A | D = 0)] \quad (12)$$

Ou

$$\beta_3 = \tau_{ATT}^D - \tau_{ATT}^A \quad (13)$$

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

$$Y_{it} = \beta_0 + \beta_1 D_i + \beta_2 t + \beta_3 D_i t + \varepsilon_{it} \quad (14)$$

Onde:

$D_i = 1$ , se recebeu o tratamento e 0 caso contrário;

$t = 1$ , se depois do tratamento e 0 se antes

Ajustando o modelo ao contexto do presente trabalho, o mesmo pode ser reapresentado da seguinte forma:

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

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 ou a 2011 (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 ou 2011 nas escolas estaduais paulistas em relação às as escolas municipais.

$X$  = vetor de covariadas.

$\varepsilon_{it}$  = Resíduo da regressão

## 8. Dados

Os bônus pagos aos servidores paulistas estão associados ao desempenho no Saesp, conforme detalhado anteriormente neste trabalho, porém, esta avaliação é restrita aos servidores da educação do Estado de São Paulo, impedindo com isto que se possa utilizar os municípios como contrafactual. Esta carência pode ser suprida pelo Sistema de Avaliação da Educação Básica – Saeb/Prova Brasil, que teve aplicação antes e depois da implantação do programa paulista para o Estado de São Paulo e também para todos os demais Estados Brasileiros simultaneamente, incluindo as escolas federais e municipais.

O Sistema de Avaliação da Educação Básica – Saeb/Prova Brasil é uma avaliação externa aplicada desde 1990, a cada dois anos, pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira – INEP.

Em 2005, o SAEB foi reestruturado, passando a Prova Brasil a avaliar de forma censitária as escolas que atendessem a critérios de quantidade mínima de estudantes na série avaliada, permitindo gerar resultados por escola.

A metodologia do Saeb/Prova Brasil baseia-se na aplicação de testes padronizados de Língua Portuguesa e Matemática e Questionários Socioeconômicos a estudantes de todas as turmas de 4ª e 8ª séries do ensino fundamental regular (em regime de 8 anos) e de 5º e 9º anos (em regime de 9 anos) de escolas públicas, desde que possuíssem pelo menos 20 alunos nas séries e anos avaliados.

Os estudantes também respondem a Questionários Socioeconômicos próprios, assim como os diretores e professores. No mesmo evento, os aplicadores da prova coletam dados sobre as condições da escola em formulário previamente elaborado pelo INEP.

No formato revisado em 2005 o Saeb realizou três edições, em 2007, 2009 e 2011. Na edição de 2007 foram avaliados 4.109.265 estudantes distribuídos nas escolas em todo território nacional.

A aplicação dos questionários visou o levantamento não apenas de informações sobre o desempenho dos alunos, como também sobre as condições intra e extraescolares que incidem sobre o processo de ensino e aprendizagem. Para isso, preencheram os questionários em 2007, além dos próprios alunos que responderam as provas, os seus 292.828 professores de língua portuguesa e matemática e os 47.596 diretores de suas unidades escolares.

Os questionários aplicados aos diretores e professores levantaram dados sobre a caracterização socioeconômica de seus respondentes, sua formação e experiência profissional,

as condições de trabalho em que se desenvolve o ambiente escolar, as práticas pedagógicas e de gestão que vêm sendo adotadas.

Com o questionário da escola foi possível saber sobre o estado de conservação e manutenção do seu prédio, bem como as condições das salas de aula, disponibilidade de computadores e se há medidas de segurança que protejam a escola e alunos internamente e no entorno.

O questionário aplicado aos alunos permitiu levantar dados sobre o seu perfil, nível de renda e estrutura familiar, bem como seus hábitos de leitura, sua trajetória escolar, as práticas escolares cotidianas e suas expectativas quanto à continuidade dos estudos.

Os testes de Matemática e de Língua Portuguesa, com foco em resolução de problemas e em leitura, respectivamente, foram elaborados a partir das Matrizes de Referência do Saeb e com base na Teoria de Resposta ao Item – TRI. A escala de desempenho é medida de 0 a 500.

Com base nas características apresentadas da Prova Brasil, considera-se que a mesma possui dados e informações suficientes para que se possa realizar os testes necessários para identificar os efeitos do PRV implantado em 2008 sobre desempenho dos diretores, professores e alunos da rede escolar estadual paulista.

## 9. 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 e 2011 (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 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 no capítulos seguintes desta pesquisa.

As características observáveis das escolas avaliadas também foram utilizadas para controlar os experimentos do professor e aluno. Para isto 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, dado que estes têm condições de interferir nas condições de uso das 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.

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 compõem um vetor próprio. Nesta seção

serão apresentados todos os controles comuns, deixando os específicos para serem apresentados na seção que irá avaliar cada um dos elementos pesquisados.

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

O controle do comportamento do PIB permite neutralizar os possíveis efeitos gerados por um melhor ou pior desempenho econômico do município que, por consequência, pode afetar o orçamento público destinado à educação municipal para custear salários, investimentos, manutenção etc. Todos estes itens podem interferir no desempenho de qualquer um dos elementos avaliados. Os valores individualizados do PIB per capita dos 38 municípios que fazem parte deste estudo se apresentaram da seguinte forma:

Tabela 5- PIB per capita municipal

Município	2007	2009	2011
Americana	26.558	29.850	32.224
Araraquara	17.191	20.898	24.837
Barueri	100.806	99.596	131.291
Bauru	15.233	18.906	23.036
Campinas	26.133	29.732	37.166
Carapicuíba	6.186	7.813	10.603
Cotia	24.704	29.814	34.242
Diadema	22.371	25.066	30.333
Embu	10.431	14.868	21.773
Franca	11.205	12.753	17.803
Guarujá	9.800	11.131	14.135
Guarulhos	22.202	24.994	35.248
Indaiatuba	23.427	28.055	29.070
Itapevi	10.007	17.671	31.467
Itaquaquecetuba	6.928	7.964	11.420
Jacareí	19.657	22.706	26.934
Jundiaí	40.704	47.396	58.352
Limeira	19.109	20.428	26.836
Marília	12.892	15.197	19.887
Mauá	13.395	15.750	18.125
Mogi das Cruzes	15.251	20.552	24.596
Osasco	35.218	43.994	58.822
Piracicaba	21.766	26.031	31.486
Praia Grande	9.184	11.143	14.904
Presidente Prudente	14.652	17.434	21.528
Ribeirão Preto	23.692	26.084	30.209

Município	2007	2009	2011
Santo André	20.044	21.844	26.035
Santos	47.108	54.055	75.195
São Bernardo do Campo	32.677	35.680	47.176
São Carlos	16.441	20.519	24.654
São José do Rio Preto	16.210	18.776	23.615
São José dos Campos	30.195	35.751	39.588
São Vicente	6.740	8.762	10.687
Sorocaba	21.447	24.272	30.166
Sumaré	24.791	28.629	32.326
Suzano	19.798	19.553	21.838
Taboão da Serra	16.338	17.751	20.863
Taubaté	25.610	30.446	34.680

Fonte: Instituto Brasileiro de Geografia e Estatística - IBGE

As outras 7 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 e 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. As 7 variáveis selecionadas do questionário do aluno foram as seguintes:

Tabela 6- Controles comuns ao diretor, professor e aluno

Proporção de Alunos com a Característica	2007		2009		2011	
	%	Nº Obs	%	Nº Obs	%	Nº Obs
Branco ou pardo	83,5	387.186	84,8	371.715	86,7	354.333
Idade adequada para a série	61,9	394.078	79,4	379.417	65,5	376.548
Tem 1 ou mais carros em casa	55,7	390.646	58,6	377.350	63,0	375.132
Tem computador com internet em casa	33,8	389.992	45,9	379.597	62,7	375.862
Tem pelo menos uma diarista em casa	10,8	378.145	8,9	380.367	8,8	375.785
Moram na mesma casa entre 3 e 6 pessoas, contando com o aluno	79,0	384.535	79,2	369.763	78,9	374.402
Não trabalha fora de casa	85,7	384.572	86,8	359.401	87,3	366.759

Fonte: Elaboração do autor com base nos dados do Inep

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 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 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. Extraíu-se 11 variáveis do questionário da escola como segue:

Tabela 7- Controles comuns ao professor e aluno – Condições de uso da escola

Proporção de Escolas com a Característica	2007		2009		2011	
	%	Nº Obs	%	Nº Obs	%	Nº Obs
Estado de conservação das paredes adequada	68,0	2.762	71,7	2.653	76,1	2.922
Salas de aula iluminadas (natural ou artificialmente)	95,6	2.757	96,2	2.605	97,3	2.644
Salas de aula arejadas	91,0	2.750	87,3	2.588	93,4	2.647
Bom estado de conservação da entrada do prédio	70,5	2.760	67,2	2.641	74,0	2.907
Bom estado de conservação das salas de aula	65,0	2.709	67,5	2.647	72,3	2.906
Bom estado de conservação dos banheiros	51,3	2.700	58,4	2.637	61,6	2.901
Bom estado de conservação da cozinha	69,7	2.698	74,2	2.621	75,8	2.890
Existe controle de entrada e saída de alunos na escola	98,6	2.769	98,2	2.609	99,0	2.906
Existe controle de entrada e saída de pessoas estranhas na escola	96,7	2.763	96,3	2.620	97,7	2.905
Escola adota medidas de segurança para proteger alunos nas imediações da escola	37,1	2.739	41,4	2.653	46,6	2.866
Quantidade de computadores com internet para uso dos professores superior a 1	64,9	2.773	94,5	2.889	91,9	2.885

Fonte: Elaboração do autor com base nos dados do Inep

Neste mesmo grupo de controles comuns<sup>5</sup> aos professores e alunos também foram adicionados os controles que refletem o 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.

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

Após a realização dos testes com escolas municipais, foi reaplicado a mesma avaliação, porém utilizando como base de comparação escolas estaduais de outros Estados ao invés de escolas municipais. Este procedimento visou testar a robustez dos resultados obtidos com as escolas municipais, avaliando seu grau de simetria.

## **9.1. Testes do efeito do programa sobre o desempenho dos diretores**

O desempenho dos diretores será avaliado segundo a adoção, por parte dos mesmos, de práticas gerenciais na gestão das escolas.

As práticas gerenciais utilizadas para avaliar o desempenho dos diretores podem ser identificadas a partir dos estudos apresentados por Sammons, Hillman & Mortimore (1995), Coleman, Collinge & 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:

- 1) Liderança profissional do diretor, que deve ter: (a) propósitos fortes e bem definidos; (b) envolver outras pessoas do seu grupo de auxiliares na tomada de decisão, em especial o professor, com um abordagem participativa e (c)

---

<sup>5</sup> Por recomendação da banca examinadora, foi sugerido controlar 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.



exercer sua autoridade profissional no processo de ensino (professor) e aprendizagem (aluno), incluindo a definição de projeto pedagógico, estratégias de ensino e monitoramento do progresso dos alunos;

2) Compartilhamento da visão e objetivos da escola. Escolas são mais efetivas quando os profissionais envolvidos constroem de forma consensuada os objetivos e valores e quando eles colocam isto em prática através de consistentes e colaborativas maneiras de trabalhar e de tomar decisão;

3) Formação de um ambiente apropriado para o ensino, com uma atmosfera organizada e atraente para o trabalho;

4) Concentração no ensino e na aprendizagem, com foco na sua qualidade, na maximização do tempo dedicado na aprendizagem e ênfase nos conteúdos pedagógicos;

5) Parceria escola-casa, envolvendo os pais dos alunos no processo de aprendizagem;

6) Construção de uma organização de aprendizagem, o que é válido para alunos, professores e diretores, pois estes também necessitam constantemente aprender e renovar conhecimentos.

De acordo com estas características, foram identificados os seguintes itens do questionário do diretor:

- Questão 22 – Se promoveu alguma atividade de formação continuada na escola? (Relacionadas à liderança do diretor e a ser uma organização de aprendizagem);
- Questão 24 – Neste ano, quantas vezes o conselho de escola se reuniu? (Relacionada à liderança do diretor, compartilhamento da visão e objetivos, parceria escola-casa);
- Questão 29 – Neste ano, quantas vezes se reuniram os conselhos de classe? (Relacionada à liderança do diretor, compartilhamento da visão e objetivos, formação de um ambiente apropriado para o ensino);
- Questão 30 – Se há projeto pedagógico da escola, como foi elaborado e por quem? (Relacionado à liderança do diretor, ao compartilhamento da visão e objetivos e à concentração no ensino e na aprendizagem);
- Questão 43 – Se a escola desenvolve, regularmente, algum programa de apoio ou reforço de aprendizagem para os alunos (monitoria, aula de reforço etc.)?

(Relacionado à liderança do diretor, a ser uma organização de aprendizagem e ao compartilhamento da visão e objetivos);

- Questão 94 - Quem escolheu os livros didáticos utilizados nesta escola?

(Relacionado à liderança do diretor e ao compartilhamento da visão e objetivos).

Para uma melhor compreensão do potencial destas variáveis, pode-se observar a proporção das respostas dos diretores nas questões identificadas em 2007, 2009 e 2011.

Tabela 8- Variáveis Dependentes do Diretor

Proporção de Diretores com a Característica	2007		2009		2011	
	%	Nº Obs	%	Nº Obs	%	Nº Obs
	22. Promove atividades de formação continuada	61,0	2.646	64,7	2.096	70,0
24. Reúne o conselho de escola mais de 3 vezes no ano	88,0	2.682	91,3	2.115	87,9	2.515
29. Reúne o conselho de classe mais de 3 vezes no ano	96,5	2.674	97,1	2.111	94,0	2.519
30. Trabalhou no projeto pedagógico da escola pessoalmente	65,5	2.667	60,7	2.092	70,2	2.485
43. Desenvolve reforço de aprendizagem para os alunos	96,0	2.642	96,4	2.129	98,2	2.524
94. Escolha conjunta dos livros didáticos	19,1	2.658	21,8	2.109	19,0	2.504

Fonte: Elaboração do autor com base nos dados do Inep

Nota: O número que identifica a questão equivale ao formulário do INEP

Nota-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.

Definida a composição da variável dependente, ou das variáveis dependentes, o próximo passo é aplicar o método de Pareamento (*matching*), que agrupa os indivíduos de acordo com suas características, representadas pelo vetor de covariadas.

Para dar maior robustez aos resultados foi aplicado o Pareamento pelo método do Escore de Propensão (*Propensity Score*) pareado pelo vizinho mais próximo (*nearest*

*neighbor*), sem reposição, conforme descrito no capítulo 7 deste trabalho, que trata do modelo empírico.

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.

A tabela a seguir apresenta o vetor de covariadas  $X_{ist}$  obtido a partir dos questionários dos diretores, com a proporção de Diretores por características controladas.

Tabela 9- Descrição das Covariadas do Diretor

Proporção de Diretores com a Característica	2007		2009		2011	
	%	Nº Obs	%	Nº Obs	%	Nº Obs
Masculino	15,1	2.686	15,6	2.134	15,4	2.525
Menor que 30 anos	1,6	2.695	2,2	2.117	4,6	2.441
Entre 30 e 50 anos	60,2	2.695	56,7	2.117	53,5	2.441
Maior que 50 anos	38,2	2.695	41,1	2.117	41,9	2.441
Branco ou pardo	92,2	2.659	94,7	2.120	93,6	2.504
Nível Escolaridade Ensino Superior	99,5	2.627	99,7	2.049	98,8	2.519
Ensino Superior em Pedagogia ou Normal Superior	55,7	2.627	53,2	2.049	62,1	2.519
Curso Superior em Instituição Pública	10,1	2.678	10,2	2.124	11,2	2.508
Curso Superior Presencial	97,7	2.689	97,4	2.130	97,3	2.519
Pós graduação de Longa Duração	64,7	2.670	70,0	2.123	71,8	2.514
Pós graduação na área da educação	65,8	2.650	67,5	2.096	69,6	2.486
Participação em atividades de formação continuada	95,1	2.676	89,6	2.130	92,2	2.514
Renda Individual Até R\$ 1.100 (2007)	0,3	2.677	0,2	2.115	2,3	1.910
Renda Individual Até R\$ 1.700 (2007)	6,1	2.677	4,2	2.115	1,7	1.910
Renda Individual Até R\$ 2.300 (2007)	30,0	2.677	18,7	2.115	7,6	1.910
Renda Individual Superior a R\$ 2.301 (2007)	63,6	2.677	76,8	2.115	88,4	1.910

Proporção de Diretores com a Característica	2007		2009		2011	
	%	Nº Obs	%	Nº Obs	%	Nº Obs
Renda Familiar Até R\$ 1.500 (2007)	0,9	2.688	1,0	2.122	2,3	1.925
Renda Familiar Até R\$ 2.300 (2007)	12,2	2.688	7,9	2.122	3,7	1.925
Renda Familiar Até R\$ 3.100 (2007)	29,1	2.688	22,8	2.122	12,8	1.925
Renda Familiar Superior a R\$ 3.101 (2007)	57,8	2.688	68,3	2.122	81,1	1.925
Não possui outras fontes de renda	69,6	2.692	71,5	2.132	72,8	2.508
Até 2 anos de experiência como diretor	11,0	2.687	14,2	2.128	12,8	2.498
Entre 2 e 4 anos de experiência como diretor	17,9	2.687	11,6	2.128	16,8	2.498
Mais de 5 anos de experiência como diretor	71,1	2.687	74,2	2.128	70,4	2.498
Até 2 anos de experiência como diretor nesta escola	33,1	2.673	33,3	2.129	31,9	2.511
Entre 2 e 4 anos de experiência como diretor nesta escola	32,3	2.673	21,7	2.129	24,0	2.511
Mais de 5 anos de experiência como diretor nesta escola	34,5	2.673	45,0	2.129	44,1	2.511
Trabalha tempo não integral	0,7	2.677	0,6	2.119	0,8	2.508
Assumiu cargo por meritocracia	57,5	2.664	61,1	2.111	56,1	2.502
25% dos professores com vínculo estável na escola	15,3	2.619	13,4	2.111	8,8	2.499
50% dos professores com vínculo estável na escola	17,1	2.619	16,5	2.111	16,2	2.499
75% dos professores com vínculo estável na escola	23,7	2.619	24,9	2.111	28,0	2.499
90% dos professores com vínculo estável na escola	23,9	2.619	26,0	2.111	27,8	2.499
100% dos professores com vínculo estável na escola	19,9	2.619	19,2	2.111	19,2	2.499
Renda do Diretor/Renda Familiar <87,5%	42,9	2.625	57,5	2.076	59,7	1.619

Fonte: Elaboração própria com base nos dados do INEP

Destaca-se a partir dos resultados tabelados que os diretores das escolas estaduais são predominantemente mulheres e acima dos 30 anos, sendo que cerca de 2/3 possuem curso de pós-graduação. Mais de 70% deles não possui outras fontes de renda e mais da metade são responsáveis por ao menos  $\frac{3}{4}$  da renda familiar. Também pode ser notado que mais da metade

possui ao menos 5 anos de experiência como diretor e quase a totalidade deles trabalha em tempo integral.

Dado o expressivo número de covariadas, numa primeira tentativa de estabelecer o critério de seleção das características a serem pareadas, realizou-se testes de médias para identificar aquelas com resultados não significantes, ou seja, foi utilizado para pareamento características que demonstraram ser estatisticamente semelhantes no grupo de tratamento e de controle.

Esse processo de seleção de variáveis foi realizado por tentativas sucessivas até se encontrar um conjunto composto somente por resultados não significantes, o que indicava que a média dos desvios de uma características observada no tratado e no controle eram iguais ou próximas.

Este processo se mostrou eficiente somente para alguns dos objetos testados, selecionando em média 12 variáveis. Porém, em alguns casos, houve dificuldade de se selecionar variáveis semelhantes e relacionadas ao resultado, como foi o caso do professor de matemática da 8ª série, que teve somente duas variáveis selecionadas. Isto prejudicou a qualidade do pareamento, pois com tão poucas exigências, muitos foram pareados sem considerar características importantes para este trabalho.

Poder-se-ia forçar o emparelhamento com variáveis consideradas associadas com o resultado, como por exemplo se o diretor tem dedicação exclusiva à escola, mesmo que se apresentassem com diferenças de médias. Porém, estabelecer critérios para identificar o que é associado ao resultado e seu grau de associação teria grande dose de subjetividade, que poderia gerar resultados questionáveis.

Outros meios de se fazer pareamento foram tentados, tais como comparando escolas que estivessem fisicamente próximas, para evitar comparar escolas de bairros de classe alta, mesmo que poucas, com escolas de periferia. Para isto poderia se utilizar a distância euclidiana calculada a partir das coordenadas geográficas que indicam a localização da escola. Porém, estes dados somente foram encontradas para cerca de 300 escolas, sendo que muitas dessas foram eliminadas da base por conta de dados incompletos.

Também se tentou fazer o pareamento usando a distância de Mahalanobis, que utiliza como medida a covariância entre as características observáveis. O par é gerado quando se encontra um indivíduo tratado e outro de controle com semelhantes covariâncias. Esse método é eficiente para pequenas amostras e poucas covariadas (menos de 8, segundo Rubin, 1979) devido ao problema da multidimensionalidade. Porém, este problema está presente nesta pesquisa, pois há, no caso do aluno, 67 covariadas e cerca de 1,4 milhão de indivíduos na

base de dados após sua depuração. Para estas 67 características observáveis o número de possíveis combinações para o vetor de covariadas será  $2^{67}$ , que é um número gigantesco de iterações, difícil de ser processado em computadores comuns.

Mesmo com esta dificuldade, nos testes realizados com os dados dos diretores e professores, que são menores do que os dos alunos, os resultados obtidos utilizando o método do *propensity score matching* com ou sem a técnica de Mahalanobis, dado que ambas podem ser combinadas, chegou-se a resultados muito semelhantes. Esta é uma prova robusta de que o pareamento utilizando Mahalanobis não produz diferenças significantes.

Outra tentativa foi feita com o uso do Código de Endereçamento Postal – CEP, porém a ausência desta informação para todas as escolas, inclusive para os demais Estados brasileiros, restringiu o seu uso. Como havia disponibilidade do CEP para as escolas paulistas, aplicou-se os testes com e sem esta informação como uma variável de controle. Novamente os resultados se apresentaram muito semelhantes. Além do mais, como o CEP é um dado de localização geográfica da escola, a determinação da distância máxima que poderia ser considerada adequada para formação de pares de escolas socioeconomicamente semelhante, seria arbitrária, o que enfraqueceria os resultados. Saber qual é a distância adequada exigiria uma pesquisa à parte.

Diante dessas restrições, optou-se por realizar os testes pelo método do *propensity score matching*, pareado pelo vizinho mais próximo, sem reposição, conforme descrito no capítulo 7 deste trabalho, que trata do modelo empírico.

Para grandes amostras, a literatura consultada indica fazer o pareamento com todas as características observáveis. Stuart (2010) recomenda o seguinte:

When matching using propensity scores there is little cost to including variables that are actually unassociated with treatment assignment, as they will be of little influence in the propensity score model. Including variables that are actually unassociated with the outcome can yield slight increases in variance. However, excluding a potentially important confounder can be very costly in terms of increased bias. Researchers should thus be liberal in terms of including variables that may be associated with treatment assignment and/or the outcomes. Some examples of matching have 50 or even 100 covariates included in the procedure (e.g., Rubin, 2001).

Acompanhando essa orientação, realizou-se teste de sensibilidade para confirmar a recomendação, notando que o acréscimo ou subtração de covariadas de forma individualizada exerce baixa influência no resultado. Por conta disso, optou-se por incluir no modelo todo o

conjunto de covariadas específicas do diretor e mais o vetor de covariadas comuns apresentadas no capítulo 9 deste trabalho.

Após o emparelhamento, do universo de 47.596 observações possíveis, 5.280 foram selecionadas pelo método do *propensity score*. São as escolas com características similares.

O resultado se apresentou com a seguinte composição:

Tabela 10- Área de Suporte Comum - Diretores.

Condição	Suporte Comum		Total
	Fora	Dentro	
Controles	0	1.742	1.742
Tratados	1.796	1.742	3.538
Total	1.796	3.484	5.280

Fonte: elaboração do autor com base nos dados do Inep

Mesmo tendo demonstrado que o pareamento foi feito com rigor, os resultados entre grupo de tratamento e controle no período de pré-tratamento (2007) se mostraram estatisticamente distintos e podem ser vistos na tabela seguinte:

Tabela 11- Resultado do Pareamento no Desempenho do Diretor - 2007

Variável	Amostra	Tratado	Controle	Diferença	D. Pad.	T-stat
22. Promove atividades de formação continuada	Antes	0,61214	0,64418	<b>-0,03204</b>	0,02187	-1,47
	Depois	0,62566	0,64418	<b>-0,01852</b>	0,02478	-0,75
24. Reúne o conselho de escola mais de 3 vezes no ano	Antes	0,93357	0,78836	<b>0,14521</b>	0,01419	10,23
	Depois	0,94048	0,78836	<b>0,15212</b>	0,01718	8,85
29. Reúne o conselho de classe mais de 3 vezes no ano	Antes	0,99286	0,91402	<b>0,07884</b>	0,00810	9,74
	Depois	0,99471	0,91402	<b>0,08069</b>	0,01054	7,66
30. Trabalhou no projeto pedagógico da escola pessoalmente	Antes	0,67429	0,62037	<b>0,05392</b>	0,02143	2,52
	Depois	0,67857	0,62037	<b>0,05820</b>	0,02451	2,37
43. Desenvolve reforço de aprendizagem para os alunos	Antes	0,97571	0,92328	<b>0,05243</b>	0,00906	5,79
	Depois	0,98016	0,92328	<b>0,05688</b>	0,01094	5,20
94. Escolha conjunta dos livros didáticos	Antes	0,15571	0,26058	<b>-0,10487</b>	0,01766	-5,94
	Depois	0,16667	0,26058	<b>-0,09392</b>	0,02096	-4,48

Fonte: Elaboração do autor com base nos dados do Inep

Em cada variável dependente da tabela se faz o teste de diferenças entre médias antes e depois do pareamento. As diferenças, destacadas em negrito, antes do pareamento são esperadas, pois indivíduos com características distintas são comparados, mas não deveria acontecer depois de pareados. Isto aponta para a existência de variáveis não observáveis que afetaram os resultados, ou seja, características do diretor que não puderam ser pareadas no experimento proporcionam a formação de pares não tão perfeitos.

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 vai avaliar 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 a partir do seguinte modelo:

$$Y_{ist} = \beta_0 + \beta_1 EESP_s + \beta_2 Ano_t + \beta_3 EESP_s * Ano_t + B_4 A_{vst} + B_5 X_{ist} + \varepsilon_{ist}$$

Onde:

$Y_{ist}$  é a variável dependente e identifica o efeito do modelo no desempenho do diretor, no qual  $i$ ,  $s$ ,  $t$  denotam o diretor, se é uma escola estadual ou municipal e se é referente a antes ou depois da implantação do programa;

$EESP_s = 1$ , se é diretor de uma escola estadual paulista (ou seja, se recebeu tratamento) e 0 se escola municipal paulista;

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

$A_{vst}$  é o vetor de covariadas das 7 características que controlam o entorno do aluno (ambiente socioeconômico);

$X_{ist}$  é o vetor de covariadas com as demais características do diretor;

$\varepsilon_{ist}$  é o resíduo da regressão

O  $\beta_3$  é o coeficiente calculado para a variável de interesse, estimada pela interação das duas variáveis anteriores: se é escola estadual paulista e se a observação se refere ao período posterior à implantação do programa. Esta variável mensurará o efeito do PRV sobre o desempenho dos diretores;



### 9.1.1. Resultados dos diretores

Com base na seleção obtida pelo pareamento por escore de propensão, nesta seção são apresentados os resultados da estimação pelo método de Diferenças em Diferenças. A variável de interesse é uma *dummy* resultante da interação do Ano (0 se antes ou 1 se depois do programa) com o Estado (dependência administrativa da escola: 0 se municipal ou 1 se estadual). O coeficiente desta variável indica o efeito no desempenho dos indivíduos tratados, ou seja, diretor participante do programa de remuneração variável das escolas estaduais paulistas após a implantação do programa em 2008. Este procedimento será repetido para cada uma das 6 questões, que são as variáveis dependentes, obtendo com isto 6 coeficientes distintos e não comparáveis entre si.

Vale lembrar que, além das covariadas apresentadas na seção anterior, específicas do diretor, controla-se o experimento também pelas covariadas do perfil do aluno e o PIB per capita, que são comuns aos três elementos desta pesquisa (diretor, professor e aluno)

Dado a quantidade de variáveis de controle, a tabela a seguir apresenta apenas os resultados da variável de interesse (Ano\*Estado) nas variáveis dependentes. As questões estão numeradas conforme constam no questionário que lhes deram origem. Os resultados completos estão disponíveis para consulta.

Tabela 12- Coeficientes estimados - Diretores

Variáveis Dependentes	Coeficientes
22. Promove atividades de formação continuada	0,0134 (0,0317)
24. Reúne o conselho de escola mais de 3 vezes no ano	-0,0858 *** (0,0222)
29. Reúne o conselho de classe mais de 3 vezes no ano	0,0135 (0,0149)
30. Trabalhou no projeto pedagógico da escola pessoalmente	-0,1009 *** (0,0328)
43. Desenvolve reforço de aprendizagem para os alunos	-0,0267 * (0,0137)
94. Escolha conjunta dos livros didáticos	0,0767 *** (0,0282)

Fonte: Elaboração do autor com base nos dados do INEP

(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, conforme demonstrado anteriormente na tabela 9.4, há uma quantidade expressiva de diretores em 2007 que não participava da elaboração do projeto pedagógico, mas poderiam fazê-lo nos anos seguintes. Porém, isto não aconteceu, ao contrário, houve redução de 10,09% na participação dos diretores nesta atividade.

O efeito positivo significativo é notado apenas na participação do diretor da escolha dos livros didáticos (questão 94). Este resultado é relevante, pois, assim como constatado na questão 30, nesta questão também há 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 7,67% na participação dos diretores na escolha dos livros didáticos.

## **9.2. Testes do efeito do programa sobre o desempenho dos professores**

Dado que o professor não possui avaliação específica, 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, na 4ª e na 8ª série. As avaliações foram feitas de forma individualizada por série e disciplina.

Identificou-se no questionário aplicado ao professor um grupo de 9 questões<sup>6</sup> sobre as atividades visando a aplicação do conteúdo programático de português e outro grupo de 10 questões sobre o conteúdo de matemática. Em ambas 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, Engenart, Furst Hill & Kathwohl, 1956)*”.

---

<sup>6</sup> Originalmente eram 10 questões, porém em 2011 a questão 27 foi excluída, impedindo sua avaliação comparativa com 2007 e 2009.

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 & 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. No quadro seguinte encontra-se a categorização atual da Taxonomia de Bloom.

<p><b>1. Lembrar:</b> Relacionado a reconhecer e reproduzir ideias e conteúdos. Reconhecer requer distinguir e selecionar uma determinada informação e reproduzir ou recordar está mais relacionado à busca por uma informação relevante memorizada. Representado pelos seguintes verbos no gerúndio: Reconhecendo e Reproduzindo.</p>
<p><b>2. Entender:</b> Relacionado a estabelecer uma conexão entre o novo e o conhecimento previamente adquirido. A informação é entendida quando o aprendiz consegue reproduzi-la com suas “próprias palavras”. Representado pelos seguintes verbos no gerúndio: Interpretando, Exemplificando, Classificando, Resumindo, Inferindo, Comparando e Explicando.</p>
<p><b>3. Aplicar:</b> Relacionado a executar ou usar um procedimento numa situação específica e pode também abordar a aplicação de um conhecimento numa situação nova. Representado pelos seguintes verbos no gerúndio: Executando e Implementando.</p>
<p><b>4. Analisar:</b> Relacionado a dividir a informação em partes relevantes e irrelevantes, importantes e menos importantes e entender a inter-relação existente entre as partes. Representado pelos seguintes verbos no gerúndio: Diferenciando, Organizando, Atribuindo e Concluindo.</p>
<p><b>5. Avaliar:</b> Relacionado a realizar julgamentos baseados em critérios e padrões qualitativos e quantitativos ou de eficiência e eficácia. Representado pelos seguintes verbos no gerúndio: Checando e Criticando.</p>
<p><b>6. Criar:</b> Significa colocar elementos junto com o objetivo de criar uma nova visão, uma nova solução, estrutura ou modelo utilizando conhecimentos e habilidades previamente adquiridos. Envolve o desenvolvimento de ideias novas e originais, produtos e métodos por meio da percepção da interdisciplinaridade e da interdependência de conceitos. Representado pelos seguintes verbos no gerúndio: Generalizando, Planejando e Produzindo.</p>

Quadro 1 - Estrutura do processo cognitivo na taxonomia de Bloom – revisada.  
Fonte: Ferraz & Belhot (2010)

As questões utilizadas para avaliar o desempenho do professor aderentes à taxonomia de Bloom, que possibilitam ao aluno realizar as identificadas atividades, foram as seguintes:

Tabela 13- Atividades Desenvolvidas em Língua Portuguesa

Proporção de Professores que realizam a atividade	2007		2009		2011	
	%	Nº Obs	%	Nº Obs	%	Nº Obs
24. Copiar textos extensos do livro didático ou do quadro	28,2	12.161	69,0	10.530	60,2	11.281
25. Conversar sobre textos de jornais e revistas	91,7	12.191	90,8	10.546	92,1	11.192
26. Fazer exercícios de gramática relacionados com textos de jornais ou revistas	87,0	12.164	78,7	10.386	89,4	11.087
27. Questão sem correspondente em 2011						
28. Ler, discutir com os colegas e escrever textos relacionados com o desenvolvimento de projeto temático	84,6	12.141	81,2	10.372	85,2	11.061
29. Ler contos, crônicas, poesias ou romances	93,2	12.231	80,9	10.346	92,6	11.031
30. Conversar sobre contos, crônicas, poesias ou romances	92,5	12.144	80,3	10.336	92,0	11.010
31. Usar contos, crônicas, poesias ou romances para exercitar aspectos da gramática	89,1	12.128	76,1	10.309	89,5	10.998
32. Fixar nomes de conceitos gramaticais e linguísticos	72,4	12.108	66,3	10.293	82,1	10.984
33. Discutir um texto, explorando as diferenças entre fatos e opiniões	95,6	12.139	86,1	10.309	95,9	11.005

Fonte: Elaboração própria com base nos dados do INEP

Tabela 14- Atividades Desenvolvidas em Matemática

Proporção de Professores que realizam a atividade	2007		2009		2011	
	%	Nº Obs	%	Nº Obs	%	Nº Obs
34. Fazer exercícios para automatizar procedimentos	86,8	12.039	87,5	10.591	93,5	11.754

Proporção de Professores que realizam a atividade	2007		2009		2011	
	%	Nº Obs	%	Nº Obs	%	Nº Obs
35. Lidar com problemas que exigem raciocínios diferentes e mais complexos que a maioria dos exemplos usuais	95,5	12.207	94,8	10.573	96,7	11.659
36. Falar sobre suas soluções, discutindo caminhos usados para encontrá-las	97,8	12.132	96,7	10.550	98,4	11.637
37. Gravar as regras que permitem obter as respostas certas dos cálculos e problemas	77,4	12.109	70,4	10.276	84,5	11.499
38. Lidar com temas que aparecem em jornais e/ou revistas, discutindo a relação dos temas com a Matemática.	79,4	12.126	78,6	10.312	82,4	11.473
39. Interpretar resultados numéricos obtidos para dar uma resposta adequada ao problema.	96,1	12.166	86,2	10.258	95,1	11.458
40. Lidar com situações que lhes sejam familiares e que apresentem temas do interesse dos alunos.	94,7	12.142	94,6	10.371	94,3	11.516
41. Experimentar diferentes modos de resolver um problema ou de efetuar um cálculo.	97,9	12.180	87,9	10.277	95,8	11.449
42. Aprimorar a precisão e a velocidade de execução de cálculos.	85,4	12.123	78,1	10.252	86,8	11.444
43. Experimentar diferentes ações (coletar informações recortar, analisar, explorar, discutir, manipular etc.) para resolver problemas.	85,4	12.171	85,7	10.315	87,9	11.463

Fonte: Elaboração própria com base nos dados do INEP

Em ambos os quadros apresentados, nota-se que há predominância das respostas em que o respondente afirma que realiza a atividade. Mesmo assim, a distribuição demonstra que haveria condições para mudança do percentual dos professores que executam cada atividade nas avaliações seguintes. Desta forma, a avaliação do desempenho baseada nestas questões indica ser adequada para averiguar o desempenho do professor de português e matemática, tanto da 4ª quanto da 8ª série.

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, tanto em português quanto em matemática, e se apresentou com a seguinte configuração:

<b>Alternativa</b>	<b>Frequência</b>
(A)	Semanalmente
(B)	Algumas vezes por mês
(C)	Uma vez por mês
(D)	Uma vez por bimestre
(E)	Nunca

Quadro 2 – Modelo de resposta questionário do professor

Fonte: Elaboração do autor com base nos dados do INEP

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 questionário 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 utilizado 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. A questão foi transformada em uma variável *dummy*, tratando como satisfatório as respostas na alternativa C e D, indicando que 60% ou mais foi dado em sala de aula, e insatisfatória para as alternativas A e B, que indicavam que o conteúdo ministrado ficou abaixo de 60%. 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 citados no parágrafo anterior. Para eles, 60% do conteúdo pedagógico é o mínimo aceitável.

A questão 55, que foi utilizada para avaliar o tanto o professor de português quanto de matemática, se apresenta da seguinte forma:

Tabela 15- Conteúdo Programático Desenvolvido

Proporção de Professores que realizam a atividade	2007		2009		2011	
	%	Nº Obs.	%	Nº Obs.	%	Nº Obs.
55. Ministrou menos de 60% do conteúdo programático	13,6	24.195	17,1	10.943	11,5	15.721

Fonte: Elaboração do autor com base nos dados do INEP

Com base nestas questões, serão avaliados o desempenho dos professores: português 4ª série, matemática 4ª série, português 8ª série e matemática 8ª série.

As avaliações em separado se justificam porque os professores da 4ª e da 8ª série, não importando se de português ou matemática, são, quase sempre, pessoas distintas. Esta diferenciação é menos intensa na 4ª série, 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 4ª e 8ª série e matemática 4ª e 8ª série (variável dependente). A média agrupa o resultado dos professores 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.

Definido a composição das variáveis dependentes, o próximo passo é fazer o pareamento pelo método do Escore de Propensão (*Propensity Score*), pareado pelo vizinho mais próximo (*nearest neighbor*), sem reposição. Este é o mesmo procedimento utilizado quando da avaliação do diretor e também será utilizado quando da avaliação do aluno, na etapa seguinte deste trabalho.

As características representadas pelas covariadas se mostraram em quantidade expressiva, sendo identificadas 60 para os professores, estando entre essas as variáveis controladas em comum e as específicas.

A tabela a seguir apresenta as variáveis controladas obtidas a partir dos questionários dos professores.

Tabela 16- Descrição das Covariadas do Professor

Proporção de Professores com a característica	2007		2009		2011	
	%	Nº Obs	%	Nº Obs	%	Nº Obs
Sexo Masculino	14,5	24.452	12,6	11.168	16,4	16.108
Menor que 30 anos	13,3	24.412	12,1	11.049	9,7	16.078
Entre 30 e 50 anos	67,0	24.412	67,9	11.049	68,5	16.078
Maior que 50 anos	19,7	24.412	20,0	11.049	21,8	16.078
Branco ou pardo	91,3	24.289	91,6	11.048	91,2	15.963
Escolaridade Ensino Superior	92,7	20.849	93,9	9.294	96,3	14.704
Ensino Superior Licenciatura em Matemática	25,1	20.849	25,3	9.294	31,1	14.704
Ensino Superior Licenciatura em Letras	28,2	20.849	28,8	9.294	34,9	14.704
Curso Superior em Instituição Pública	14,6	24.222	13,3	11.036	13,2	15.856
Curso Superior Presencial	92,5	24.227	93,3	11.042	94,8	15.850
Pós graduação de Longa Duração	31,1	24.147	33,6	10.777	40,0	15.608
Participação em atividades de formação continuada	83,1	24.311	75,1	10.987	76,5	15.951
Renda Individual na Escola Até R\$ 700 (2007)	4,0	24.253	3,2	11.037	3,3	12.727
Renda Individual na Escola Até R\$ 1.100 (2007)	15,7	24.253	12,7	11.037	8,6	12.727
Renda Individual na Escola Até R\$ 1.500 (2007)	40,0	24.253	38,1	11.037	27,9	12.727
Renda Individual na Escola Superior a R\$ 1.501 (2007)	40,3	24.253	46,0	11.037	60,2	12.727
Não possui outras fontes de renda	58,7	24.274	58,4	11.001	61,6	15.316
Renda Individual Até R\$ 900 (2007)	6,8	9.895	6,0	6.268	10,4	6.998
Renda Individual Até R\$ 1.300 (2007)	12,8	9.895	15,5	6.268	12,6	6.998
Renda Individual Até R\$ 1.900 (2007)	25,4	9.895	42,4	6.268	34,2	6.998
Renda Individual Superior a R\$ 1.901 (2007)	55,0	9.895	36,2	6.268	42,8	6.998
Até 2 anos de experiência como professor	4,4	24.342	3,6	11.057	6,4	11.148
Entre 2 e 4 anos de experiência como professor	9,9	24.342	10,4	11.057	10,0	11.148



Proporção de Professores com a característica	2007		2009		2011	
	%	Nº Obs	%	Nº Obs	%	Nº Obs
Mais de 5 anos de experiência como professor	85,8	24.342	86,0	11.057	83,6	11.148
Até 2 anos de experiência como professor nesta escola	47,3	24.411	44,1	11.090	30,8	15.586
Entre 2 e 4 anos de experiência como professor nesta escola	21,8	24.411	23,9	11.090	25,2	15.586
Mais de 5 anos de experiência como professor nesta escola	30,9	24.411	31,9	11.090	43,9	15.586
Até 2 anos de experiência como professor de alunos da série estudada	23,4	24.405	20,8	11.080	37,4	16.047
Entre 2 e 4 anos de experiência como professor de alunos da série estudada	19,9	24.405	20,3	11.080	23,9	16.047
Mais de 5 anos de experiência como professor de alunos da série estudada	56,7	24.405	58,8	11.080	38,7	16.047
Carga horária na escola inferior a 30 horas semanais	40,1	23.870	44,6	10.954	35,8	15.904
Carga horária na escola superior a 30 e inferior a 40 horas semanais	46,1	23.870	41,6	10.954	43,9	15.904
Carga horária na escola superior a 40 horas semanais	13,7	23.870	13,7	10.954	20,3	15.904
Trabalha em apenas 1 escola	60,4	24.360	59,5	11.019	65,9	16.080
Trabalha tempo não integral	68,2	9.514	44,5	7.957	37,1	14.578
Professor Celetista ou estatutário (não temporário ou informal)	75,1	23.876	74,4	10.792	77,7	15.779
Todos os alunos das turmas em que leciona têm livros didáticos	68,9	24.179	70,7	9.038	83,0	15.785
Os alunos receberam os livros no início do ano letivo	86,3	24.140	79,0	9.231	95,3	15.535
Renda do professor nessa escolar/Renda Total Professor<75%	66,0	8.944	48,6	5.704	48,7	5.218

Fonte: Elaboração do autor com base nos dados do INEP

Destaca-se nesta tabela que os professores das escolas são predominantemente mulheres e tem entre 30 e 50 anos, com a grande maioria sendo formada de brancos ou pardos e tem nível de ensino superior. Mais de 60% deles não possui outras fontes de renda. Também pode ser notado que mais da metade possui ao menos 5 anos de experiência como professor e quase a metade deles trabalha em tempo integral.

Na sequência dos testes, dado que o processo de pareamento para selecionar covariadas com diferenças entre médias não significantes foi abandonado na avaliação do

diretor e para manter coerência metodológica, optou-se pelo pareamento utilizando todo o vetor de covariadas, conforme explicado na seção 9.1, sobre o desempenho dos diretores.

Após o emparelhamento, do universo de 292.828 observações possíveis, foram selecionadas pelo método do *propensity score* o seguinte quantitativo com características similares:

Tabela 17- Quantitativo de professores selecionados para avaliação

Disciplina e Série	Quantidade
Professor de Português da 4ª série	3.075
Professor de Português da 8ª série	2.524
Professor de Matemática da 4ª série	3.024
Professor de Matemática da 8ª série	2.670

Fonte: Elaboração do autor com base nos dados do INEP

Entre estes selecionados, serão excluídos aqueles com escore de propensão superior ao menor escore máximo entre tratados e controles ou inferior ao maior mínimo entre os dois e não comporão a área de suporte comum. O resultado se apresentou com a seguinte composição:

Tabela 18- Área de Suporte Comum - Professores.

	Condição	Suporte Comum		Total
		Fora	Dentro	
Prof. Português 4ª série	Controles	0	1.772	1.772
	Tratados	27	1.276	1.303
	Total	27	3.048	3.075
Prof. Português 8ª série	Controles	0	346	346
	Tratados	1.832	346	2.178
	Total	1.832	692	2.524
Prof. Matemática 4ª série	Controles	0	1.746	1.746
	Tratados	21	1.257	1.278
	Total	21	2.993	3.024
Prof. Matemática 8ª série	Controles	0	384	384
	Tratados	1.902	384	2.286
	Total	1.902	768	2.670

Fonte: Elaboração do autor com base nos dados do INEP

O pareamento se mostrou mais rigoroso para todos os grupos avaliados do que para os diretores, sendo que para os professores de 8ª série, seja de português ou matemática, menos de 400 indivíduos foram selecionados. Esse aumento no rigor se deve a um maior número de covariadas, sendo 60 para os professores enquanto para os diretores foram 47.

Após o pareamento há necessidade de testar se há diferenças nos resultados entre grupo de tratamento e controle no período de pré-tratamento (2007). Os resultados de 2007 podem ser vistos nas 4 tabelas seguintes. Considerando que o interesse maior está em saber se há diferença depois de realizado o pareamento, e dado a quantidade de resultados a ser apresentada, para simplificar a leitura as tabelas apresentam apenas o resultado de cada variável dependente depois o pareamento. Os dados completos estão disponíveis para consulta. As diferenças estão destacadas na coluna em negrito.

Tabela 19- Resultado do Pareamento do Professor de Português da 4ª série

Variável	Tratado	Controle	<b>Diferença</b>	Desv. Pad.	T-estat
Copiar textos extensos do livro didático ou do quadro	0,2409	0,3113	<b>-0,0704</b>	0,0291	-2,41
Conversar sobre textos de jornais e revistas	0,9595	0,9595	<b>0,0000</b>	0,0129	0
Fazer exercícios de gramática relacionados com textos de jornais ou revistas	0,9104	0,8870	<b>0,0235</b>	0,0197	1,19
Ler, discutir com os colegas e escrever textos relacionados com o desenvolvimento de projeto temático	0,9104	0,8870	<b>0,0235</b>	0,0197	1,19
Ler contos, crônicas, poesias ou romances	0,9808	0,9531	<b>0,0277</b>	0,0117	2,38
Conversar sobre contos, crônicas, poesias ou romances	0,9723	0,9467	<b>0,0256</b>	0,0129	1,99
Usar contos, crônicas, poesias ou romances para exercitar aspectos da gramática	0,9211	0,9062	<b>0,0149</b>	0,0184	0,81
Fixar nomes de conceitos gramaticais e linguísticos	0,6588	0,6738	<b>-0,0149</b>	0,0308	-0,48
Discutir um texto, explorando as diferenças entre fatos e opiniões	0,9744	0,9765	<b>-0,0021</b>	0,0101	-0,21
Ministrou menos de 60% do conteúdo programático	0,1279	0,1109	<b>0,0171</b>	0,0212	0,8

Fonte: Elaboração do autor com base nos dados do INEP

Tabela 20- Resultado do Pareamento do Professor de Português da 8ª série

Variável	Tratado	Controle	Diferença	Desv. Pad.	T-estat
Copiar textos extensos do livro didático ou do quadro	0,2874	0,1257	<b>0,1617</b>	0,0435	3,71
Conversar sobre textos de jornais e revistas	0,8024	0,9102	<b>-0,1078</b>	0,0380	-2,83
Fazer exercícios de gramática relacionados com textos de jornais ou revistas	0,8144	0,8263	<b>-0,0120</b>	0,0421	-0,28
Ler, discutir com os colegas e escrever textos relacionados com o desenvolvimento de projeto temático	0,7725	0,7126	<b>0,0599</b>	0,0479	1,25
Ler contos, crônicas, poesias ou romances	0,8263	0,8084	<b>0,0180</b>	0,0424	0,42
Conversar sobre contos, crônicas, poesias ou romances	0,8383	0,7784	<b>0,0599</b>	0,0431	1,39
Usar contos, crônicas, poesias ou romances para exercitar aspectos da gramática	0,7964	0,8084	<b>-0,0120</b>	0,0437	-0,27
Fixar nomes de conceitos gramaticais e linguísticos	0,7066	0,6527	<b>0,0539</b>	0,0511	1,05
Discutir um texto, explorando as diferenças entre fatos e opiniões	0,9102	0,9341	<b>-0,0240</b>	0,0294	-0,82
Ministrou menos de 60% do conteúdo programático	0,1138	0,1198	<b>-0,0060</b>	0,0352	-0,17

Fonte: Elaboração do autor com base nos dados do INEP

Tabela 21- Resultado do Pareamento do Professor de Matemática da 4ª série

Variável	Tratado	Controle	Diferença	Desv. Pad.	T-estat
Fazer exercícios para automatizar procedimentos	0,7926	0,8210	<b>-0,0284</b>	0,0261	-1,09
Lidar com problemas que exigem raciocínios diferentes e mais complexos que a maioria dos exemplos usuais	0,9694	0,9563	<b>0,0131</b>	0,0125	1,05
Falar sobre suas soluções, discutindo caminhos usados para encontra-las	0,9956	0,9934	<b>0,0022</b>	0,0049	0,45
Gravar as regras que permitem obter as respostas certas dos cálculos e problemas	0,7009	0,6834	<b>0,0175</b>	0,0305	0,57
Lidar com temas que aparecem em jornais e/ou revistas, discutindo a relação dos temas com a Matemática.	0,8690	0,8777	<b>-0,0087</b>	0,0220	-0,4

Variável	Tratado	Controle	Diferença	Desv. Pad.	T-estat
Interpretar resultados numéricos obtidos para dar uma resposta adequada ao problema.	0,9716	0,9672	<b>0,0044</b>	0,0114	0,38
Lidar com situações que lhes sejam familiares e que apresentem temas do interesse dos alunos.	0,9847	0,9891	<b>-0,0044</b>	0,0075	-0,58
Experimentar diferentes modos de resolver um problema ou de efetuar um cálculo.	0,9978	0,9869	<b>0,0109</b>	0,0057	1,9
Aprimorar a precisão e a velocidade de execução de cálculos.	0,8712	0,8100	<b>0,0611</b>	0,0241	2,53
Experimentar diferentes ações (coletar informações recortar, analisar, explorar, discutir, manipular etc.) para resolver problemas.	0,9279	0,9301	<b>-0,0022</b>	0,0170	-0,13
Ministrou menos de 60% do conteúdo programático	0,1419	0,1135	<b>0,0284</b>	0,0221	1,29

Fonte: Elaboração do autor com base nos dados do INEP

Tabela 22- Resultado do Pareamento do Professor de Matemática da 8ª série

Variável	Tratado	Controle	Diferença	Desv. Pad.	T-estat
Fazer exercícios para automatizar procedimentos	0,7926	0,8210	<b>-0,0284</b>	0,0261	-1,09
Lidar com problemas que exigem raciocínios diferentes e mais complexos que a maioria dos exemplos usuais	0,9694	0,9563	<b>0,0131</b>	0,0125	1,05
Falar sobre suas soluções, discutindo caminhos usados para encontra-las	0,9956	0,9934	<b>0,0022</b>	0,0049	0,45
Gravar as regras que permitem obter as respostas certas dos cálculos e problemas	0,7009	0,6834	<b>0,0175</b>	0,0305	0,57
Lidar com temas que aparecem em jornais e/ou revistas, discutindo a relação dos temas com a Matemática.	0,8690	0,8777	<b>-0,0087</b>	0,0220	-0,4
Interpretar resultados numéricos obtidos para dar uma resposta adequada ao problema.	0,9716	0,9672	<b>0,0044</b>	0,0114	0,38
Lidar com situações que lhes sejam familiares e que apresentem temas do interesse dos alunos.	0,9847	0,9891	<b>-0,0044</b>	0,0075	-0,58
Experimentar diferentes modos de resolver um problema ou de efetuar um cálculo.	0,9978	0,9869	<b>0,0109</b>	0,0057	1,9

Variável	Tratado	Controle	Diferença	Desv. Pad.	T-estat
Aprimorar a precisão e a velocidade de execução de cálculos.	0,8712	0,8100	<b>0,0611</b>	0,0241	2,53
Experimentar diferentes ações (coletar informações recortar, analisar, explorar, discutir, manipular etc.) para resolver problemas.	0,9279	0,9301	<b>-0,0022</b>	0,0170	-0,13
Ministrou menos de 60% do conteúdo programático	0,1419	0,1135	<b>0,0284</b>	0,0221	1,29

Fonte: Elaboração do autor com base nos dados do INEP

Os resultados nas 4 tabelas se mostraram significantes para algumas variáveis, indicando que há diferenças de médias entre indivíduos tratados e controles, o que não deveria ocorrer, pois o PRV ainda não havia sido implantado, o que veio acontecer somente em 2008. Esse problema também foi relatado no experimento do diretor.

As diferenças indicam para a existência de variáveis não observáveis que afetaram os resultados, ou seja, características do professor que não puderam ser controlados no experimento proporcionam a formação de pares não tão perfeitos. Para eliminar este viés foi aplicado o método do estimador de Diferenças em Diferenças (DD), que permite eliminar a influência das características não observadas com efeitos fixos no tempo.

Desta forma, de acordo com o apresentado nesta seção, pode-se avaliar isoladamente o impacto do programa sobre o desempenho dos professores baseado nas técnicas pedagógicas empregadas na aplicação do conteúdo ministrado em português e em matemática.

O estimador DD foi calculado pelo método dos Mínimos Quadrados Ordinários a partir do seguinte modelo:

$$Y_{igst} = \beta_0 + \beta_1 EESP_s + \beta_2 Ano_t + \beta_3 EESP_s * Ano_t + B_4 E_{st} + B_5 D_{vst} + B_6 A_{vst} + B_7 X_{ist} + \varepsilon_{ist}$$

Onde:

$Y_{igst}$  é a variável dependente e identifica o efeito do modelo no desempenho do professor, no qual  $i$ ,  $g$ ,  $s$ ,  $t$  denotam o professor, a série, se é uma escola estadual ou municipal e se é referente a antes ou depois da implantação do programa;

$EESP_s = 1$ , se é escola estadual paulista (ou seja, se recebeu tratamento) e 0 caso seja escola municipal;

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

O  $\beta_3$  é o coeficiente da variável de interesse que é composta pela interação das duas variáveis anteriores: se é escola estadual paulista e se a observação se refere ao período posterior à implantação do programa. Esta variável mensurará o efeito do PRV sobre o desempenho dos professores;

$E_{st}$  vetor de covariadas com 11 características que controlam as condições de uso da escola (manutenção, computadores e segurança)

$D_{vst}$  vetor de 6 covariadas que avaliam o desempenho do diretor de uma escola estadual paulista depois do programa;

$A_{vst}$  vetor de covariadas das 7 características que controlam o entorno do aluno (ambiente socioeconômico);

$X_{ist}$  é o vetor de covariadas com as demais características do professor;

$\varepsilon_{ist}$  é o resíduo da regressão.

### 9.2.1. Resultados dos professores

Os resultados da estimação pelo método de Diferenças em Diferenças pode ser observado na variável de interesse, apresentada na forma de uma *dummy* resultante da interação do Ano (0 se antes ou 1 se depois do programa) com o Estado (dependência administrativa da escola: 0 se municipal ou 1 se estadual). O coeficiente desta variável indicou o efeito no desempenho dos indivíduos tratados, ou seja, professor participante do programa de remuneração variável das escolas estaduais paulista após a implantação do programa em 2008. Este procedimento será replicado para cada uma das 9 questões referentes às atividades desenvolvidas em língua portuguesa tanto na 4ª quanto na 8ª série, que são as variáveis dependentes, obtendo com isto 9 coeficientes distintos por série. O mesmo procedimento será realizado para as 10 questões referentes às atividades de matemática.

Além das covariadas apresentadas na seção anterior, específicas do professor, controla-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 na seção 9.1.1 deste trabalho.

Dado a quantidade de variáveis de controle, a tabela a seguir apresenta apenas os coeficientes da variável de interesse (Ano\*Estado), considerando cada variável dependente como representativa de um critério de avaliação do desempenho do professor. As questões estão numeradas conforme consta no questionário que lhe deu origem. Os resultados completos estão disponíveis para consulta.

Tabela 23- Coeficientes estimados – Professor de Português

Variáveis Dependentes	4ª Série	8ª Série
24. Copiar textos extensos do livro didático ou do quadro	-0,0563 (0,0351)	-0,0764 (0,0685)
25. Conversar sobre textos de jornais e revistas	0,0046 (0,0151)	0,1727*** (0,0597)
26. Fazer exercícios de gramática relacionados com textos de jornais ou revistas	-0,0558 ** (0,0226)	0,0993 (0,0629)
28. Ler, discutir com os colegas e escrever textos relacionados com o desenvolvimento de projeto temático	0,0188 (0,0228)	0,0560 (0,0735)
29. Ler contos, crônicas, poesias ou romances	-0,0154 (0,0126)	-0,0722 (0,0595)
30. Conversar sobre contos, crônicas, poesias ou romances	-0,0047 (0,0142)	-0,0795 (0,0640)
31. Usar contos, crônicas, poesias ou romances para exercitar aspectos da gramática	-0,0416 ** (0,0211)	-0,0630 (0,0661)
32. Fixar nomes de conceitos gramaticais e linguísticos	-0,1006 *** (0,0350)	-0,1360* (0,0750)
33. Discutir um texto, explorando as diferenças entre fatos e opiniões	0,0111 (0,0114)	0,0739 (0,0537)
55. Ministrou menos de 60% do conteúdo programático	-0,0214 (0,0258)	0,0195 (0,0531)

Fonte: Elaboração do autor com base nos dados do INEP

(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 da 4ª série tiveram efeito negativo e estatisticamente significativo no desempenho mensurado pelas questões 26, 31 e 32. Nas demais 7 questões utilizadas para avaliar o desempenho do professor da 4ª série não se constatou significância estatística.

Em nenhuma das atividades previstas os professores de português da 4ª série melhoraram seu desempenho.



Quanto aos professores da 8ª série de português, o efeito é positivo e significativo para 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. A questão 32 apresenta resultado negativo somente a 10% de nível de significância. As demais 8 questões não apresentaram resultados significantes.

Os resultados dos professores de português da 8ª série são pouco expressivos, indicando que o PRV, por si só, não gerou influência relevante no desempenho do professor de português da 8ª série.

Tabela 24- Coeficientes estimados – Professor de Matemática

Variáveis Dependentes	4ª Série	8ª Série
34.Fazer exercícios para automatizar procedimentos	-0,0411 (0,0294)	0,0389 (0,0379)
35.Lidar com problemas que exigem raciocínios diferentes e mais complexos que a maioria dos exemplos usuais	-0,0066 (0,0136)	-0,0237 (0,0346)
36.Falar sobre suas soluções, discutindo caminhos usados para encontra-las	0,0020 (0,0081)	-0,0217 (0,0201)
37.Gravar as regras que permitem obter as respostas certas dos cálculos e problemas	-0,0980 *** (0,0356)	-0,0789 (0,0562)
38.Lidar com temas que aparecem em jornais e/ou revistas, discutindo a relação dos temas com a Matemática.	0,0218 (0,0247)	0,1109 (0,0703)
39.Interpretar resultados numéricos obtidos para dar uma resposta adequada ao problema.	-0,0153 (0,0131)	0,0641* (0,0338)
40.Lidar com situações que lhes sejam familiares e que apresentem temas do interesse dos alunos.	-0,0015 (0,0099)	-0,0515 (0,0451)
41.Experimentar diferentes modos de resolver um problema ou de efetuar um cálculo.	-0,0041 (0,0076)	0,0139 (0,0365)
42.Aprimorar a precisão e a velocidade de execução de cálculos.	-0,0472 * (0,0267)	0,0167 (0,0585)
43.Experimentar diferentes ações (coletar informações recortar, analisar, explorar, discutir, manipular etc.) para resolver problemas.	0,0009 (0,0192)	0,0651 (0,0638)
55. Ministrou menos de 60% do conteúdo programático	-0,0334 (0,0261)	0,0181 (0,0532)

Fonte: Elaboração do autor com base nos dados do INEP

(1) Desvio padrão entre parênteses

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

Pela tabela anterior os professores de matemática da 4ª série tiveram efeito negativo e estatisticamente significativo nas questões 37 e 42, sendo que nesta última apenas ao nível de significância de 10%. Os demais 8 coeficientes não indicam efeitos significantes proporcionados pelo programa.

Em nenhuma das atividades previstas os professores de matemática da 4ª série melhoraram seu desempenho.

Quanto aos professores da 8ª série de matemática, o efeito é positivo e significativo apenas para as questões 39, assim mesmo apenas a 10% de significância e indicou que o programa impactou em 6,41% o resultado deste quesito avaliado.

Os resultados dos professores de matemática tanto da 4ª quanto da 8ª série são pouco expressivos, indicando que o PRV, por si só, não gerou influência relevante no desempenho do professor de matemática.

### **9.3. Testes do efeito do programa sobre o desempenho dos alunos**

A avaliação do efeito do programa sobre os alunos não é propriamente o objeto de interesse direto 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 refletir no desempenho do aluno, torna-se necessário avaliar o efeito finalístico do desempenho dos diretores e professores 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 foram utilizados os dados da Prova Brasil, que é um teste padronizado. Conforme demonstra Hanushek e Rivkin (2010), Hanushek e Woessmann (2008), as notas em testes padronizados para medir habilidades cognitivas são altamente relacionadas a fatores como os rendimentos escolares. Neste mesmo sentido, Moriconi (2012) afirma em sua pesquisa o seguinte:

Para medir a qualidade de atributos dos indivíduos, os melhores instrumentos que se tem conhecimento são os testes padronizados de larga escala. Esses testes geram medidas de habilidades cognitivas dos indivíduos, como a leitura e resolução de problemas matemáticos, as quais permitem com que esses atributos sejam quantificados e os alunos recebam um valor que traduz a qualidade de seu atributo em uma escala numérica. Por transformar a qualidade dos atributos dos indivíduos em quantidades que podem ser medidas antes e depois de um processo educacional, esses resultados têm grande potencial para serem usados como insumos e produtos nas funções de produção educacionais – como de fato se dá na grande maioria dos estudos da área.

Para saber se as notas dos alunos da rede estadual paulista é um bom instrumento de avaliação, é necessário comparar seus desempenhos em relação às notas dos alunos da rede municipal. Com isto, pode-se identificar se há condições dos alunos estaduais alterarem suas notas por conta de estímulos extrínsecos. Caso os alunos estaduais estiverem com notas muito boas, no topo, será difícil conseguir grandes avanços no seu desempenho, pois o mesmo já está próximo da sua excelência.

O resultado comparativo das notas médias dos alunos estaduais e municipais se apresentou da seguinte forma:

Tabela 25- Proficiência comparativa do Aluno 4ª série

Variável	2007		2009		2011	
	Estado	Municípios	Estado	Municípios	Estado	Municípios
Língua Portuguesa	178,31	184,17	190,39	191,83	194,61	199,62
Matemática	195,23	202,52	214,37	213,32	216,79	219,50

Fonte: Elaboração do autor com base nos dados do INEP

Tabela 26- Proficiência comparativa do Aluno 8ª série

Variável	2007		2009		2011	
	Estado	Municípios	Estado	Municípios	Estado	Municípios
Língua Portuguesa	232,02	241,35	239,79	250,34	242,10	253,59
Matemática	242,33	253,43	242,27	253,95	245,29	259,85

Fonte: Elaboração do autor com base nos dados do INEP

Os dados apresentados indicam que, com exceção do aluno de 8ª série em matemática no ano de 2009, todas as demais notas dos alunos da rede estadual estão abaixo das notas nos municípios. Portanto, não se encontram no limite superior, havendo condições de serem melhoradas. Assim, as mesmas podem servir como instrumento de avaliação de desempenho dos alunos da rede estadual.

Baseado nestas referências, o aluno foi avaliado pela proficiência em português e em matemática separadamente e nas respectivas séries. As variáveis de interesse foram as proficiências em Português e Matemática dos alunos da 4ª e 8ª série apuradas pela Prova Brasil, nos anos de 2007, antes do programa, e 2009 e 2011, depois do programa.

A estatística descritiva das variáveis de interesse do aluno se apresentou da seguinte forma:

Tabela 27- Variáveis dependentes – Aluno 4 série

Variável	2007		2009		2011	
	Média	Nº Obs	Média	Nº Obs	Média	Nº Obs
Proficiência em Língua Portuguesa	181,42	203.258	191,21	198.036	197,62	174.610
Proficiência em Matemática	199,09	203.258	213,78	198.036	218,42	174.610

Fonte: Elaboração do autor com base nos dados do INEP

Tabela 28- Variáveis dependentes – Aluno 8 série

Variável	2007		2009		2011	
	Média	Nº Obs	Média	Nº Obs	Média	Nº Obs
Proficiência em Língua Portuguesa	233,05	181.245	241,05	160.122	243,57	181.592
Proficiência em Matemática	243,55	181.245	243,68	160.122	247,16	181.592

Fonte: Elaboração do autor com base nos dados do INEP

A avaliação do desempenho em cada disciplina foi feita pela média aritmética por escola das notas dos alunos nos quatro segmentos apresentados; português 4ª e 8ª série e matemática 4ª e 8ª série (variável dependente). 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 10 características em português e 11 em matemática, sem o diretor. Tais desempenhos são os mesmos que foram apurados nas duas seções anteriores a esta.

Coerente com a avaliação dos diretores e professores, aplicou-se o método de Pareamento por Escore de Propensão (*Propensity Score*), pareado pelo vizinho mais próximo (*nearest neighbor*), sem reposição.

As características representadas pelas covariadas se mostraram em quantidade expressiva, sendo identificadas 67 para os alunos, estando entre essas as variáveis controladas em comum e as específicas. Porém, a quantidade de variáveis controladas se alteram para 71 ou 72, caso os controles sejam feitos pelos resultados dos professores de português ou matemática ao invés do diretor.

A tabela a seguir apresenta as características que mereceram destaques e que compõem o vetor de covariadas  $X_{ist}$  obtido a partir dos questionários dos alunos, com a proporção de alunos por características controladas. A tabela completa pode ser observada no APÊNDICE.

Tabela 29- Covariadas selecionadas dos alunos

Proporção de alunos com a característica	2007		2009		2011	
	%	No Obs	%	No Obs	%	No Obs
Idade adequada para a série	91,0	394.078	91,8	379.417	92,5	376.548
Tem computador com internet em casa	33,8	389.992	45,9	379.597	62,7	375.862
Tem pelo menos uma diarista em casa	10,8	378.145	8,9	380.367	8,8	375.785
Não sabe informar a escolaridade da mãe	24,1	380.798	25,8	365.840	27,1	367.887
Não sabe informar a escolaridade do Pai	34,3	384.532	35,3	370.932	34,5	339.554
Presença frequente dos responsáveis em reunião dos pais	69,9	387.196	69,3	374.074	67,8	372.604
Aluno não trabalha fora de casa	85,7	384.572	86,8	359.401	87,3	366.759
Começou a estudar na pré-escola	43,7	382.403	44,0	366.321	47,0	369.813
Nunca foi reprovado	82,5	386.385	81,4	368.525	82,2	371.397
Se 8 série - Após a 8 série pretende apenas trabalhar	2,7	183.306	2,4	150.777	2,5	187.500

Fonte: Elaboração do autor com base nos dados do INEP

Pode-se notar que mais de 90% tem idade adequada para a série, o que indica condições para que o professor consiga bom rendimento em sala, aproximadamente 2/3 tem apenas 1 banheiro em casa e pouco menos de 10% tem diarista em casa, o que indica limitações das condições financeiras, podendo dispendir parte do seu tempo, em que poderia estar estudando, para auxiliar nas tarefas domésticas.

Por outro lado, em 2011 cerca de 63% tinha computador com internet em casa, contra 33% em 2007, o que indica um aumento expressivo da disponibilidade desse bem e que pode auxiliar no aprendizado do aluno.

Cerca de 25% não soube informar a escolaridade da mãe, passando para 35% quando perguntado sobre a escolaridade do pai. Quase 70% informou também que os

responsáveis participam frequentemente das reuniões na escola, lembrando que este item foi um dos utilizados para mensurar o desempenho do diretor.

Nota-se também que mais de 85% não trabalha fora de casa, ou seja, os outros 15% trabalham, mesmo sabendo que são crianças com até 13 anos.

Mais de 80% indica que começou a estudar no maternal ou pré-escola, o que facilita o aprendizado. Mais de 80% nunca foi reprovado, portanto cerca de 20% já foi reprovado, e este aluno tem desempenho inferior ao demais. O coeficiente desta variável de controle se mostrou o mais robusto entre os outros 67 avaliados.

Alinhado com a metodologia utilizada para avaliar os diretores e professores, o pareamento foi feito utilizando todo o vetor de covariadas, dado que a tentativa de parer covariadas com diferenças entre médias não significantes se mostrou parcialmente eficaz, conforme explicitado na seção 9.1.

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 30- Quantitativo de Alunos selecionados para avaliação

Controle	Disciplina e Série	Quantidade
Diretor	Português e Matemática - 4ª série	223.183
	Português e Matemática - 8ª série	294.720
Professor	Português da 4ª série	222.360
	Português da 8ª série	266.984
	Matemática da 4ª série	294.720
	Matemática da 8ª série	266.047

Fonte: Elaboração do autor com base nos dados do INEP

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 da 4ª série será o mesmo, não importando se a disciplina é português ou matemática. Porém, há escolas que oferecem apenas um dos ciclos do ensino fundamental (4ª ou 8ª série), 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.

Entre os selecionados, foram excluídos aqueles com escore de propensão superior ao menor escore máximo entre tratados e controles ou inferior ao maior mínimo entre os dois. O resultado se apresentou com a seguinte composição:

Tabela 31- Área de Suporte Comum – Alunos

Controle	Série	Condição	Suporte Comum		Total
			Fora	Dentro	
Diretor	4ª Português e Matemática	Controles	-	118.080	118.080
		Tratados	15	105.085	105.103
		Total	15	223.168	223.183
	8ª Português e Matemática	Controles	-	28.149	28.149
		Tratados	238.422	28.149	266.571
		Total	238.422	56.298	294.720
Professor	4ª Português	Controles	-	127.519	127.519
		Tratados	183	94.658	94.841
		Total	183	222.177	222.360
	8ª Português	Controles	-	18.734	18.734
		Tratados	229.516	18.734	248.250
		Total	229.516	37.468	266.984
	4ª Matemática	Controles	-	127.412	127.412
		Tratados	94	94.815	94.909
		Total	94	222.227	222.321
	8ª Matemática	Controles	-	18.452	18.452
		Tratados	229.143	18.452	247.595
		Total	229.143	36.904	266.047

Fonte: Elaboração do autor com base nos dados do INEP

O pareamento se mostrou mais rigoroso para os alunos da 8ª série, seja controlado pelo diretor ou professor. A principal justificativa para esse rigor é o fato de haver menos escolas de 8ª série nos municípios, dado que este tem foco maior nas séries iniciais de ensino.

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.

Novamente fez-se o teste para saber se há diferenças nos resultados entre grupo de tratamento e controle no período de pré-tratamento (2007). Os resultados de 2007 podem ser vistos na tabela seguinte:

Tabela 32- Resultado Pareamento do Aluno – 2007 – controlado pelo diretor

Controle	Disciplina	Série	Tratado	Controle	Diferença	Desv. Pad.	T-estat
Diretor	Português	4 <sup>a</sup>	188,53	192,94	<b>-4,4098</b>	0,2941	-14,99
		8 <sup>a</sup>	243,63	243,89	<b>-0,2595</b>	0,6185	-0,42
	Matemática	4 <sup>a</sup>	204,90	210,60	<b>-5,7065</b>	0,3081	-18,52
		8 <sup>a</sup>	253,12	255,15	<b>-2,0290</b>	0,6116	-3,32
Professor	Português	4 <sup>a</sup>	190,13	192,45	<b>-2,3160</b>	0,3149	-7,35
		8 <sup>a</sup>	244,57	243,10	<b>1,4680</b>	0,7127	2,06
	Matemática	4 <sup>a</sup>	206,30	210,02	<b>-3,7262</b>	0,3294	-11,31
		8 <sup>a</sup>	255,95	255,19	<b>0,7623</b>	0,7093	1,07

Fonte: Elaboração do autor com base nos dados do INEP

Os testes de médias se mostraram com diferenças estatísticas significantes para quase todas as séries e disciplinas avaliadas, com exceção do aluno de português 8<sup>a</sup> série controlado pelo desempenho do diretor (t estatístico = -0,42)<sup>7</sup> e do aluno de matemática 8<sup>a</sup> série controlado pelo professor (t estatístico = 1,07). Tais diferenças não deveriam ocorrer, pois o PRV ainda não havia sido implantado, o que veio acontecer somente em 2008. Esse problema também foi constatado no experimento do diretor e professor.

As diferenças indicam a existência de variáveis não observáveis que afetaram os resultados, ou seja, características dos alunos que não puderam ser controladas no experimento proporcionam diferentes desempenhos entre tratados e controles. Para eliminar este viés aplicou-se o método do estimador de Diferenças em Diferenças (DD), que permite eliminar a influência das características não observadas com efeitos fixos no tempo.

Desta forma, de acordo com o apresentado nesta seção, pode-se avaliar isoladamente o impacto do programa sobre o desempenho dos alunos baseado na sua proficiência em português e matemática, tanto na 4<sup>a</sup> quanto na 8<sup>a</sup> série.

O estimador DD será calculado pelo método dos Mínimos Quadrados Ordinários a partir do seguinte modelo:

$$Y_{igdst} = \beta_0 + \beta_1 EESP_s + \beta_2 Ano_t + \beta_3 EESP_s * Ano_t + B_4 Y_{ivst} + B_5 X_{ist} + \varepsilon_{ist}$$

<sup>7</sup> Para ser considerado significativo, o t estatístico deverá ser igual ou superior a 1,57



Onde:

$Y_{igdst}$  é a variável dependente e identifica o efeito do modelo no desempenho do aluno, no qual  $i, g, d, s, t$  denotam o aluno, a série, a disciplina avaliada, se é uma escola estadual ou municipal e se é referente a antes ou depois da implantação do programa;

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

$Ano_t = 1$ , se é referente a 2009 ou a 2011 (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 das duas variáveis anteriores: se é escola estadual paulista e se a observação se refere ao período posterior à implantação do programa. Esta variável mensurará o efeito do PRV sobre o desempenho dos alunos;

$Y_{ivst}$  é o vetor em que  $i$  identifica se diretor ou professor,  $v$  será uma das variáveis dependentes do modelo do diretor ou professor,  $s$  a escola e  $t$  o ano correspondente;

$X_{ist}$  é o vetor de covariadas com as demais características do aluno;

$\varepsilon_{ist}$  é o resíduo da regressão

### 9.3.1. Resultados dos alunos

Os resultados da estimação pelo método de Diferenças em Diferenças pode ser observado na variável de interesse, que é uma *dummy* resultante da interação do Ano (0 se antes ou 1 se depois do programa) com o Estado (dependência administrativa da escola: 0 se municipal ou 1 se estadual). O coeficiente desta variável indica 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 será aplicado individualizadamente nas disciplinas de português e matemática, que são as variáveis dependentes, tanto na 4ª quanto na 8ª série, 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. A estatística descritiva da proficiência do aluno foi apresentada na seção anterior na tabela 9.23 e está incluída no Apêndice.

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.

Dado a quantidade de variáveis de controle, a tabela a seguir apresenta apenas os resultados da variável de interesse (Ano\*Estado) em relação às variáveis dependentes. Os resultados completos estão disponíveis para consulta.

Tabela 33- Coeficientes estimados – Aluno

Variáveis Dependentes	Controle	4a Série	8a Série
Proficiência Língua Portuguesa	Diretor	0,3965 (0,3502)	-4,4271 (0,7818) ***
	Professor	-0,4643 (0,3570)	-4,6135 (0,9209) ***
Proficiência Matemática	Diretor	4,1737 *** (0,3700)	-5,5710 *** (0,7873)
	Professor	3,3059 *** (0,3755)	-6,7245 *** (0,9457)

Fonte: Elaboração do autor com base nos dados do INEP

(1) Desvio padrão entre parênteses

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

Dos 8 resultados obtidos, apenas os desempenhos dos alunos de português da 4ª série controlado tanto pelo efeito do diretor como do professor não foram estatisticamente significantes. Todos os outros 6 coeficientes foram significantes a 1% de confiança.

Os efeitos positivos e significantes foram notados nas proficiência em matemática 4ª série, indicando que o aluno pode aumentar sua nota em 4,17 pontos<sup>8</sup>, se controlado pelo desempenho do diretor, e 3,30 pontos, se controlado pelo desempenho do professor.

Na 8ª série todos os coeficientes se apresentaram negativos e significantes, não importando se controlado pelo diretor ou professor, seja em matemática ou português, indicando que o PRV paulista não proporcionou melhoria do desempenho destes alunos.

<sup>8</sup> A escala SAEB varia de 0 a 500 pontos



## 10. Análise de Robustez

Para assegurar que os resultados sejam consistentes, recomenda-se que os testes sejam refeitos utilizando outro grupo de controle com características semelhantes, mantendo o mesmo grupo de tratamento.

A análise dos dados indicou que as escolas municipais dos 38 municípios paulistas poderiam ser substituídas por outras regiões, como por exemplo, a região centro oeste, que apresentou características socioeconômicas semelhantes. Os Estados da região sul, com exceção do Estado do Paraná, também demonstraram condições de exercerem a função de controles.

Porém, a base mais ampla para servir de grupo de controle pode ser encontrada entre as escolas dos demais Estados que não adotaram qualquer modalidade de PRV, em substituição às escolas municipais.

Para a perfeita seleção, foi realizada pesquisa em todas as Secretarias de Educação dos demais 25 estados brasileiros e no Distrito Federal, excluindo aqueles que adotaram programa de remuneração variável.

Nos Estados do Amazonas, Pernambuco, Minas Gerais e Espírito Santo foram encontrados programas de remuneração variável associados ao desempenho do servidor.

Nos Estados do Acre, Piauí e Paraíba não foi possível confirmar se há ou não PRV, dado que os respondentes não souberam informar com confiança se há um programa efetivo de incentivos no Estado. Para assegurar que os dados dos controles não contivessem viés, estes Estados também foram excluídos do grupo de controle, pois foram considerados como se tivessem adotado o PRV. Os três Estados juntos representam aproximadamente 5% do total de alunos, professores e diretores. Tais exclusões não comprometeram o resultado da pesquisa.

No total 7 Estados foram excluídos do grupo de controle, considerando os que têm PRV e aqueles em que não foi possível confirmar se há PRV.

O resultado encontrado foi o seguinte:

Tabela 34- Estados do grupo de controle

Código Estado	Estado	Nº de Escolas Estaduais	
		Tratados	Controles
11	Rondônia		403
14	Roraima		115
15	Pará		573

Código Estado	Estado	Nº de Escolas Estaduais	
		Tratados	Controles
16	Amapá		150
17	Tocantins		511
21	Maranhão		737
23	Ceará		505
24	Rio Grande do Norte		629
27	Alagoas		361
28	Sergipe		393
29	Bahia		1.179
33	Rio de Janeiro		1.275
35	São Paulo	5.348	
41	Paraná		1.423
42	Santa Catarina		1.425
43	Rio Grande do Sul		2.854
50	Mato Grosso do Sul		585
51	Mato Grosso		741
52	Goiás		1.088
53	Distrito Federal		418
	Total	5.348	15.365

Fonte: Elaboração do autor com base nos dados do INEP

Após a depuração dos dados obtidos no site do INEP, com exclusão das observações que continham erros de preenchimento ou estavam incompletas (*missing*), nestes Estados foram encontradas, baseado nos dados de 2007, os seguintes quantitativos de observações em condições de serem avaliadas:

Tabela 35- Dados disponíveis em 2007

	São Paulo	19 Estados	Total
Diretores	5.348	15.365	20.713
Professores	38.434	61.835	100.269
Alunos	608.915	840.009	1.448.924

Fonte: Elaboração do autor com base nos dados do INEP

As variáveis controladas são as mesmas utilizadas nos testes com os municípios, porém, foi utilizado o dado por Estado para o PIB per capita a preços constantes.

Além destes controles, foi feito o controle por Estado visando neutralizar possíveis diferenças nas características sociais da população de cada Estado, pois a análise dos dados demonstraram que Estados da região norte e nordeste tiveram evolução maior na nota do aluno, porém médias inferiores. Para estes Estados, com uma mesma política visando incentivar a

melhoria do desempenho do aluno, seria possível conseguir avanços mais significativos do que em Estados do Sul/Sudeste. Porém, quando a nota dos alunos daqueles Estados estiverem no mesmo patamar destes, possivelmente apresentarão progressos semelhantes. Os rendimentos são crescentes, mas a taxas decrescentes.

Este entendimento está alinhado com o praticado pelo Saresp no pagamento do bônus da educação, pois mantém entre os premiados as escolas que estavam e continuam entre os 10% mais bem colocados, mesmo que tenham piorado seu desempenho.

Os testes de robustez foram realizados seguindo a mesma sequência dos testes com os dados municipais, ou seja, são apresentados primeiramente os testes do diretor, seguidos do professor e aluno.

### 10.1. Testes sobre o desempenho dos diretores

Utilizando-se das mesmas variáveis dependentes empregadas nos testes com os municípios, pode-se observar a proporção de respostas dos diretores nas questões identificadas em 2007, 2009 e 2011.

Tabela 36- Variáveis Dependentes do Diretor

Proporção de Diretores com a Característica	2007		2009		2011	
	%	Nº Obs	%	Nº Obs	%	Nº Obs
22. Promove atividades de formação continuada	63,8	20.155	68,4	15.408	70,1	17.834
24. Reúne o conselho de escola mais de 3 vezes no ano	80,4	20.458	85,4	15.573	82,1	17.866
29. Reúne o conselho de classe mais de 3 vezes no ano	77,5	20.314	77,8	15.618	80,2	17.957
30. Trabalhou no projeto pedagógico da escola pessoalmente	64,9	20.361	63,1	15.480	72,8	17.802
43. Desenvolve reforço de aprendizagem para os alunos	74,3	20.144	77,5	15.682	80,3	18.010
94. Escolha conjunta dos livros didáticos	12,9	20.539	14,6	15.057	12,0	17.959

Fonte: Elaboração do autor com base nos dados do INEP

Nota-se nestas variáveis de interesse que a proporção dos diretores não ficou concentrada próxima a 100%, o que foi constatado nos testes utilizando a base de dados com escolas municipais, em específico nas questões 29 e 43, que se apresentaram em 2007 nos municípios com mais de 96% de diretores afirmando que realizam estas atividades. Nos Estados estas proporções caíram para 77,5% e 74,3% respectivamente.

O próximo passo é aplicar o método de Pareamento (*matching*), que agrupa os indivíduos de acordo com suas características, representadas pelo vetor de covariadas, pareados pelo vizinho mais próximo (*nearest neighbor*), sem reposição.

Após o emparelhamento, do universo de 57.145 observações possíveis, 37.057 foram selecionadas pelo método do *propensity score*. Nota-se que houve um maior número de escolas pareadas do que nos testes dos municípios. Como os Estados em geral assumem preferencialmente as séries mais avançadas, deixando o primeiro ciclo do ensino fundamental para os municípios, ao parear entre Estados encontrou-se escolas com características semelhantes em maior quantidade.

O resultado se apresentou com a seguinte composição:

Tabela 37- Área de Suporte Comum - Diretores.

Condição	Suporte Comum		Total
	Fora	Dentro	
Controles	0	25.947	25.947
Tratados	3	11.107	11.110
Total	3	37.054	37.057

Fonte: Elaboração do autor com base nos dados do INEP

Também foram realizados os testes de médias para avaliar se as diferenças entre grupo de tratamento e controle no período de pré-tratamento (2007) eram estatisticamente distintas, o que não deveria ocorrer, pois o PRV ainda não havia sido implantado, o que veio acontecer somente em 2008. Novamente os testes demonstraram que havia diferenças, indicando a existência de variáveis não observáveis afetando o resultado do pareamento. Os resultados dos testes estão disponíveis para consulta.

Diante desta situação, os efeitos das características não observáveis foram eliminados utilizando o estimador Diferenças em Diferenças (DD) a partir do seguinte modelo:

$$Y_{ist} = \beta_0 + \beta_1 EESP_s + \beta_2 Ano_t + \beta_3 EESP_s * Ano_t + B_4 A_{vst} + B_5 X_{ist} + \varepsilon_{ist}$$

Onde:

$Y_{ist}$  é a variável dependente e identifica o efeito do modelo no desempenho do diretor, no qual  $i$ ,  $s$ ,  $t$  denotam o diretor, se é uma escola estadual ou municipal e se é referente a antes ou depois da implantação do programa;

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

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

O  $\beta_3$  é o coeficiente calculado da variável de interesse estimada pela interação das duas variáveis anteriores: se é escola estadual paulista e se a observação se refere ao período posterior à implantação do programa. Esta variável mensurará o efeito do PRV sobre o desempenho dos diretores;

$A_{vst}$  vetor de covariadas das 7 características que controlam o entorno do aluno (ambiente socioeconômico);

$X_{ist}$  é o vetor de covariadas com as demais características do diretor;

$\varepsilon_{ist}$  é o resíduo da regressão

### 10.1.1. Resultados dos diretores

Com base na seleção obtida pelo pareamento por escore de propensão, nesta seção serão apresentados os resultados da estimação pelo método de Diferenças em Diferenças. A variável de interesse é uma *dummy* resultante da interação do Ano (0 se antes ou 1 se depois do programa) com o Estado (0 se escola for dos 19 Estados ou 1 se Estadual paulista). O coeficiente desta variável indica o efeito no desempenho dos indivíduos tratados, ou seja, diretor participante do programa de remuneração variável das escolas estaduais paulista após a implantação do programa em 2008.

A tabela com os resultados dos Estados se apresenta comparativa com os resultados obtidos nos Municípios para testar a robustez dos resultados.

Tabela 38- Coeficientes estimados - Diretores



Variáveis Dependentes	Municípios	Estados
22. Promove atividades de formação continuada	0,0134 (0,0317)	0,0296 *** (0,0112)
24. Reúne o conselho de escola mais de 3 vezes no ano	-0,0858 *** (0,0222)	-0,0580 *** (0,0077)
29. Reúne o conselho de classe mais de 3 vezes no ano	0,0135 (0,0149)	-0,0063 (0,0057)
30. Trabalhou no projeto pedagógico da escola pessoalmente	-0,1009 *** (0,0328)	-0,0840 *** (0,0116)
43. Desenvolve reforço de aprendizagem para os alunos	-0,0267 * (0,0137)	-0,0235 *** (0,0071)
94. Escolha conjunta dos livros didáticos	0,0767 *** (0,0282)	0,0417 *** (0,0088)

Fonte: Elaboração do autor com base nos dados do INEP

(1) Desvio padrão entre parênteses

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

Os testes de robustez mostraram que a replicação dos testes utilizando outro grupo de controle e ainda ampliando o grupo de tratamento para todas as escolas estaduais paulista que puderam ser pareadas gerou resultados semelhantes, com exceção da questão 22, que passou a ser significativa apenas nos Estados.

Os demais coeficientes dos testes são similares entre si, sendo que aqueles significantes nos municípios se mantiveram nos Estados, inclusive respeitando o mesmo sinal (positivo ou negativo) e com coeficientes próximos entre si.

Em outras palavras, os resultados são semelhantes sejam eles obtidos pela comparação de diretores de escolas estaduais com de escolas municipais em 38 municípios paulistas selecionados ou pela comparação dos diretores de escolas estaduais paulistas com diretores de escolas de outros 19 Estados.

## 10.2. Testes sobre o desempenho dos professores

Para testar a robustez dos resultados dos professores, utilizou-se das mesmas variáveis empregadas nos testes com os municípios. Com base na tabela a seguir pode-se observar a proporção de respostas dos professores nos Estados selecionados, conforme apresentado na seção 10, nas questões identificadas em 2007, 2009 e 2011.

Tabela 39- Atividades Desenvolvidas em Língua Portuguesa

Proporção de Professores que realizam a atividade	2007		2009		2011	
	%	Nº Obs	%	Nº Obs	%	Nº Obs
24. Copiar textos extensos do livro didático ou do quadro	31,7	56.461	68,0	45.842	60,9	44.450
25. Conversar sobre textos de jornais e revistas	89,6	56.583	86,6	45.823	88,1	43.900
26. Fazer exercícios de gramática relacionados com textos de jornais ou revistas	84,2	56.267	72,4	44.484	86,4	43.235
27. Questão sem correspondente em 2011	78,8	56.059	70,7	44.099	77,1	42.930
28. Ler, discutir com os colegas e escrever textos relacionados com o desenvolvimento de projeto temático	87,8	56.410	72,1	44.059	89,1	42.998
29. Ler contos, crônicas, poesias ou romances	86,4	56.061	71,5	44.007	88,0	42.845
30. Conversar sobre contos, crônicas, poesias ou romances	84,6	56.116	68,2	43.930	86,7	42.763
31. Usar contos, crônicas, poesias ou romances para exercitar aspectos da gramática	75,0	55.773	63,4	43.802	82,5	42.688
32. Fixar nomes de conceitos gramaticais e linguísticos	93,9	55.994	80,2	43.926	94,0	42.744
33. Discutir um texto, explorando as diferenças entre fatos e opiniões	31,7	56.461	68,0	45.842	60,9	44.450

Fonte: Elaboração do autor com base nos dados do INEP

Tabela 40- Atividades Desenvolvidas em Matemática

Proporção de Professores que realizam a atividade	2007		2009		2011	
	%	Nº Obs	%	Nº Obs	%	Nº Obs
34. Fazer exercícios para automatizar procedimentos	90,8	56.163	90,7	46.800	95,1	47.250
35. Lidar com problemas que exigem raciocínios diferentes e mais complexos que a maioria dos exemplos usuais	93,8	56.660	92,8	46.446	95,8	46.535

Proporção de Professores que realizam a atividade	2007		2009		2011	
	%	Nº Obs	%	Nº Obs	%	Nº Obs
36. Falar sobre suas soluções, discutindo caminhos usados para encontra-las	96,9	56.158	94,7	46.346	97,6	46.400
37. Gravar as regras que permitem obter as respostas certas dos cálculos e problemas	82,0	55.778	72,9	44.235	87,3	45.255
38. Lidar com temas que aparecem em jornais e/ou revistas, discutindo a relação dos temas com a Matemática.	73,8	55.776	72,3	44.009	79,5	44.961
39. Interpretar resultados numéricos obtidos para dar uma resposta adequada ao problema.	95,1	55.887	81,8	43.602	94,0	44.820
40. Lidar com situações que lhes sejam familiares e que apresentem temas do interesse dos alunos.	92,2	55.837	91,5	44.473	92,7	45.204
41. Experimentar diferentes modos de resolver um problema ou de efetuar um cálculo.	96,4	55.976	83,2	43.622	94,4	44.772
42. Aprimorar a precisão e a velocidade de execução de cálculos.	86,8	55.743	75,4	43.387	87,7	44.656
43. Experimentar diferentes ações (coletar informações recortar, analisar, explorar, discutir, manipular etc.) para resolver problemas.	80,5	56.013	78,7	43.864	84,7	44.834

Fonte: Elaboração do autor com base nos dados do INEP

A questão 55, que foi utilizada para avaliar o tanto o professor de português quanto de matemática, se apresenta da seguinte forma:

Tabela 41- Conteúdo Programático Desenvolvido

Proporção de Professores que realizam a atividade	2007		2009		2011	
	%	Nº Obs.	%	Nº Obs.	%	Nº Obs.
55. Ministrou menos de 60% do conteúdo programático	17,2	98.355	19,4	49.313	15,3	67.474

Fonte: Elaboração do autor com base nos dados do INEP

Em ambos os quadros apresentados, nota-se que há predominância das respostas em que o respondente afirma que realiza a atividade, porém esta predominância é menos acentuada do que foi notada na base de dados dos municípios, ou seja, a proporção de professores que executam cada uma das atividades listadas é menor, via de regra.

Esta distribuição indica haver condições para mudança do percentual dos professores que executam cada atividade nas avaliações dos anos seguintes (2009 e 2011).

No pareamento pelo vizinho mais próximo (*nearest neighbor*), sem reposição, do universo de 244.787 observações possíveis, 8.134 foram selecionadas pelo método do *propensity score*. Nota-se que também houve um maior número de professores pareados do que nos testes dos municípios. Como os Estados em geral assumem preferencialmente as séries mais avançadas, deixando o primeiro ciclo do ensino fundamental para os municípios, ao parear entre Estados encontrou-se escolas com características semelhantes em maior quantidade.

O resultado se apresentou com a seguinte composição:

Tabela 42- Área de Suporte Comum - Professores.

	Condição	Suporte Comum		Total
		Fora	Dentro	
Prof.	Controles	0	4.980	25.947
Português	Tratados	39	3.115	11.110
4ª série	Total	39	8.095	8.134
Prof.	Controles	0	9.843	9.843
Português	Tratados	98	5.405	5.503
8ª série	Total	98	15.248	15.346
Prof.	Controles	0	4.943	4.943
Matemática	Tratados	41	3.085	3.126
4ª série	Total	41	8.028	8.069
Prof.	Controles	0	9.994	9.994
Matemática	Tratados	6	5.847	5.853
8ª série	Total	6	15.841	15.847

Fonte: Elaboração do autor com base nos dados do INEP

Foram realizados os testes de médias para avaliar se as diferenças entre grupo de tratamento e controle no período de pré-tratamento (2007) eram estatisticamente distintas. Os

testes demonstraram que havia diferenças, indicando a existência de variáveis não observáveis afetando o resultado do pareamento. Os resultados dos testes estão disponíveis para consulta.

Diante desta situação, eliminou-se os efeitos das características não observáveis utilizando o estimador Diferenças em Diferenças (DD) a partir do seguinte modelo:

$$Y_{igst} = \beta_0 + \beta_1 EESP_s + \beta_2 Ano_t + \beta_3 EESP_s * Ano_t + B_4 E_{st} + B_5 D_{vst} + B_6 A_{vst} + B_7 X_{ist} + \varepsilon_{ist}$$

Onde:

$Y_{igst}$  é a variável dependente e identifica o efeito do modelo no desempenho do professor, no qual  $i, g, s, t$  denotam o professor, a série, se é uma escola estadual ou municipal e se é referente a antes ou depois da implantação do programa;

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

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

O  $\beta_3$  é o coeficiente calculado da variável de interesse, estimada pela interação das duas variáveis anteriores: se é escola estadual paulista e se a observação se refere ao período posterior à implantação do programa. Esta variável mensurará o efeito do PRV sobre o desempenho dos professores;

$E_{st}$  vetor de covariadas com 11 características que controlam as condições de uso da escola (manutenção, computadores e segurança)

$D_{vst}$  vetor de 6 covariadas que avaliam o desempenho do diretor de uma escola estadual paulista depois do programa;

$A_{vst}$  vetor de covariadas das 7 características que controlam o entorno do aluno (ambiente socioeconômico);

$X_{ist}$  é o vetor de covariadas com as demais características do professor;

$\varepsilon_{ist}$  é o resíduo da regressão.

### 10.2.1. Resultados dos professores

Com base na seleção obtida pelo pareamento por escore de propensão, pode-se estimar os resultados pelo método de Diferenças em Diferenças. O coeficiente da *dummy* resultante da interação do Ano (0 se antes ou 1 se depois do programa) com o Estado (0 se escola for dos 19 Estados ou 1 se Estadual paulista), indica o efeito no desempenho dos indivíduos tratados, ou seja, professor participante do programa de remuneração variável das escolas estaduais paulista após a implantação do programa em 2008.

A tabela com os resultados dos Estados se apresenta comparativa com os resultados obtidos nos Municípios para testar a robustez dos resultados, sendo primeiramente apresentado o comparativo dos professores de português, tanto da 4ª quanto da 8ª série, e na tabela seguinte os resultados dos professores de matemática.

Tabela 43- Coeficientes estimados - Professores Português

Variáveis Dependentes	4ª série		8ª série	
	Municípios	Estados	Municípios	Estados
24. Copiar textos extensos do livro didático ou do quadro	-0,0563 (0,0351)	-0,0519 ** (0,0249)	-0,0764 (0,0685)	-0,0595 *** (0,0170)
25. Conversar sobre textos de jornais e revistas	0,0046 (0,0151)	0,0294 ** (0,0129)	0,1727 *** (0,0597)	0,0250 * (0,0134)
26. Fazer exercícios de gramática relacionados com textos de jornais ou revistas	-0,0558 ** (0,0226)	-0,0231 (0,0170)	0,0993 (0,0629)	0,0615 *** (0,0155)
28. Ler, discutir com os colegas e escrever textos relacionados com o desenvolvimento de projeto temático	0,0188 (0,0228)	0,0400 ** (0,0184)	0,0560 (0,0735)	0,0622 *** (0,0172)
29. Ler contos, crônicas, poesias ou romances	-0,0154 (0,0126)	0,0043 (0,0114)	-0,0722 (0,0595)	0,0035 (0,0141)
30. Conversar sobre contos, crônicas, poesias ou romances	-0,0047 (0,0142)	-0,0051 (0,0132)	-0,0795 (0,0640)	-0,0032 (0,0143)
	-0,0416 **	-0,0195	-0,0630	0,0174

Variáveis Dependentes	4ª série		8ª série	
	Municípios	Estados	Municípios	Estados
31. Usar contos, crônicas, poesias ou romances para exercitar aspectos da gramática	(0,0211)	(0,0168)	(0,0661)	(0,0152)
32. Fixar nomes de conceitos gramaticais e linguísticos	-0,1006 *** (0,0350)	-0,0254 (0,0234)	-0,1360 * (0,0750)	0,0294 * (0,0174)
33. Discutir um texto, explorando as diferenças entre fatos e opiniões	0,0111 (0,0114)	0,0143 (0,0096)	0,0739 (0,0537)	0,0434 *** (0,0117)
55. Ministrou menos de 60% do conteúdo programático	-0,0214 (0,0258)	-0,0676 *** (0,0193)	0,0195 (0,0531)	-0,0087 (0,0143)

Fonte: Elaboração do autor com base nos dados do INEP

(1) Desvio padrão entre parênteses

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

Os testes de robustez mostraram que a reaplicação dos testes utilizando outro grupo de controle e ainda ampliando o grupo de tratamento para todas as escolas estaduais paulistas que puderam ser pareadas gerou alguns resultados que não podem ser afirmados com confiança como sendo semelhantes. Uma das possibilidades que justificaria as diferenças seria a adoção por parte dos Estados de políticas educacionais distintas e que não puderam ter seus efeitos capturados pelas variáveis observáveis.

Há coeficientes significantes similares, sejam dos Estados ou dos municípios, sendo que aqueles positivos nos municípios se mantiveram positivos e próximos nos Estados. O mesmo pode ser observado na comparação dos coeficientes negativos.

Para os coeficientes que se apresentaram não significantes observa-se que as questões 29, 31 e 55 da 8ª série há inversão do sinal de negativo para positivo, se município ou estado. Dado que os coeficientes se apresentam muito pequenos, e por estarem muito próximos ao zero, são sensíveis para mais ou menos a qualquer alteração na base de dados, por menor que seja. Neste caso, pode-se afirmar que a inversão de sinais não é relevante.

Tabela 44- Coeficientes estimados - Professores Matemática

Variáveis Dependentes	4ª série		8ª série	
	Municípios	Estados	Municípios	Estados
34.Fazer exercícios para automatizar procedimentos	-0,0411 (0,0294)	-0,0014 (0,0184)	0,0389 (0,0379)	-0,0022 (0,0094)
35.Lidar com problemas que exigem raciocínios diferentes e mais complexos que a maioria dos exemplos usuais	-0,0066 (0,0136)	0,0063 (0,0099)	-0,0237 (0,0346)	-0,0164 * (0,0093)
36.Falar sobre suas soluções, discutindo caminhos usados para encontrá-las	0,0020 (0,0081)	0,0136 (0,0077)	* -0,0217 (0,0201)	-0,0033 (0,0072)
37.Gravar as regras que permitem obter as respostas certas dos cálculos e problemas	-0,0980 *** (0,0356)	-0,0539 ** (0,0236)	-0,0789 (0,0562)	-0,0044 (0,0143)
38.Lidar com temas que aparecem em jornais e/ou revistas, discutindo a relação dos temas com a Matemática.	0,0218 (0,0247)	0,0156 (0,0180)	0,1109 (0,0703)	-0,0265 (0,0176)
39.Interpretar resultados numéricos obtidos para dar uma resposta adequada ao problema.	-0,0153 (0,0131)	0,0291 *** (0,0099)	0,0641 * (0,0338)	0,0363 *** (0,0099)
40.Lidar com situações que lhes sejam familiares e que apresentem temas do interesse dos alunos.	-0,0015 (0,0099)	0,0081 (0,0086)	-0,0515 (0,0451)	-0,0198 * (0,0115)
41.Experimentar diferentes modos de resolver um problema ou de efetuar um cálculo.	-0,0041 (0,0076)	0,0302 *** (0,0079)	0,0139 (0,0365)	0,0228 ** (0,0099)
42.Aprimorar a precisão e a velocidade de execução de cálculos.	-0,0472 * (0,0267)	0,0203 (0,0182)	0,0167 (0,0585)	0,0171 (0,0145)
43.Experimentar diferentes ações (coletar informações recortar, analisar, explorar, discutir, manipular etc.) para resolver problemas.	0,0009 (0,0192)	0,0129 (0,0148)	0,0651 (0,0638)	-0,0007 (0,0163)



Variáveis Dependentes	4ª série		8ª série	
	Municípios	Estados	Municípios	Estados
55.Ministrou menos de 60% do conteúdo programático	-0,0334 (0,0261)	-0,0670 *** (0,0196)	0,0181 (0,0532)	0,0227 (0,0142)

Fonte: Elaboração do autor com base nos dados do INEP

(1) Desvio padrão entre parênteses

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

Os mesmos comentários utilizados na tabela com os resultados do professor de português se aplicam a esta última. Novamente os testes de robustez mostraram que a replicação dos testes utilizando outro grupo de controle e ainda ampliando o grupo de tratamento para todas as escolas estaduais paulista que puderam ser pareadas gerou alguns resultados que não podem ser afirmados com confiança como sendo semelhantes. Uma das possibilidades que justificaria as diferenças seria a adoção por parte dos Estados de políticas educacionais distintas e que não puderam ter seus efeitos capturados pelas variáveis observáveis.

Há coeficientes significantes similares, sejam dos Estados ou dos municípios, sendo que aqueles positivos nos municípios se mantiveram positivos e próximos nos Estados. O mesmo pode ser observado na comparação dos coeficientes negativos.

### 10.3. Testes sobre o desempenho dos alunos

Para saber se as notas dos alunos da rede estadual paulista é um bom instrumento de avaliação, é necessário comparar seus desempenhos em relação às notas dos alunos dos demais estados participantes do teste para saber se há condições de mobilidade na escala ou se já se encontram no limite superior, o que caracteriza um situação em que as melhorias são mais difíceis de serem conseguidas e, quando se consegue, são mais tímidas.

O resultado comparativo das notas médias dos alunos paulistas com os outros estados se apresentou da seguinte forma:

Tabela 45- Proficiência comparativa do Aluno 4ª série entre Estados

Variável	2007		2009		2011	
	SP	Outros	SP	Outros	SP	Outros
Língua Portuguesa	176,92	174,35	189,31	182,08	192,37	189,08
Matemática	193,97	190,98	212,88	200,62	213,92	206,55

Fonte: Elaboração do autor com base nos dados do INEP

Tabela 10.13 – Proficiência comparativa do Aluno 8ª série entre Estados

Variável	2007		2009		2011	
	SP	Outros	SP	Outros	SP	Outros
Língua Portuguesa	231,95	228,44	240,31	238,74	242,07	238,55
Matemática	242,60	240,09	242,85	241,73	245,59	245,48

Fonte: Elaboração do autor com base nos dados do INEP

Os dados apresentados indicam que as notas dos alunos paulistas estão acima da média nacional, diferente da comparação com os 38 municípios paulistas, que estava abaixo. Porém, se for analisado os dados individualizados por estado, nota-se que os estados da região sul, mais destacadamente o Estado do Paraná, possuem notas superiores aos paulistas. Portanto, as notas dos alunos paulistas não se encontram no limite superior, havendo condições de serem melhoradas, podendo ser utilizadas como instrumento de avaliação de desempenho dos alunos paulistas.

Na sequência, utilizando-se das mesmas variáveis dependentes empregadas nos testes com os municípios, pode-se observar a proporção de respostas dos alunos nas questões identificadas em 2007.

Tabela 46- Variáveis dependentes – Aluno ano de 2007

Série	Disciplina	Obs	Média	Desvio Padrão	Mínimo	Máximo
4ª	Português	539.414	175,43	41,55	73,50	347,37
	Matemática	539.414	192,23	42,95	86,98	367,54
8ª	Português	909.510	229,93	43,90	112,66	395,89
	Matemática	909.510	241,16	42,57	138,74	417,11

Fonte: Elaboração do autor com base nos dados do INEP

O comportamento das variáveis de interesse se mantém semelhante aos dos municípios.

Na sequência, reaplicou-se o método de Pareamento (*matching*), agrupando os indivíduos de acordo com suas características, representadas pelo vetor de covariadas, pareados pelo vizinho mais próximo (*nearest neighbor*), sem reposição.

Após o emparelhamento, do universo de 4.412.671 observações possíveis, 504.632 foram selecionadas pelo método do *propensity score*, o que significa um aumento sensível se comparado com o número de alunos pareados nos testes dos municípios, dado a vocação dos Estados em assumirem preferencialmente as séries mais avançadas, deixando o primeiro ciclo do ensino fundamental para os municípios.

O resultado se apresentou com a seguinte composição:

Tabela 47- Área de Suporte Comum - Alunos.

Controle	Série	Condição	Suporte Comum		Total
			Fora	Dentro	
Diretor	4ª Port e Mat	Controles	0	261.236	261.236
		Tratados	104	243.292	243.396
		Total	104	504.528	504.632
	8ª Port e Mat	Controles	0	710.765	710.765
		Tratados	59	632.030	632.089
		Total	59	1.342.795	1.342.854
Professor	4ª Port	Controles	0	179.718	179.718
		Tratados	40.983	179.718	220.701
		Total	40.983	359.436	400.419
	8ª Port	Controles	0	627.214	627.214
		Tratados	99	607.986	608.085
		Total	99	1.235.200	1.235.299
	4ª Mat	Controles	0	179.545	179.545
		Tratados	41.174	179.545	220.719
		Total	41.174	359.090	400.264
	8ª Mat	Controles	0	629.674	629.674
		Tratados	92	608.027	608.119
		Total	92	1.237.701	1.237.793

Fonte: Elaboração do autor com base nos dados do INEP

Na realização dos testes de médias para avaliar se havia diferenças entre grupo de tratamento e controle no período de pré-tratamento (2007) eram estatisticamente distintas, demonstrou haver médias diferentes, indicando a existência de variáveis não observáveis afetando o resultado do pareamento. Os resultados dos testes estão disponíveis para consulta.

Por conta dessa constatação, eliminou-se os efeitos das características não observáveis utilizando o estimador Diferenças em Diferenças (DD) a partir do seguinte modelo:

$$Y_{igdst} = \beta_0 + \beta_1 EESP_s + \beta_2 Ano_t + \beta_3 EESP_s * Ano_t + B_4 Y_{ivst} + B_5 X_{ist} + \varepsilon_{ist}$$

Onde:

$Y_{igdst}$  é a variável dependente e identifica o efeito do modelo no desempenho do aluno, no qual  $i, g, d, s, t$  denotam o aluno, a série, a disciplina avaliada, se é uma escola estadual ou municipal e se é referente a antes ou depois da implantação do programa;

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

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

O  $\beta_3$  é o coeficiente calculado da variável de interesse, estimada pela interação das duas variáveis anteriores: se é escola estadual paulista e se a observação se refere ao período posterior à implantação do programa. Esta variável mensurará o efeito do PRV sobre o desempenho dos professores;

$Y_{ivst}$  é o vetor em que  $i$  identifica se diretor ou professor,  $v$  será uma das variáveis dependentes do modelo do diretor ou professor,  $s$  a escola e  $t$  o ano correspondente;

$X_{ist}$  é o vetor de covariadas com as demais características do aluno;

$\varepsilon_{ist}$  é o resíduo da regressão

### 10.3.1. Resultados dos Alunos

Os resultados da estimação pelo método de Diferenças em Diferenças são observados através da *dummy* resultante da interação do Ano (0 se antes ou 1 se depois do programa) com o Estado (dependência administrativa da escola: 0 se municipal ou 1 se estadual). O coeficiente desta variável indica 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 individualizadamente nas proficiências de português e matemática, que são as variáveis dependentes, tanto na 4ª quanto na 8ª série, 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 covariadas específicas do aluno, controla-se o experimento também pelo PIB per capita estadual, 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 com os resultados dos Estados é apresentada de forma a facilitar a comparação com os resultados obtidos nos Municípios para testar a robustez dos resultados da variável de interesse (Ano\*Estado) nas variáveis dependentes.

Tabela 48- Coeficientes estimados – Aluno

Variáveis Dependentes	Controle	4ª série		8ª série	
		Municípios	Estados	Municípios	Estados
Proficiência Língua Portuguesa	Diretor	0,3965 (0,3502)	0,0043 (0,2226)	-4,4271 *** (0,7818)	-0,7125 *** (0,1448)
	Professor	-0,4643 (0,3570)	-0,8677 *** (0,2623)	-4,6135 *** (0,9209)	-0,9826 *** (0,1510)
Proficiência Matemática	Diretor	4,1737 *** (0,3700)	4,4020 *** (0,2332)	-5,5710 *** (0,7873)	-1,5351 *** (0,1422)
	Professor	3,3059 *** (0,3755)	3,3101 *** (0,2736)	-6,7245 *** (0,9457)	-1,5638 *** (0,1471)

Fonte: Elaboração do autor com base nos dados do INEP

(1) Desvio padrão entre parênteses

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

A reaplicação dos testes utilizando outro grupo de controle e ainda ampliando o grupo de tratamento para todas as escolas estaduais paulista que puderam ser pareadas gerou resultados semelhantes.

Os coeficientes dos estados e municípios são similares, sendo que aqueles positivos para municípios se mantiveram positivos e próximos nos Estados. O mesmo pode ser observado na comparação dos coeficientes negativos e para o único não significativo, que é proficiência em língua portuguesa da 4ª série. Até mesmo no nível de significância eles são iguais para todas as variáveis.

Em outras palavras, os desempenhos são semelhantes sejam eles obtidos pela comparação de alunos de escolas estaduais com de escolas municipais em 38 municípios paulista selecionados ou pela comparação de aluno de escolas estaduais paulistas com alunos de escolas de outros 19 Estados, demonstrando que os resultados obtidos são robustos.

## 11. Conclusões e Recomendações

Esta pesquisa teve como objetivo analisar os efeitos dos incentivos vinculados à programas de remuneração variável sobre o desempenho do servidor público. Nesta conclusão, primeiramente será feita uma resenha de como se desenvolveu a pesquisa, passando em seguida para seus resultados, análises e recomendações.

Inicialmente foi realizada uma revisão da literatura objetivando identificar teorias associada ao tema. Esta indicou que Programas de Remuneração Variável – PRV podem contribuir para reduzir assimetrias entre agente e principal, minimizando os conflitos de agência, desde que observadas algumas recomendações já consolidadas pela literatura quanto à forma e grau em que os incentivos são disponibilizados aos trabalhador.

Baseados na teoria dos incentivos, constatou-se que a motivação extrínseca, que refere-se a fazer algo porque isso leva o indivíduo a um resultado diferenciado e não porque é interessante ou agradável, ajuda a compreender as variações no desempenho dos adultos, fase em que a limitação das liberdades de escolha é maior por conta das demandas sociais e familiares, que exigem que os indivíduos assumam responsabilidades por tarefas não intrinsecamente interessantes.

Também foi possível conhecer a evolução e os modelos de recompensa para o trabalhador, com a recomendação de que o PRV, na visão moderna, deve associar as competências do trabalhador aos objetivos estratégicos da organização. Desta forma, as competências podem ser vinculadas aos resultados da organização (*outcomes*), podendo ser mensurados os impactos causados pela variação no desempenho do trabalhador.

Quanto aos modelos de PRVs indicados pela literatura, as recomendações básicas mais comumente notadas sugerem: alinhar a estratégia organizacional com os resultados considerados importantes; avaliar os resultados individuais e corporativos de forma equilibrada; simplificar o modelo de forma que possa ser facilmente entendido; definir indicadores precisos e metas factíveis; estabelecer linha direta de comunicação com os envolvidos e; preparar os gestores e a organização para implantar o PRV.

Na revisão bibliográfica também foi possível confirmar que há dubiedade nos resultados dos diversos estudos que analisaram PRV, alternando resultados positivos com negativos e neutros. Nestes, foram encontrados resultados qualitativos positivos com maior frequência se comparados com os estudos que utilizaram métodos quantitativos para avaliação dos impactos. Ainda assim, os resultados quantitativos positivos são maioria. Identificou-se que as empresas,

por objetivarem o lucro, conseguem associar o desempenho dos trabalhadores mais facilmente aos seus interesses, porém, no serviço público, os resultados são mais difusos e difíceis de mensurar, o que dá margem a programas que custam muito para os governos sem gerar impactos expressivos na motivação do servidor, podendo até mesmo desmotivá-lo.

Na literatura que trata das experiências brasileiras, há poucos casos relatados e somente em Oshiro (2011) foi encontrado com avaliação quantitativa dos impactos causados pelo PRV, que apresentou resultados parcialmente positivos.

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.

Para verificação da hipótese foram realizados testes no PRV implantado em 2008, que prevê o pagamento de bônus aos diretores e professores da rede estadual paulista. Os testes basearam-se nos dados obtidos na Prova Brasil aplicada nos anos de 2007, antes do programa, que foi implantado em 2008, e nos dados das provas aplicadas em 2009 e 2011, depois da implantação do programa. Também foi testado se o desempenho dos diretores e professores refletiu no desempenho dos alunos.

A base de dados original utilizada para o desenvolvimento dos testes é composta de 47.596 diretores, 292.828 professores e 4.109.265 alunos. O número de características observáveis identificadas a partir dos questionários específicos de cada elemento avaliado foram de 47 para diretores, 60 para professores e 67 para os alunos. Esta ampla base de dados encontrada e a expressiva quantidade de características observáveis permitiram a realização de testes que apresentassem resultados mais confiáveis.

Os grupos de tratamento, composto por aqueles que estavam sob a influência do PRV, foram formados pelos diretores, professores e alunos das escolas estaduais paulista. Os grupos de controle foram formados pelos mesmos elementos, porém pertencentes às escolas dos 38 municípios paulistas identificados com população superior a 200 mil habitantes. As análises foram segmentadas para cada grupo de elemento avaliado.

Para avaliar seus desempenhos foi utilizada a metodologia de pareamento por escore de propensão, que permitiu comparar indivíduos tratados e não tratados (controles) com características semelhantes. Sobre o resultado obtido por este método, foi aplicada uma segunda metodologia denominada Diferenças em Diferenças, que permitiu capturar os efeitos das características não observáveis que não variam no tempo.

Uma vez apresentada um breve relato do desenvolvimento desta pesquisa, passa-se para a análise geral dos resultados, seguidos das recomendações.



## 11.1. Desempenho dos diretores

Os resultados no desempenho dos diretores foram obtidos pela avaliação de 6 características independentes, identificadas a partir do questionário próprio e seguindo as recomendações dos estudos apresentados por Sammons, Hillman & Mortimore (1995), Coleman, Collinge & Scifert (1993) e em Hanushek (2003), que tratam da efetividade das escolas.

Os coeficientes indicaram efeitos negativos e significantes em 3 delas, sendo que entre estas os resultados mais expressivos foram encontrados nas questões que avaliavam o desempenho quanto à “promoção de reuniões do conselho de escola”, que apresentou redução de 8,58%, e se “trabalhou no projeto pedagógico da escola”, indicando que participar do PRV paulista reduz o desempenho do diretor em 10,09% se comparado com os diretores de escolas municipais, que não participaram do PRV. Menos negativo se apresenta o resultado do desempenho do diretor quanto ao “desenvolvimento de reforço na aprendizagem do aluno”, que indicou redução de 2,67%.

Por estes resultados é possível afirmar que o PRV serviu para desestimular a melhora do desempenho nestes quesitos avaliados, contrariando o propósito do programa.

Na análise das possíveis justificativas, 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 & Deci (2000), os jovens são mais preocupados em obter altos salários, mas à medida 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 com ressalvas já que não é incorporado ao valor pago na aposentadoria, porém mais estudos precisam ser feitos para testar esta suposição.

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

(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 estar se sentindo 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, equipara-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.

Em oposição aos resultados anteriores, a atividade de “escolher de forma conjunta os livros didáticos para os alunos”, indicou efeito positivo e significativo. Ser participante do PRV proporciona 7,67% de aumento no desempenho se comparado com os não participantes.

Esse resultado positivo isolado requer que mais pesquisas sejam feitas para confirmá-lo. Pode ser que variáveis não observáveis, que não se mantiveram fixas no tempo, estejam influenciando este resultado, como por exemplo, a adoção durante esse período avaliado de política por parte da Secretaria Estadual de Educação atribuindo ao diretor maior responsabilidade por esta atividade.

## **11.2. Desempenho dos professores**

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. A seleção destas atividades seguiu a recomendação dada pela Taxonomia de Bloom apresentada por Krathwohl (2002). Representando estas atividades, foram selecionadas 9 questões sobre o conteúdo dado em português e 10 questões sobre o conteúdo de matemática. Além destas, a questão que avaliou o percentual do conteúdo pedagógico dado em sala de aula foi utilizada para ambas as disciplinas.

Dos 20 resultados obtidos (10 da 4ª série e outros 10 da 8ª série) intencionando avaliar o desempenho dos professores de português, 5 resultados foram significantes, sendo 4 negativos e um positivo.

Dentre os resultados negativos, 3 são relativos às atividades dos professores de 4ª série. Sendo que os professores de português da 4ª série que participaram do PRV reduziram em 5,58% a atividade “fazer exercícios de gramática relacionados com textos de jornais ou revistas”, se comparado com os professores da rede municipal de ensino dos 38 municípios que participaram desta pesquisa. A atividade de “usar contos, crônicas, poesias ou romances para exercitar aspectos da gramática” foi reduzida em 4,16% e a atividade de “fixar nomes de conceitos gramaticais e linguísticos” foi reduzida em 10,06% pelos professores. Esta última atividade também teve impacto negativo em 13,60% nos professores da 8ª série.

Notadamente o efeito para os professores de português da 4ª série não foram bons. Em nenhuma das 10 atividades avaliadas houve efeitos positivos. Na 8ª série apenas a atividade de “conversar sobre textos de jornais e revistas” teve evolução de 17,27% dos professores. Pela taxonomia de Bloom, esta atividade está relacionada a capacidade de reconhecer e reproduzir ideias e conteúdos. 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 22 itens avaliados (11 na 4ª série e 11 na 8ª), apenas 3 foram significantes, sendo que dois deles negativos. Na 4ª série a atividade de “gravar as regras que permitem obter as respostas certas dos cálculos e problemas” foi reduzida em 9,80% pelos professores de matemática da rede estadual paulista se comparado com os da rede municipal. Também nesta mesma série, a atividade de “aprimorar a precisão e a velocidade de execução de cálculos” foi reduzida em 4,72% pelos professores que participaram do PRV.

Na 8ª série foi identificado apenas um item com significância estatística, que se refere a “interpretar resultados numéricos obtidos para dar uma resposta adequada ao problema”, que proporcionou uma melhora de 6,41% na sua execução pelos professores de matemática participantes do PRV. Este é um resultado relevante, pois desenvolve no aluno, segundo a taxonomia de Bloom, a capacidade de dividir a informação em partes relevantes e irrelevantes e entender a inter-relação existente entre as partes. Atinge o nível 4 de dificuldade, na escala crescente de 1 a 5 da referida taxonomia.

Numa análise conjunta das duas disciplinas, para a 4ª série, 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.

Para a 8ª série, houve ganhos consistentes em duas atividades: “conversar sobre textos de jornais e revistas”, em português, e “interpretar resultados numéricos obtidos para dar uma resposta adequada ao problema”, em matemática. Estes mesmos resultados são considerados

robustos, pois são confirmados nos testes quando se substitui os municípios por outros estados da federação, mesmo considerando que os ganhos apurados nos estados, 2,50% e 3,63% respectivamente, são menores.

Ao se fazer uma avaliação do desempenho dos professores em geral, nota-se que das 8 atividades com resultados estatisticamente significantes, 6 foram negativos e 2 positivos, sendo que destes dois últimos, um está relacionado a uma atividade de baixo nível de dificuldade, segundo a taxonomia de Bloom. Com estes resultados não se pode afirmar que o PRV gera contribuição relevante para o desempenho do professor.

Entre as justificativas possíveis, 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 2011. 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 a 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 poderão ter 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 pela Secretaria de Educação do Estado de São Paulo foram os seguintes<sup>9</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

---

<sup>9</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>

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 (2009) pagou-se 3% do salário anual e no Quênia apresentado por Glewwe (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.

Encerrando a análise do professor, sugere-se a realização de pesquisa qualitativa com os professores e outros representantes da categoria, para identificar sua percepção quanto aos propósitos e benefícios do programa. A visão de quem está participando do PRV irá contribuir para melhor entender os resultados desta pesquisa.

### **11.3. Desempenho dos alunos**

Por fim, foi analisado o desempenho dos alunos na proficiência em português e matemática, tanto da 4ª quanto da 8ª série, com base nos resultados dos testes padronizados aplicados a cada 2 anos pelo INEP, visando identificar se as alterações no desempenho dos diretores e professores impactam o desempenho do aluno.

Nesta etapa foram realizadas duas avaliações do desempenho dos alunos. Uma controlando o desempenho dos diretores e outra replicando o mesmo teste substituindo o diretor pelo desempenho dos professores.

Das 8 avaliações realizadas, em 6 foram encontrados resultados significantes. Apenas na 4ª série em português não se pode afirmar se o PRV paulista tem algum efeito.

Os benefícios mais animadores foram notados nos alunos da 4ª série na disciplina de matemática, que tiveram aumento de 4,17 pontos na escala do Saeb (varia de 0 a 500), se controlado pelo desempenho dos diretores, e 3,30 pontos se controlado pelo desempenho dos professores. Alinhado com estes resultados, Oshiro (2011) apurou ganhos de 6,35 pontos na nota do aluno.

Todos os 4 resultados da 8ª série, indistintamente, seja em português ou matemática, com ou sem diretores ou professores, são negativos. Em português os alunos reduziram 4,42 e

4,61 pontos respectivamente, se controlado pelo diretores e professores. Em matemática houve redução de 5,57 e 6,72 pontos respectivamente.

A tentativa de isolar o desempenho ora pelo diretor ora pelo professor não produziu alterações significativas nos resultados. Por este desenho de pesquisa, ambos exercem influência semelhante no desempenho dos alunos.

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 da 8ª série foram menos negativos, mas continuaram significantes ao nível de confiança de 1%.

Diante destes resultados do aluno é possível inferir que, mesmo não se tendo demonstrado associação direta do desempenho do diretor na proficiência do aluno e mesmo que os professores de matemática da 4ª série não tenham demonstrado melhora em seu desempenho, os alunos da 4ª série desta mesma disciplina se beneficiaram da implantação do PRV em 2008. Para os alunos da 4ª série de português e os alunos da 8ª série não há evidências de que os incentivos aos diretores e professores tiveram algum impacto na sua proficiência.

Baseado nos três resultados apresentados, seja do diretor, professor ou aluno e retornando à hipótese de pesquisa, pode-se concluir que a remuneração variável tem efeitos positivos no desempenho do servidor público apenas em atividades específicas. No caso do servidores da educação do Estado de São Paulo constatou-se efeitos positivos no diretor na “escolha conjunta dos livros didáticos”. No caso do professor de português da 8ª série constatou-se melhoria na atividade de “conversar sobre textos de jornais e revistas” com os alunos e no professor de matemática da 8ª série houve melhora em desenvolver a atividade de “interpretar resultados numéricos obtidos para dar uma resposta adequada ao problema”.

Também pode-se concluir que o programa de remuneração variável proporcionou melhora na proficiência dos alunos da 4ª série 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 da 4ª série de matemática, nenhum dos seus coeficientes se apresentou positivo e significativo nos 11 itens utilizados para avaliar o seu desempenho.

Baseado na teoria da agência apresentada anteriormente neste trabalho, 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 programa como instrumento de motivação extrínseca não pode ser percebido de forma evidente.

## 11.4. Recomendações

No decorrer deste estudo surgiram questões que podem ser aprofundadas em futuras pesquisas. A principal delas refere-se a necessidade de pesquisa qualitativa para desvendar as opiniões dos participantes do programa de remuneração variável analisado a fim de identificar, pela ótica de quem se beneficia do programa, suas vantagens e desvantagens, bem como saber quais seriam os motivadores extrínsecos que impactariam no seu desempenho.

Também sugere-se desenvolver mais pesquisas em busca da identificação do grau de importância que cada atividade associada ao desempenho dos diretores e professores exerce efetivamente. Nesta pesquisa as atividades selecionadas foram tratadas como de igual importância no modelo. Porém, possivelmente, há algumas mais importantes do que outras. Se esta graduação de nível de importância for viável, pode-se concentrar o esforço em novas pesquisas nas atividades que geram mais impacto no desempenho do servidor.

Finalizando, sugere-se aos gestores de programas de remuneração variável revisar o modelo periodicamente, podendo ser cada dois anos, que é o prazo mínimo necessário para se captar a efetividade de um programa, assim como demonstrado por Fryer (2011), Glewwe (2010) e Muralidharan (2009). Tais revisões evitarão que o propósito do bônus seja distorcido pelo efeito *crowding-out*, descrito por Weibel *et al.* (2009), que é a apropriação do benefício como um direito permanente e não relacionado à melhoria do desempenho.

## Referências Bibliográficas

ANGRIST, Joshua D; PISCHKE, Jörn-Steffen. **Mostly Harmless Econometrics, An Empiricist's Companion**. Princeton University Press, 2009.

ASSIS, Luis Otávio M; NETO, Mario T. R. **Principais Características do Sistema de Remuneração Variável no Choque de Gestão em Minas Gerais: O acordo de resultados e o prêmio por produtividade**. *Gestão & Regionalidade*, v. 26, n. 76, p. 75-90, jan-abr, 2010.

**Remuneração Variável no Setor Público: Investigação das Causas do Fracasso e Implicações para o Estado Brasileiro**. *Revista Eletrônica de Gestão Organizacional*, Vol. 9, No. 3 p. 585 - 614, set. / dez. 2011.

**Efeitos Organizacionais da Implantação de Metas e Remuneração Variável por Desempenho: o caso da segurança pública em Minas Gerais**. Tese de mestrado. Fundação Getulio Vargas/EAESP, 2012.

BEHN, R. D. **The big questions of public management**. *Public Administration Review*, 55(4):313-24, 1995.

**Why public managers must measure outputs**. *Public Management Report*, vol.1, n. 10, June, 2004.

BÉNABOU, Roland; TIROLE, Jean. **Intrinsic and Extrinsic Motivation**. *The Review of Economic Studies*, Vol. 70, No. 3 (Jul., 2003), pp 489-520.

BESANKO, David et al. **A Economia da Estratégia**. 5ª ed. Porto Alegre: Bookman, 2012

BOYLE, Richard. **Managing public sector performance**. Dublin : Institute of Public Administration, 1989.

BOWMAN, James S. **The Success of Failure: The Paradox of Performance Pay**. *Review of Public Personnel Administration*. 30; 70, 2010,

BURGESS, Simon et al. **Evaluating the impact of performance-related pay for teachers in England**. *Labour Economics* 16 (2009) 251–261

CADSBY, C. Bram. **Sorting and incentive effects of pay for performance: an experimental investigation** –*Academy of Management Journal*. Vol. 50, No. 2, 387–405, 2007.



CAMERON, A. Colin; TRIVEDI, Pravin K. **Microeconometrics Using Stata**. Texas-USA : Stata Press, 2010.

COLEMAN, P., COLLINGE, J. & SCIFERT, T. (1993) **Seeking the Levers of Change: Participant attitudes and school improvement**. *School Effectiveness and School Improvement*, International Journal of Research, Policy and Practice. Volume 4, (1): 59-83.

COURTY, Pascal; MARSCHKE, Gerald. **Performance Incentives with Award Constraints**. *The Journal of Human Resources*, Vol. 37, No 4 (Autumn, 2002), pp 812-845.

CREWSON, Philip E. **Public-Service Motivation: Building Empirical Evidence of Incidence and Effect**. *Journal of Public Administration Research and Theory* (4-1997), p 499-518

CURY, A. **Organização e métodos: uma visão holística**. São Paulo : Atlas, 1994.

DIXIT, Avinash. **Incentives and Organizations in the Public Sector: And Interpretative Review**. *The Journal of Human Resources*. Vol. 37. No 4 (Autumm, 2002), p 696-727

DUTRA, Joel de Souza. **Competências: Conceitos e Instrumentos para a Gestão de Pessoas na Empresa Moderna**. São Paulo: Atlas, 2007.

EHRENBERG, Ronald G.; SMITH, Robert S. **A moderna economia do trabalho – teoria e política pública**. 5. ed. São Paulo, Makron Books, 2000.

FERNANDES, Bruno Henrique R., HIPÓLITO, José Antonio M. **Dimensões de Avaliação de Pessoa e o Conceito de Competências**. In: DUTRA, Joel Souza; FLEURY, Maria Tereza; RUAS, Roberto. *Competências: Conceitos, Métodos e Experiências*. São Paulo: Atlas, 2008.

FLANNERY, Thomas; HOFRICHTER, David; PLATTEN, Paul. **Pessoas. Desempenho e salários: as mudanças na forma de remuneração nas empresas**. São Paulo: Futura, 1997.

FOGUEL, Miguel Nathan. **Diferenças em Diferenças**. In: MENEZES FILHO, Naercio. *Avaliação Econômica de Projetos Sociais*. São Paulo: Dinâmica, 2012.

FRYER, Roland G. **Teacher incentives and student achievement: Evidence from New York City Public Schools**. National Bureau of Economic Research. Cambridge, MA. March 2011.

GIBBONS, Robert. **Incentives and Careers in Organizations**. In: KREPS, David M.; WALLIS, Kenneth F. *Advances in Economics and Econometrics: Theory and Applications*. Cambridge University Press: 1997.

\_\_\_\_\_. **Incentives in Organizations**. National Bureau of Economic Research. Working Paper 6695. August 1998.

GLEWWE, Paul; ILIAS, Nauman; KREMER, Michael. **Teacher Incentives**. American Economic Journal: Applied Economics 2. July 2010, 205-227.

JENKINS, G. Douglas *et al.* **Are Financial Incentives Related to Performance?** A meta-empirical Review of empirical research. Journal of Applied Psychology, v. 83, n. 5, p. 777-787. 1998.

HANUSHEK, Eric. **Teacher Characteristics and Gains in Student Achievement: Estimation Using Micro Data**. The American Economic Review, May, 1971, 280-288.

\_\_\_\_\_. **The Economics of Schooling: Production and Efficiency in Public Schools**. Journal Economic Literature, September 1986, 1141-1177.

\_\_\_\_\_. **The Failure of Input-Based Schooling Policies**. The Economic Journal, February 2003, F64-F98.

\_\_\_\_\_ & Woessmann, L. **The Role of Cognitive Skills in Economic Development**. Journal of Economic Literature, 46(3), 607-68, 2008.

\_\_\_\_\_ & RIVKIN, S. G. **Generalizations about Using Value-Added Measures of Teacher Quality**. American Economic Review, 100(2), 267-271, 2010.

HASNAIN, Zahid; MANNING, Nick; PIERSKALLA, Jan H. **Performance-related Pay in the Public Sector**. A Review of Theory and Evidence. World Bank, abril 2012.

HIPÓLITO, José Antonio M. **Administração Salarial: A Remuneração por Competências como Diferencial Competitivo**. São Paulo: Atlas, 2001.

HOOD, Christopher. **Gaming in Targetworld: The Targets Approach to Managing British Public Services**. Public Administration Review, july/august, 2006.

HOUSTON, David J. **Public-Service Motivation: A Multivariate Test**. Journal of Public Administration Research and Theory (4-2000), pp 713-727

KARAWEJCZYK, Tamára Cecília. Seminário de Gestão por Competência. Curso de Especialização em Gestão de Recursos Humanos. Passo Fundo: UPF, 2005.

LAVY, Victor. **Performance pay and teachers' effort, productivity and grading ethics.** National Bureau of Economic Research. Cambridge, MA. June 2004.

LAZEAR, Edward P. **Performance Pay and Productivity.** American Economic Review, v. 90, n. 5, p. 1346-1361. 2000,

LEWIS, Gregory B., FRANK, Sue A. **Who Wants to Work for the Government?** Public Administration Review. Vol. 62, july/august 2002.

LONGO, Francisco. **Mérito e Flexibilidade: A Gestão de Pessoas no Setor Público.** São Paulo: Edições Fundap, 2007.

MORICONI, Gabriela M. **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 Getulio Vargas/EAESP, 2012.

MARSDEN, David. **The paradox of performance related pay systems: 'why do we keep adopting them in the face of evidence that they fail to motivate?'** in Hood, C; MARGETTS, H; **Paradoxes of modernization: unintended consequences of public policy reforms.** Oxford, UK: Oxford University Press 2010 (In Press).

MOYNIHAN, Donald P. **Normative Model in Decline? Public Service Motivation in the Age of Governance.** In: PERRY, James L.; HONDEGHEM, Annie. Motivation in Public Management. Oxford, 2008.

MURALIDHARAN, Karthik; SUNDARARAMAN, Venkatesh. **Teacher performance pay: experimental evidence from India.** National Bureau of Economic Research. Cambridge, MA. September 2009.

OECD- Organisation for Economic Co-Operation and Development. **Performance-related pay: policies for government employees.** Paris: OECD Publishing, 2005a.

\_\_\_\_\_ **Performance-related pay: policies for government employees (Brief).** Paris: OECD Publishing, 2005b.

OSHIRO, Cláudia H. SCORZAFAVE, Luiz G. **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. 2011.

PACHECO, Regina Sílvia V. **Mensuração de desempenho no setor público: os termos de um debate.** In: II Congresso Consad de Gestão Pública. Brasília: Anais do Consad, 2009.

\_\_\_\_\_. **Brasil: avanços da contratualização de resultados no setor público.** Texto apresentado ao XI Congresso Internacional del CLAD, Ciudad de Guatemala, 2006.

PANDEY, Sanjay K.; STZYK, Edmund C. **Antecedents and Correlates of Public Service Motivation.** In: PERRY, James L.; HONDEGHEM, Annie. *Motivation in Public Management.* Oxford, 2008.

PERRY, James L.; WISE, Lois R. **The Motivational Bases of Public Service.** *Public Administration Review* (May/June 1990), pp 367-373

\_\_\_\_\_. **Measuring Public Service Motivation: An Assessment of Construct Reliability and Validity.** *Journal of Public Administration Research and Theory* (6-1996), p 5-22.

\_\_\_\_\_. ENGBERS, Trent A. & JUN, So Yun. **Back to the future?** Performance-related pay, empirical research, and the perils of persistence. *Public Administration Review*, v. 69, n.1, p. 1-31, Washington, January/February, 2009.

PLOTHOW, Célia Bueno A., HIPÓLITO, José Antonio M. **Remuneração por Competências: Alternativas e Implicações.** In: DUTRA, Joel Souza; FLEURY, Maria Tereza; RUAS, Roberto. *Competências: Conceitos, Métodos e Experiências.* São Paulo: Atlas, 2008.

POLLITT, Christopher. **Performance management in practice: a comparative study of executive agencies.** *Journal of Public Administration Research and Theory*, Oxford, June, 2005.

PRENDERGAST, Canice. **The Provision of Incentives in Firms.** *Journal of Economic Literature*, Vol. 37, No. 1 (Mar., 1999), pp. 7-63.

RIVKIN, Steven G. HANUSHEK, Eric. **Teachers, schools, and academic achievement.** *Journal of the Econometric Society*, Vol. 73, No. 2 (March, 2005), 417–458.

RYAN, Richard M.; DECI, Edward L. **Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions.** *Contemporary Educational Psychology* 25, 54-67. 2000.

SAMMONS, Pam. HILLMAN, Josh. MORTIMORE, Peter. **Key Characteristics of Effective Schools.** *School Effectiveness & School Improvement Journal*, march 1995

SOUZA, Maria Zélia A. *et al.* **Cargos, Carreiras e Remuneração.** Rio de Janeiro: FGV, 2005.

SCHUSTER, J. R & ZINGHEIM, P. K. **The new pay: linking employee and organizational performance.** New York: Lexington Books, 1995.

STUART, Elizabeth A. **Matching methods for causal inference: A review and a look forward.** National Institute of Health, February 1, 2010.

TROSA, Sylvie. *Gestão pública por resultados. Quando o Estado se compromete.* Rio de Janeiro: Revan; Brasília: ENAP, 2001.

VANDENABEELE, Wouter; WALLE, Steven Van de Walle. **International Differences in Public Service Motivation: Comparing Regions Across the World.** In: PERRY, James L.; HONDEGHEM, Annie. *Motivation in Public Management.* Oxford, 2008.

VIGDOR, Jacob L. **Teacher Salary Bonuses in North Carolina.** 2008.

VIELUF, Svenja *et al.* **Teaching Practices and Pedagogical Innovation.** OECD Publishing, 2012.

WEIBEL, Antoinette; ROST, Katja; OSTERLOH, Margit. **Pay for Performance in the Public Sector – Benefits and (Hidden) Costs.** *Journal of Public Administration Research and Theory*, v. 20, p. 387-412, Oxford, may, 2009.

WOOD JUNIOR, Thomaz; PICARELLI FILHO, Vicente. **Remuneração Estratégica: A Nova Vantagem Competitiva.** 3ª Ed. São Paulo: Atlas, 2004a.

\_\_\_\_\_. **Remuneração e Carreira por Habilidades e por Competências.** 3ª Ed. São Paulo: Atlas, 2004b.

XAVIER, Paulo Roberto; SILVA, Mateus Oliveira; NAKAHARA, Júlio Massaaki. **Remuneração Variável: Quando os Resultados Falam mais Alto.** São Paulo: Makron Books, 1999.

## Apêndice A

Proporção de Alunos com a Característica	2007		2009		2011	
	%	Nº Obs	%	Nº Obs	%	Nº Obs
Sexo Masculino	50,4	391.876	50,7	378.364	50,6	372.191
Branco ou pardo	83,5	387.186	84,8	371.715	86,7	354.333
Idade adequada para a série	91,0	394.078	91,8	379.417	92,5	376.548
Tem 1 ou mais carros em casa	55,7	390.646	58,6	377.350	63,0	375.132
Tem computador com internet em casa	33,8	389.992	45,9	379.597	62,7	375.862
Não tem banheiro em casa	0,6	391.175	1,0	375.341	0,7	376.028
Tem 1 banheiro em casa	64,8	391.175	66,6	375.341	66,1	376.028
Tem mais de 1 banheiro em casa	34,6	391.175	32,4	375.341	33,2	376.028
Tem pelo menos uma diarista em casa	10,8	378.145	8,9	380.367	8,8	375.785
Não tem quartos para dormir em casa	1,0	383.631	0,0	373.093	0,9	373.479
Tem 1 quarto para dormir em casa	17,3	383.631	17,3	373.093	15,6	373.479
Tem mais de 1 quarto para dormir em casa	81,7	383.631	82,7	373.093	83,5	373.479
Moram na mesma casa entre 3 e 6 pessoas, contando com o aluno	79,0	384.535	79,2	369.763	78,9	374.402
Mora com a Mãe	92,6	390.673	92,6	367.775	92,8	375.291
Mãe nunca estudou ou não completou a 4 série	8,3	380.798	7,1	365.840	9,4	367.887
Mãe completou a 4 série mas não completou a 8 série	23,0	380.798	21,3	365.840	16,1	367.887
Mãe completou a 8 série mas não completou o ensino médio	17,3	380.798	17,3	365.840	16,1	367.887
Mãe completou o ensino médio mas não completou a faculdade	18,2	380.798	19,6	365.840	21,9	367.887
Mãe completou a faculdade	9,1	380.798	8,9	365.840	9,4	367.887
Não sabe informar a escolaridade da mãe	24,1	380.798	25,8	365.840	27,1	367.887
Mãe ou mulher responsável sabe ler e escrever	96,2	384.136	96,8	370.600	96,4	370.749
Mãe ou mulher responsável é vista lendo	89,6	378.454	89,5	366.089	90,3	368.531
Mora com o Pai	68,6	386.403	67,3	370.067	68,3	370.710

Proporção de Alunos com a Característica	2007		2009		2011	
	%	Nº Obs	%	Nº Obs	%	Nº Obs
Pai nunca estudou ou não completou a 4 série	7,1	384.532	6,8	370.932	9,5	339.554
Pai completou a 4 série mas não completou a 8 série	17,1	384.532	16,8	370.932	12,3	339.554
Pai completou a 8 série mas não completou o ensino médio	15,3	384.532	15,4	370.932	15,1	339.554
Pai completou o ensino médio mas não completou a faculdade	16,0	384.532	16,4	370.932	18,7	339.554
Pai completou a faculdade	10,2	384.532	9,3	370.932	9,9	339.554
Não sabe informar a escolaridade do Pai	34,3	384.532	35,3	370.932	34,5	339.554
Pai ou homem responsável sabe ler e escrever	93,9	385.437	94,4	351.126	92,5	339.430
Pai ou homem responsável é visto lendo	79,2	385.631	78,2	369.862	80,8	337.239
Presença frequente dos responsáveis em reunião dos pais	69,9	387.196	69,3	374.074	67,8	372.604
Responsáveis incentivam os estudos	98,0	387.925	98,2	367.495	98,5	357.243
Responsáveis incentivam fazer deveres de casa e trabalhos	96,0	390.322	96,2	375.031	96,2	371.203
Responsáveis incentivam a ler	93,3	386.351	93,9	374.147	93,5	355.326
Responsáveis incentivam a presença na escola	97,2	386.435	97,5	370.533	97,6	371.993
Responsáveis conversam sobre acontecimentos da escola	97,2	386.435	97,5	370.533	97,6	371.993
Aluno não faz trabalho doméstico	20,0	375.492	19,3	367.173	17,9	370.623
Aluno gasta até 1 hora por dia de aula em trabalhos domésticos	39,8	375.492	44,4	367.173	46,1	370.623
Aluno gasta mais de 1 hora por dia de aula em trabalhos domésticos	40,2	375.492	36,3	367.173	36,0	370.623
Não trabalha fora de casa	85,7	384.572	86,8	359.401	87,3	366.759
Começou a estudar no Maternal	40,1	382.403	40,0	366.321	37,4	369.813
Começou a estudar na pré-escola	43,7	382.403	44,0	366.321	47,0	369.813
Começou a estudar na primeira série	13,4	382.403	12,2	366.321	12,8	369.813
Começou a estudar depois da primeira série	2,8	382.403	3,8	366.321	2,8	369.813

Proporção de Alunos com a Característica	2007		2009		2011	
	%	Nº Obs	%	Nº Obs	%	Nº Obs
Se 4 série - Sempre estudou na mesma escola	65,3	202.969	66,2	204.618	82,0	184.384
Se 4 série - Mudou de escola mas sempre em escolas públicas	26,8	202.969	26,0	204.618	5,0	184.384
Se 4 série - Mudou de escola e já estudou em escolas particulares	7,9	202.969	7,8	204.618	13,0	184.384
Se 8 série - Sempre estudou em escolas públicas	92,9	183.010	93,1	165.432	92,7	187.083
Se 8 série - Sempre estudou em escolas particulares	1,2	183.010	1,3	165.432	1,4	187.083
Se 8 série - Estudou em escolas públicas e particulares	5,9	183.010	5,6	165.432	6,0	187.083
Nunca foi reprovado	82,5	386.385	81,4	368.525	82,2	371.397
Apenas se 8 série - Após a 8 série pretende continuar estudando	21,0	183.306	32,5	150.777	25,8	187.500
Apenas se 8 série - Após a 8 série pretende continuar estudando e trabalhar	64,5	183.306	65,1	150.777	62,1	187.500
Apenas se 8 série - Após a 8 série pretende apenas trabalhar ou não sabe o que vai fazer	2,7	183.306	2,4	150.777	2,5	187.500