UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE DE
RIBEIRÃO PRETO
DEPARTAMENTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA – ÁREA: ECONOMIA
APLICADA

RENATA DE LACERDA ANTUNES BORGES

Produtividade científica e infraestrutura tecnológica

ORIENTADOR: PROF. DR. SÉRGIO KANNEBLEY JÚNIOR

RIBEIRÃO PRETO

2015

Prof. Dr. Marco Antônio Zago
Reitor da Universidade de São Paulo

Prof. Dr. Dante Pinheiro Martinelli
Diretor da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto

Prof. Dr. Renato Leite Marcondes
Chefe do Departamento de Economia

RENATA DE LACERDA ANTUNES BORGES

Produtividade científica e infraestrutura tecnológica

Dissertação apresentada ao Programa de Pós-Graduação em Economia – Área: Economia Aplicada da Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto da Universidade de São Paulo, para obtenção do título de Mestre em Ciências.

ORIENTADOR: PROF. DR. SÉRGIO KANNEBLEY JÚNIOR

RIBEIRÃO PRETO

2015

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

FICHA CATALOGRÁFICA

Borges, Renata de Lacerda Antunes

Produtividade científica e infraestrutura tecnológica. Ribeirão Preto, 2015.

116 p. : il. ; 30 cm

Dissertação de Mestrado, apresentada à Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto/USP.

Orientador: Kannebley Júnior, Sérgio.

1. Produtividade Científica. 2. Infraestrutura de pesquisa. 3. Sistemas de Inovação.

DEDICATÓRIA

Dedico esta dissertação à minha família, ao Prof. Dr. Maurício Jorge de Souza Pinto (in memoriam) e a todos os meus mestres educadores que me orientaram durante a vida.

AGRADECIMENTOS

Agradeço, primeiramente, a Deus pela minha saúde e dos meus familiares, além de força para superar as dificuldades. Aos meus pais, Maria, Nilo e Nelson, e aos meus irmãos, Diogo e Thiago pelo amor, incentivo e apoio incondicional.

Também gostaria de agradecer a Universidade de São Paulo, pela oportunidade de realização do mestrado, em especial, aos professores Sérgio Sakurai, Elaine Pazzelo e Cláudio Lucinda por terem confiado no meu potencial ao selecionar-me ao curso. Aproveito para agradecer a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro no primeiro ano de mestrado.

Meu eterno obrigada ao prof. Dr. Sérgio Kannebley Jr. pela disposição ao aceitar me orientar, apoio em todos os momentos, confiança e dedicação à elaboração desta dissertação. Ademais, ao Instituto de Pesquisa Econômica Aplicada (IPEA), mais diretamente a Fernanda De Negri, pela disponibilidade da base de dados e pelos esclarecimentos fornecidos sempre quando preciso.

Aos professores participantes do grupo CEPER – Rudinei Toneto Jr, Sérgio Sakurai e Luciano Nakabashi – por todo o aprendizado durante esses quase dois anos de trabalho. Meu muito obrigada aos funcionários da FEA-RP/USP pela presteza e esclarecimentos durante este processo. Em especial, ao André Pignata, que foi o responsável pela extração dos dados que fizeram parte deste estudo, sem ele tudo seria mais difícil.

Aos meus colegas de mestrado, gostaria de agradecer todo o aprendizado compartilhado. A todos os meus amigos por entender minha ausência em momentos importantes e pela amizade dedicada sempre.

Ao meu companheiro, Caio Lopes, por todo o carinho e dedicação, além dos vários momentos de ensinamentos acadêmicos e ajudas essenciais na base de dados. Tudo compartilhado de forma paciente, sabendo ouvir, me estimular e acreditar sempre no meu potencial.

RESUMO

BORGES, R. L. A. **Produtividade científica e infraestrutura tecnológica**. 2015. 116 f. Dissertação (Mestrado) - Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, São Paulo, 2015.

O Sistema de Inovação (SI) brasileiro é caracterizado por um estágio intermediário de desenvolvimento e pela forte participação do Setor Público, refletindo assim em seus grandes dispêndios em C&T. O SI é considerado, no cenário atual, como determinante ao desenvolvimento e crescimento econômico das economias modernas. Porém, há poucos estudos sobre o papel desempenhado pelas características dos pesquisadores e das infraestruturas de P&D sobre a produção científica e tecnológica. Portanto, o objetivo dessa pesquisa é avaliar os efeitos de tais características sobre a produtividade científica individual dos coordenadores pertencentes às 1.756 infraestruturas de pesquisa no Brasil. Para tal, o trabalho utiliza um mapeamento inédito das infraestruturas nacionais realizado pelo Ministério da Ciência, Tecnologia e Inovação em parceria com o CNPq e o IPEA, e a extração das características individuais dos pesquisadores por meio da Plataforma *Lattes*. As infraestruturas de pesquisa foram agrupadas por meio da Análise de Cluster, o qual gerou três clusters que se distinguem pelo escopo de suas atividades, áreas científicas e composição do capital físico e humano. Os pesquisadores com maiores produtividades são àqueles presentes no cluster 2 (multiatividades, Engenharia e Ciências Biológicas, estrutura organizacional piramidal), em detrimento aos presentes no cluster 1 (pesquisa e desenvolvimento de tecnologia, Multiárea) ou 3 (intensivos em pesquisa, Engenharias e Ciências Exatas e da Terra, estrutura organizacional horizontal). Os resultados dos modelos de regressão, OLS e Tobit, estimados indicam que quanto maior o número de estudantes bolsistas, maior a produtividade do coordenador, bem como a relação capital-trabalho da infraestrutura. Por fim, o presente estudo contribui ao preencher uma lacuna de informação sobre o sistema brasileiro de C&T e após exercícios econométricos foi possível concluir que as dimensões do SI nacional, relacionando capital humano, independente da qualificação, e físico, são determinantes aos níveis de produtividade dos coordenadores.

Palavras-chave: produtividade científica, infraestrutura de pesquisa, sistemas de inovação.

ABSTRACT

BORGES, R. L. A. **Scientific productivity and technological infrastructure**. 2015. 116 f. Dissertation (Master's Degree) – Faculdade de Economia Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo, São Paulo, 2015.

The Brazilian innovation system (IS) is characterized by an intermediate stage of development and the strong participation of the public sector, reflecting its large share in expenditure on S & T. The IS is considered, in the current scenario, a key factor to development and economic growth of modern economies. However, there are few studies on the role played by the characteristics of the researchers and research infrastructure on scientific productivity. Therefore, the aim of this research is to evaluate the effects of these characteristics on individual scientific productivity of the coordinators belonging to 1,756 research infrastructures in Brazil. To do this, the work uses a unique mapping of infrastructure in Brazil by the Ministry of ST & I in partnership with CNPq and IPEA, and the extraction of the individual characteristics of the researchers through the *Lattes* Platform. The research infrastructure were characterized by the Cluster Analysis, which generated three clusters distinguished by the scope of its activities, scientific areas and composition of physical and human capital. Researchers with higher productivity were present in cluster 2 (multi-activities, Engineering and Life Sciences, pyramidal organizational structure) rather than those present in the cluster 1 (research and development of technology, multi-area) or 3 (research intensive, Engineering and Exact and Earth Sciences, horizontal organizational structure). OLS and Tobit regression models results indicate that the larger the number of scholarship students and capital-labor ratio of the infrastructure the better the productivity of the coordinator. Finally, this study contributes to fill an information gap on the Brazilian system of S & T and after econometric exercise was concluded that the dimensions of the IS, relating human capital, regardless of qualification, and physical, are crucial to the productivity levels of the coordinators.

Keywords: scientific productivity, research infrastructure, innovation systems.

LISTA DE GRÁFICOS

Gráfico 1 – Histograma da quantidade de artigos publicados entre 2009 a 2013 pelo total da mão de obra pertencentes às infraestruturas.	43
Gráfico 2 – Distribuição da quantidade de artigos publicados de 2009 a 2013 pelas infraestruturas, segundo as regiões do Brasil de atuação da mão de obra das infraestruturas.	44
Gráfico 4 – Ligação Simples para dissimilaridade entre as 15 variáveis utilizadas na construção do <i>cluster</i>	96

LISTA DE QUADROS

Quadro 1 – Conjunto de variáveis de determinação dos <i>clusters</i>	48
Quadro 2: Correspondência entre o <i>cluster</i> global e cada subgrupo.	52
Quadro 3: Resultado do teste MANOVA (<i>Multivariate analysis-of-variance</i>) aplicado sobre o <i>cluster</i> fk3 gerado.	53
Quadro 4 – Teste MVREG.	94

LISTA DE TABELAS

Tabela 1 – Infraestruturas por tipo de configuração.....	28
Tabela 2 – Distribuição regional das infraestruturas.....	28
Tabela 3 – Ano de início de operação das infraestruturas.....	28
Tabela 4 – Distribuição das infraestruturas por grandes áreas do conhecimento das infraestruturas.....	29
Tabela 5 – Mão de obra, por titulação.....	29
Tabela 6 – Mão de obra, por tipo de vínculo.....	30
Tabela 7 – Mão de obra das infraestruturas, por titulação e constituição/localização em 2012.	31
Tabela 8 – Quantidade de infraestruturas, de acordo com a interação das atividades desenvolvidas na categoria “ <i>contínuo</i> ”.....	32
Tabela 9 – Atividades realizadas, segundo a grande área de atuação das infraestruturas.....	33
Tabela 10 – Distribuição das infraestruturas, de acordo com a realização ou não de cooperação com outras infraestruturas e/ou empresas do Brasil ou do exterior.....	33
Tabela 11 – Distribuição (%) das infraestruturas que realização cooperação segundo área do conhecimento de atuação.....	34
Tabela 12 – Distribuição das infraestruturas pesquisadas segundo faixa de valor estimado das infraestruturas e dos equipamentos de pesquisa (em R\$ de 2012).....	35
Tabela 13 – Estimativa do valor das infraestruturas, segundo a área de conhecimento em que atuam (em R\$ de 2012).	36
Tabela 14 – Estimativa dos custos operacionais e receitas anuais das infraestruturas (em R\$ de 2012).....	37
Tabela 15 – Financiamento recebido, por entidade financiadora (em R\$ médio de 2012).....	37
Tabela 16 – Distribuição (%) dos coordenadores, os quais avaliaram as condições físicas das infraestruturas como “ <i>muito boa</i> ” e “ <i>boa</i> ”, por a área do conhecimento.....	38
Tabela 17 – Distribuição (%) dos coordenadores, os quais avaliaram os recursos humanos das infraestruturas como “ <i>adequado</i> ”, por a área do conhecimento.	39
Tabela 18 – Distribuição da mão de obra e de artigos publicados no quinquênio (2009-2013), segundo as características de ocupação nas infraestruturas.....	41
Tabela 19 – Coordenadores, por titulação máxima.....	41
Tabela 20 – Coordenadores, por tipo de vínculo.....	42
Tabela 21 – Coordenadores, por ano de início de operação das infraestruturas.	42

Tabela 22 – Quantidade de artigos publicados de 2009 a 2013 segundo a grande área do conhecimento de atuação.....	43
Tabela 23 – Quantidade de artigos publicados no período de 2009-2013 de acordo com a Faixa de Idade de obtenção do título de doutorado dos pesquisadores.	44
Tabela 24 – Descrição média das variáveis de Capital Humano presente por infraestrutura. .	48
Tabela 25 – Descrição da variável grande área do conhecimento usada na criação dos <i>clusters</i>	49
Tabela 26 – Descrição da distribuição das variáveis de atividades, por intensidade de realização.	50
Tabela 27 – Divisão das infraestruturas de acordo com o número de grupos e processos de inicialização de escolha de K.	51
Tabela 28 – Centroides das variáveis, por grupos do <i>cluster</i> fk3.....	53
Tabela 29 – Discriminação das características de ocupação dos pesquisadores, de acordo com as tipologias de <i>cluster não-hierárquico</i>	56
Tabela 30 – Estatísticas descritivas das variáveis da produtividade científica dos coordenadores e pesquisadores permanentes.	65
Tabela 31 – Estatísticas descritivas das variáveis da produtividade científica dos coordenadores, por <i>clusters</i>	68
Tabela 32 – Descritiva das variáveis de mão de obra na infraestrutura e por <i>cluster</i>	69
Tabela 33 – Descritiva da variável relação capital-trabalho nas infraestruturas e por <i>cluster</i> . .	70
Tabela 34 – Resultados das Regressões – Efeitos Marginais da Produtividade Científica dos coordenadores contra as características coletivas e individuais.	74
Tabela 35 – Resultados das Regressões – Efeitos Marginais da Produtividade Científica dos coordenadores contra as características coletivas por <i>Cluster</i>	75
Tabela 36 – Resultados das Regressões – Efeitos Marginais da Produtividade Científica dos coordenadores contra as características coletivas e individuais por <i>Cluster</i>	76
Tabela 37 – Resultados das Regressões – Efeitos Marginais da Produtividade Científica dos coordenadores contra as características coletivas e individuais ponderados pelo IFPA.	80
Tabela 38 – Resultados das Regressões – Efeitos Marginais da Produtividade Científica dos coordenadores contra as características coletivas por <i>Cluster</i> ponderados pelo IFPA.....	81
Tabela 39 – Resultados das Regressões – Efeitos Marginais da Produtividade Científica dos coordenadores contra as características coletivas e individuais por <i>Cluster</i> ponderados pelo IFPA.	82

Tabela 40 – Matriz de similaridade usando como medida de proximidade a correlação entre os pares de variáveis utilizadas na criação do <i>cluster</i>	98
Tabela 41 – Resultados das Regressões – Efeitos Marginais da Produtividade Científica Individual contra as características coletivas e individuais (variáveis absolutas).	99
Tabela 42 – Resultados das Regressões – Efeitos Marginais da Produtividade Científica Individual contra as características coletivas por <i>Cluster</i> (variáveis absolutas).	100
Tabela 43 – Resultados das Regressões – Efeitos Marginais da Produtividade Científica Individual contra as características coletivas e individuais por <i>Cluster</i> (variáveis absolutas).	101
Tabela 44 – Resultados das Regressões – Efeitos Marginais da Produtividade Científica Individual contra as características coletivas e individuais ponderados pelo IFPA (variáveis absolutas).	102
Tabela 45 – Resultados das Regressões – Efeitos Marginais da Produtividade Científica Individual contra as características coletivas por <i>Cluster</i> , ponderados pelo IFPA (variáveis absolutas).	103
Tabela 46 – Resultados das Regressões – Efeitos Marginais da Produtividade Científica Individual contra as características coletivas e individuais por <i>Cluster</i> , ponderados pelo IFPA (variáveis absolutas).	104

SUMÁRIO

1	INTRODUÇÃO	11
2	DETERMINANTES DA PRODUTIVIDADE CIENTÍFICA	15
2.1.	Determinantes Individuais da Produtividade Científica	15
2.2.	Determinantes Coletivos da Produtividade Científica	21
3	BASE DE DADOS	26
3.1.	Mapeamento das Infraestruturas	26
3.2.	Mapeamento dos Pesquisadores	40
4	ANÁLISE DE CLUSTER: DEFINIÇÃO DE CATEGORIAS DAS INFRAESTRUTURAS	
	46	
4.1.	Construção da base de dados	47
4.2.	Cluster Não-Hierárquico.....	50
4.3.	Resultados por tipologia das Infraestruturas.....	53
5	METODOLOGIA ECONOMÉTRICA.....	57
5.1.	Modelo com Dados em <i>Cross-Section</i> para Determinantes da Produtividade Individual ...	57
6	RESULTADOS	65
6.1.	Estatísticas Descritivas	65
6.2.	Modelo de Produtividade Científica Individual.....	70
6.3.	Modelo de Produtividade Científica Individual ponderada pelo IFPA.....	77
7	CONCLUSÃO	83
	REFERÊNCIAS	86
	APÊNDICE A – PROCESSOS DE VALIDAÇÃO DOS CLUSTERS	90
	APÊNDICE B – TESTE MVREG.	93
	APÊNDICE C – CLUSTER HIERÁRQUICO.....	95
	APÊNDICE D - MODELO DE PRODUTIVIDADE CIENTÍFICA INDIVIDUAL COM	
	VARIÁVEIS ABSOLUTAS	99
	APÊNDICE E – MODELO DE PRODUTIVIDADE CIENTÍFICA INDIVIDUAL	
	PONDERADA PELO IFPA, COM VARIÁVEIS ABSOLUTAS.....	102
	ANEXO A – QUESTIONÁRIO	105

1 INTRODUÇÃO

O papel da Ciência, Tecnologia e Inovação (C, T&I) é reconhecidamente um agente diferenciador ao desenvolvimento social e crescimento econômico das economias modernas. Porém, nem todos os países encontram-se no mesmo estágio de desenvolvimento do Sistema de Inovação (SI). O Brasil, assim como a maioria dos países em desenvolvimento, possui um sistema de C, T&I em estágio intermediário de consolidação, em que a participação das universidades e de instituições públicas de pesquisa é bastante expressiva.

Dados do Ministério de Ciência, Tecnologia e Inovação (MCTI) ¹, de 2012, revelam que 52,4% da participação relativa dos dispêndios em C&T foram realizados por entes públicos e 47,6% por empresas privadas e/ou estatais. Uma das possíveis consequências dessa forma de constituição foi o avanço em posições nos *rankings* internacionais de produção científica e tecnológica, o que demonstra o potencial produtivo da infraestrutura de pesquisa nacional. Em 16 anos, o Brasil obteve um crescimento de 616,6% de artigos indexados pela *Scopus*, enquanto que a produção mundial de artigos indexada ao mesmo órgão cresceu 200%. Em resposta ao aumento da produção científica, houve acréscimo de 300% no número de pedidos de patentes do Brasil junto ao Escritório Americano de Marcas e Patentes (USPTO) no período de 2000 a 2012.

Na literatura tanto determinantes individuais quanto os coletivos buscam explicar os diferenciais de produção acadêmica por meio, de forma geral, do número de artigos publicados como *proxy* de produtividade científica. Dentre os fatores individuais citam-se: gênero (COLE; COLE, 1973; XIE; SHAUMAN, 1998 e 2003; TURNER; MAIRESSE, 2005; FOX, 2005); idade (COLE, 1979; WEISS; LILLARD, 1982; FOX, 1983; DIAMOND, 1986; LEVIN; STEPHAN, 1991; CARAYOL; MATT, 2006; GONZALEZ-BRAMBILA; VELOSO, 2007; KANNEBLEY JR.; CAROLO; NEGRI, 2014); e fatores não observáveis dos indivíduos (STEPHAN, 2012). Com relação às características coletivas, a produtividade acadêmica pode ser bastante influenciada devido às externalidades geradas quando os pesquisadores trabalham em times dentro dos laboratórios. Dessa forma, aspectos ligados ao capital humano (como estrutura de pessoal e reputação do grupo), infraestrutura física e complementariedade entre diferentes tipos de pesquisadores e entre diferentes agendas de pesquisa podem afetar a produtividade científica e tecnológica no âmbito coletivo

¹ <http://www.mcti.gov.br/index.php>.

(BONACCORSI; DARAIO, 2003; CARAYOL; MATT, 2004 e 2006; HORTA; LACY, 2011; STHEPAN, 2012).

Estudos empíricos realizados por Turner e Mairesse (2005) e Carayol e Matt (2006) avaliaram a produtividade científica de pesquisadores franceses. Ambos encontraram que o tamanho da infraestrutura, em termos de quantidade de pesquisadores, tem impacto negativo na produtividade individual. Além disso, o primeiro estudo mostrou que há fortes efeitos positivos da produção do laboratório e da participação em redes de colaboração internacional, enquanto que o segundo trabalho encontrou impactos positivos e significativos da presença de pós-doutorandos estrangeiros; das atividades de pesquisa dos colegas; e do financiamento público sobre a produção acadêmica do pesquisador. Quanto aos efeitos do Ciclo de Vida do cientista (idade), o pico de produtividade dos franceses ocorre próximo aos 52 anos.

No caso brasileiro há pouca investigação sobre o papel desempenhado pelos pesquisadores e das características das infraestruturas sobre a produção acadêmica individual. Deste modo, esse trabalho tem como objetivo de pesquisa avaliar os efeitos das características coletivas sobre a produtividade científica individual dos coordenadores pertencentes às infraestruturas de pesquisa no Brasil. Para cumprir com tal proposta, o presente trabalho irá utilizar duas bases de dados. A primeira trata-se de um Mapeamento inédito das infraestruturas de pesquisas² do Brasil realizado pelo Ministério da Ciência, Tecnologia e Inovação (MCTI) em parceria com o Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e o Instituto de Pesquisa Econômica Aplicada (IPEA). Tal base conta com 1.760 infraestruturas, sendo a maioria de tipologia de laboratório com idade média de 15 anos e situando-se na região Sudeste do país. A segunda base de dados é composta por informações individuais dos pesquisadores pertencentes às infraestruturas do mapeamento do MCTI extraídos da Plataforma *Lattes*-CNPq exclusivamente para a realização deste estudo. O público-alvo são coordenadores, os quais obtiveram o título de doutorado até o ano de 2009, totalizando 1.445 coordenadores distintos. Como *proxy* da produtividade científica individual, utilizou-se a quantidade de artigos publicados no período de 2009 a 2013 pelos coordenadores dividida pelo número de coautores.

A fim de analisar qual perfil de infraestrutura de pesquisa que propicia uma maior produção científica no Brasil, foram definidas, por meio da análise de cluster, três grupos de laboratórios. Os mesmos se distinguiram pelo escopo de suas atividades, áreas científicas e

² Conceitua-se “infraestrutura de pesquisa”, de acordo com o empregado no projeto *Meril* (*Mapping of the European Research Infrastructure Landscape*), como o conjunto de instalações físicas e condições materiais de apoio (equipamentos, recursos e serviços) utilizados pelos pesquisadores.

composição do capital físico e humano. O primeiro grupo, denominado “intensivos em pesquisa e desenvolvimento de tecnologias”, atuam em múltiplas áreas de conhecimento científico e suas atividades de pesquisa e desenvolvimento de tecnologias são realizadas de forma contínua. O segundo grupo, denominado como “laboratórios de multiatividades” atuam, principalmente, nas áreas científica de Engenharia e Ciências Biológicas, desenvolvem continuamente atividades pesquisa, ensino e desenvolvimento de tecnologia e possuem uma estrutura organizacional piramidal. O terceiro grupo, denominado “intensivos em pesquisa”, são laboratórios que atuam, principalmente, nas áreas científicas de Engenharia e Ciências Exatas e da Terra, cuja atividade de pesquisa é realizada de forma contínua, enquanto que as atividades de ensino ocorrem de modo descontínuo, além disso não há desenvolvimento de tecnologia e revelam uma estrutura organizacional horizontal.

Quanto à produtividade, pode-se dizer que as características dos laboratórios têm diferentes influências sobre a produção científica dos coordenadores dependendo da categoria de laboratório analisada. Resumidamente, encontrou-se que os pesquisadores que se dedicam à multiatividades (ensino, pesquisa e desenvolvimento de tecnologias) são mais produtivos que àqueles que pertencem aos *clusters* 1 (pesquisa e desenvolvimento de tecnologia) e 3 (intensivos em pesquisa). Nos modelos de regressão estimados, encontrou-se que a produtividade científica individual dos coordenadores é determinada, principalmente, pela sua própria produção no quinquênio anterior, além da produção dos pesquisadores permanentes no período de análise e da participação dos pesquisadores não-permanentes. Além desses fatores, quando avaliadas a produção científica dos coordenadores indexada internacionalmente, percebe-se a importância da relação capital-trabalho.

Enfim, o presente estudo contribui ao preencher uma lacuna de informação sobre o sistema brasileiro de C, T&I destacando-se, pelo escopo, dimensão e agregação de variáveis diferenciadas dos estudos internacionais. Assim, após exercícios econométricos foi possível concluir que as dimensões do SI, relacionando capital humano, independente da qualificação, e capital físico, são determinantes aos níveis de produtividade dos pesquisadores.

A estrutura do trabalho é composta por sete seções, incluindo essa Introdução. A segunda seção é dedicada a Revisão da Literatura, a qual expõe os principais determinantes, tanto individuais quanto coletivos, que influenciam a produção científica e tecnológica dos pesquisadores dos centros de pesquisa. Na terceira e quarta seções são apresentadas as bases de dados e a análise descritiva, conveniente ao escopo deste trabalho e a descrição do método de Análise de *Cluster*, respectivamente. Na quinta seção são apresentados os procedimentos

metodológicos. Na seção seguinte são demonstrados os principais resultados das estimações. Por fim, são apresentadas as conclusões gerais.

2 DETERMINANTES DA PRODUTIVIDADE CIENTÍFICA

Lotka (1926), em seus estudos empíricos, observou que a metade dos artigos são produzidos por um número bastante restrito de pesquisadores, ou seja, grande concentração da produção científica nas mãos de poucos pesquisadores/centros de pesquisa. Essa observação induziu o surgimento de vários estudos em áreas como sociologia e economia, a fim de racionalizar os fatores causadores de tal desigualdade no desempenho produtivo dos cientistas.

Diversos são os argumentos relacionados aos determinantes individuais e coletivos que tentam explicar os níveis desiguais de produtividade acadêmica, em sua maioria estão relacionados às heterogeneidades individuais dos pesquisadores, gênero, idade, infraestrutura de pesquisa, ambiente institucional, entre outros. Desta forma, o presente capítulo irá expor o arcabouço teórico no qual esses argumentos são discutidos.

2.1. Determinantes Individuais da Produtividade Científica

Inicialmente, a busca por racionalizações para os diferentes níveis de produtividade dos cientistas foi direcionada para a análise das características específicas dos pesquisadores. Para Stephan et al. (2004), dois fatores chaves levam os cientistas a se engajarem em pesquisa, são eles: o interesse em resolver um enigma da ciência e/ou o reconhecimento tanto pela academia, quanto financeiro pela descoberta em si.

Ademais, a literatura, em seus estudos, aponta como determinantes individuais da produtividade acadêmica fatores como: habilidade inata, idade, gênero, escolha de atuação na carreira profissional, entre outros. Nesta seção, serão abordados fatores, tais como as características não observáveis dos indivíduos; idade e geração; gênero; e Efeito de Mateus.

2.1.a Autosseleção e fatores não observáveis dos indivíduos

A heterogeneidade individual do pesquisador é argumentada como sendo um determinante não observável dos níveis desiguais de produtividade. Segundo a hipótese da Centelha Divina (*Sacred Spark Hypothesis*) existiria uma habilidade inata do cientista e/ou uma motivação na realização das atividades científicas que provoca tal desigualdade na produtividade. As pesquisas na área psicológica indicam que há um conjunto de fatores bibliográficos do pesquisador (anterior à vida acadêmica), como autonomia e independência, que podem resultar em diferenciais apresentados por cientistas mais produtivos e criativos³.

Stephan (2012) destaca, como outro fator importante, a persistência que o pesquisador deve possuir a fim de concluir longos períodos de trabalho. Para a autora, o estímulo pela criatividade e o esforço pela prática, através das muitas tentativas de se obter um resultado, levam o cientista a desenvolver um modo criativo de solucionar problemas. De fato, de acordo com evidências produzidas por Roach e Sauermann (2010), a partir de uma amostra de 400 doutorandos nas áreas de ciências e engenharia, existe uma diferença sistemática nas preferências desses estudantes com relação ao chamado “gosto pela ciência” (*Taste for Science*). Isto é, doutorandos com características individuais como desejo pela independência, reconhecimento pela publicação, interesse na pesquisa básica e habilidade em publicar preferiram, fortemente, a carreira acadêmica.

A crítica a essa linha de pesquisa está no questionamento de quão forte seria a correlação entre habilidade científica e produtividade a ponto de explicar tamanho diferencial de produção. Ou ainda, em razão da desconsideração de fatores sociais e organizacionais intervenientes que levariam os cientistas a traduzirem seus diferenciais de criatividade em produção científica efetiva. Segundo Fox (1983), as evidências produzidas por diversos autores demonstram que os diferenciais de talento e habilidade podem ser úteis para explicar a propensão a pesquisar, mas aparentemente não são capazes de explicar os diferentes níveis de produtividade dos pesquisadores⁴.

2.1.b Idade e Geração

³ Ainda são destacadas diferentes habilidades cognitivas dos cientistas altamente produtivos, principalmente, no aspecto qualitativo dessas diferenças.

⁴ Com relação aos aspectos de formação e/ou de trabalho do cientista são obtidas evidências sobre a importância e o prestígio do centro em que o cientista obteve seu doutorado, bem como a reputação da instituição em que atua, para afetar a sua produtividade científica.

Economistas e sociólogos têm interesse no estudo do impacto da idade sobre a produtividade científica dos pesquisadores. Fox (1983) apresenta evidências fornecidas por diversos autores quanto ao ciclo de vida dos cientistas e sua curva de produtividade ao longo do tempo. Essas evidências dividem os autores em dois grandes grupos.

O primeiro grupo (LEHMAN, 1953, apud CARAYOL; MATT, 2006; COLE, 1979) defende que as maiores, ou mais importantes, contribuições dos cientistas ocorreriam entre 30 e 40 anos de idade com posterior declínio. O estudo mais sistemático de Cole (1979) utilizou uma base de dados em *cross-section*, que englobava cientistas de seis campos do conhecimento (química, geologia, matemática, física, psicologia e sociologia), e como *proxy* da produtividade, o número de publicações entre 1965 e 1969. No trabalho seminal de Lehman (LEHMAN, 1953, apud CARAYOL; MATT, 2006), encontrou-se que os cientistas de disciplinas mais “abstratas” (como matemática e teoria física) possuem picos de produtividade mais cedo quando comparados aos de disciplinas mais aplicadas (como geologia e biologia). Para o segundo grupo existem dois picos na curva de produtividade. Pelz e Andrews (1976) encontraram que pico inicial ocorre entre 30 e 40 anos de idade, enquanto que o segundo pico ocorre de 10 a 15 anos após os 50 anos. Para Bayer e Dutton (1977), o primeiro pico ocorre 10 anos após o início da carreira dos cientistas e o último muito próximo da aposentadoria do pesquisador.

A racionalização por parte dos economistas no estudo da relação entre idade e produtividade científica e tecnológica apoia-se na Teoria do Capital Humano⁵. Desta forma, os modelos procuraram levar em conta a finitude da vida produtiva do pesquisador e investigar as implicações que isso teria sobre a alocação de tempo para a pesquisa ao longo da sua carreira acadêmica.

Os modelos de ciclos de vida propostos por Weiss e Lillard (1982), Diamond (1986) e Levin e Stephan (1991) são formulados assumindo que os cientistas se engajam em pesquisa em razão dos retornos financeiros futuros da atividade (DIAMOND, 1986) adicionados, ou não, pelo interesse associado à satisfação obtida com a resolução dos problemas científicos (STEPHAN, 1991). Resultando em uma produtividade crescente com retornos decrescentes, ou seja, uma média anual positiva do número de publicações na fase inicial da carreira acadêmica e, declínio após atingir um pico máximo (WEISS; LILLARD, 1982).

⁵ Os economistas Schultz (1961) e Becker (1962) realizaram estudos analisando como o investimento em capital humano varia ao longo do ciclo de vida dos indivíduos. Como resultados, encontraram que, dada o tempo finito de vida, o investimento atinge um pico e depois declina, formando uma curva em forma de U invertido.

Segundo Stephan (1996), apesar dos modelos apresentarem diferentes suposições sobre a formulação da função objetivo do cientista, seus resultados são similares. Conclui-se, portanto, que o estoque de capital do pesquisador atingiria um pico declinando posteriormente. Tal tendência seguida, também, pelo nível de publicação ao longo do ciclo de vida, porém sendo atenuado quanto maior for a satisfação do cientista derivada de sua atividade de pesquisa.

Levin e Stephan (1991), Turner e Mairesse (2005) e Gonzalez-Brambila e Veloso (2007) também realizaram trabalhos empíricos que ratificaram as proposições teóricas do modelo de ciclo de vida. O resultado em comum desses trabalhos é a tendência dos cientistas a ser tornarem menos produtivos à medida que envelhecem. Especificamente, no trabalho de Levin e Stephan (1991) foram analisados cientistas de seis áreas do conhecimento e todos, com exceção da área de Física de Partículas, apresentaram menores produções na fase final de suas carreiras acadêmicas. Já para os cientistas franceses (TURNER; MAIRESSE, 2005) o pico de produtividade ocorreu em torno de 52 anos de idade e para os cientistas mexicanos (GONZALEZ-BRAMBILA; VELOSO, 2007) esse pico ocorreu aos 53 anos de idade.

Posteriormente, dois fatos relevantes para o aprimoramento da teoria foram encontrados: i) os pontos dos picos variam entre as diferentes áreas do conhecimento; ii) há um efeito marginal significativo da reputação, ainda que pequeno, na determinação do número de citações, enquanto que o mesmo não se verifica para o número de publicações. Para o caso de cientistas brasileiros, utilizando uma base de dados composta por 11.416 pesquisadores no período de 2000 a 2008, Kannebley Jr., Carolo e De Negri (2013) apresentaram estimativas as quais revelaram que o pico da produtividade de um pesquisador no Brasil ocorre em torno dos 57 anos de idade.

Entretanto, conforme argumentado em Stephan (1996), os modelos de Ciclo de Vida não podem ser considerados a “pedra angular” para a modelagem do comportamento dos cientistas. Suas restrições para com esse modelo são, principalmente, devido à dificuldade de incorporação satisfatória do processo de aquisição de conhecimento científico, das limitações no reconhecimento da importância que a “prioridade” da descoberta tem para os cientistas e a relevância que a captura de recursos financeiros e humanos tem para o desenvolvimento das pesquisas⁶. Adicionalmente, os efeitos da idade também são confundidos com o efeito geracional do pesquisador. O efeito de geração ocorre em razão das diferentes bases de

⁶ Ou seja, considera importante não apenas os recursos materiais para a pesquisa, mas a montagem de redes de pesquisa, que envolvem não apenas coautores graduados, mas também estudantes de doutorado e participantes de programas de pós-doutorado.

conhecimento que os cientistas de diversas idades possuem em cada campo científico. Assim outro aspecto que merece especial atenção nas estimações de modelos de ciclo de vida são os chamados “efeitos geracionais”. A inclusão da relação entre produtividade acadêmica e os efeitos de geração é desejada, pois insumos de pesquisa e a “safra” do pesquisador podem variar e tornar certas gerações mais produtivas que outras.

Levin e Stephan (1991) estimaram um modelo controlando idade e a geração do cientista. Seus resultados, todos estatisticamente significantes, indicaram que não há evidências de que gerações mais recentes, que teoricamente possuem conhecimento mais atualizado, sejam mais produtivas que gerações mais antigas. Ademais, em vários subcampos do conhecimento pode-se encontrar que cientistas de gerações mais recentes foram menos produtivos que os mais velhos.

2.1.c Gênero

Cole e Cole (1973), Turner e Mairesse (2005) e Fox (2005), entre outros, destacaram a importância do gênero na produtividade individual do pesquisador e concluíram que mulheres possuem menos publicações e são menos citadas quando comparadas com os homens. Tal fato marcando uma grande diferença de produtividade entre gêneros. No que diz respeito à produtividade científica, a literatura aponta dois principais motivos que podem fazer com que as mulheres publiquem menos que os homens. Primeiro, as características pessoais específicas, peculiaridades da família, e quantidade de tempo alocada em pesquisa. Segundo, escassez de oportunidades sofridas por parte das mulheres, tais como as advindas das decisões de chefia nas contratações e destinação dos recursos oriundos de financiamentos (XIE; SHAUMAN, 1998 e 2003).

Esse diferencial de gênero tem diminuído com o passar dos anos. Os resultados de Xie e Shauman (1998) apontaram que, sendo a medida de produtividade a razão entre o número de publicações de mulher e homem, a produção das mulheres em relação aos homens aumentou de 60% no final dos anos 60 para 82% em 1993. A grande razão para tal fato é a tendência de distribuição mais equitativa, entre os sexos, dos recursos e dos cargos de chefia. Fox (2005) encontrou que mulheres publicaram, ou obtiveram aceites de suas publicações, em média, 8,9 artigos no período de 1990-1993, enquanto que os homens atingiram o patamar, na média dos três anos, de 11,4 artigos.

2.1.d Efeito de Mateus

A observação da lei de Lotka também pode ser vista como consequência do sistema de reconhecimento e de financiamento à pesquisa. Um estudo realizado por Stephan et al. (2007) apontou comparações entre o desenvolvimento científico e tecnológico em uma amostra de 10.962 acadêmicos americanos no período de 1990 a 1995. Desses, 90,1% apresentaram produção zero de patentes, enquanto 14,4% não tiveram produção de artigos. Uma denominação para tal fenômeno foi dada por Merton (1968) como “Efeito de Mateus” – “A todo aquele que tem, será dado mais, e terá em abundância. Mas ao que não tem, até o que tem, lhe será tirado.” – que consiste na explicação de uma “vantagem cumulativa” ao longo da carreira.

This is expressed in the principle of cumulative advantage that operates in many systems of social stratification to produce the same result: the rich get richer at a rate that makes the poor become relatively poorer. (Robert K. Merton, 1968, *The Matthew Effect in Science*, p.7).

Esse efeito enuncia o fato de que os cientistas com melhor reputação possuem maior reconhecimento, credibilidade e visibilidade quando comparado aos cientistas iniciantes ou com pouca visibilidade no meio acadêmico. Em outras palavras, ocorre um mecanismo de retroalimentação, o qual reforçaria a criação de um sistema assimétrico na distribuição e dispersão da produção científica e tecnológica nas diversas áreas do conhecimento.

Tal cenário também se aplica ao nível de centros de pesquisa, ou seja, grande alocação de capital em alguns centros, que são conhecidos por sua excelência, e a falta de recursos por parte daqueles que não obtiveram esse reconhecimento. Além disso, é possível explicar o processo de atração de capital humano especializado – alunos e/ou pesquisadores – a esses centros de excelência em pesquisa e a carência desse tipo de alunos e profissionais em outros centros. Como se pode notar, a vantagem cumulativa tem efeitos ambíguos: beneficiando um grupo de pesquisadores privilegiados, por meio do reforço competitivo em direção aos mais fortes, em detrimento dos demais.

2.2. Determinantes Coletivos da Produtividade Científica

O estudo de Carayol e Matt (2004) afirma que a produtividade do pesquisador pode se diferenciar quando são levadas em consideração as características do ambiente em que o mesmo está inserido, devido às externalidades geradas pelo grupo dentro das infraestruturas de pesquisa. Dado que a atividade de pesquisa, raramente, é realizada de forma isolada (Stephan, 2010).

Desta maneira, além dos determinantes tratados anteriormente, os quais estão relacionados às características e decisões dos próprios indivíduos (variáveis de controle interno), deve-se considerar a avaliação dos determinantes coletivos, como fatores explicativos da produtividade individual. Enfim, trata-se das características da infraestrutura de pesquisa na qual o pesquisador está inserido e as suas externalidades geradas.

2.2.a Capital Físico e Humano

Ao nível dos centros de pesquisa, as características do capital humano e físico são consideradas como fatores que influenciam a produtividade dos pesquisadores. No que tange a composição do capital físico, Stephan (2012) ratifica que o acesso a equipamentos e materiais é de suma importância para os cientistas, pois afeta significativamente a sua produtividade. Desta forma, destaca o papel da tecnologia, por meio do desenvolvimento de novos equipamentos, como viabilizadora do avanço da pesquisa:

The key role that equipment plays is one reason to stress what is sometimes referred to as the nonlinear model: scientific research leads to advances in technology, but it is new technology that often brings about advances in science. (STEPHAN, 2012, p. 84).

Devido aos diferentes objetivos da pesquisa científica, as infraestruturas tendem a se diferenciar. Tal fato é resumido por Stephan (2012) na seguinte passagem:

The role of equipment provides another dimension. In some fields, the equipment required to do research is fairly minimal, as in the case of

certain areas of math, chemistry, and fluid physics. In others, research is almost entirely organized and defined by equipment, as in the case of astronomy and high-energy experimental physics. (STEPHAN, 2012, p. 63)

Os materiais de pesquisa (células, reagentes, telescópios, etc.) surgem ao propiciar meios e incentivos para o desenvolvimento de novos estudos/descobertas. Porém, apesar dos materiais serem comprados pelos laboratórios de pesquisa, cientistas de outros centros de estudo têm a possibilidade de acesso por meio de um processo de troca e/ou parcerias, sendo tal prática tradicional no meio científico. Porém, é importante notar, mais uma vez, para a concentração da produção científica e tecnológica agora oriunda das barreiras ao acesso a equipamentos e materiais especializados, ou seja, um aumento na especialização e preço desses instrumentos de produção tende a concentrar a pesquisa em poucos centros e, portanto, nas mãos de poucos cientistas.

No que tange ao capital humano presente nos centros de pesquisa, a literatura destaca as seguintes características como fatores influentes na produtividade dos pesquisadores: composição do laboratório; diferentes tipos de cientistas (cargos e idade); presença de pesquisadores externos; e nível de qualidade dos pesquisadores internos e de suas publicações. A composição do laboratório envolve pesquisadores permanentes, não permanentes e não pesquisadores⁷. Nesse sentido, Carayol e Matt (2004) constataram que pesquisadores permanentes são mais produtivos em termos de produção de patentes e publicações, os pesquisadores não-permanentes são mais importantes na atividade de patenteamento dos laboratórios, e, por fim, os indivíduos não pesquisadores aumentam significativamente os resultados de publicação dos laboratórios. Adicionalmente, Horta e Lacy (2011) constataram que as unidades de pesquisa que possuem estudantes de doutorado e pós-doutorado são aquelas cuja equipe é mais produtiva. Quanto aos pesquisadores externos às infraestruturas, Carayol e Matt (2006) encontraram que nem os doutores nacionais, nem estrangeiros possuem influência significativa na produtividade dos pesquisadores permanentes. Porém, quanto aos produtos da pesquisa – artigos e/ou patentes – encontraram que dentro do laboratório, a presença de pesquisadores externos afetou significativamente os permanentes. Ou seja, os estrangeiros impactaram a produção científica e os nacionais influenciaram a produção tecnológica.

⁷ Os pesquisadores permanentes são compostos por professores (dedicados ao ensino e pesquisa) e pesquisadores dedicados exclusivamente à pesquisa, enquanto os pesquisadores não permanentes envolvem os alunos de pós-graduação. Os indivíduos não pesquisadores correspondem aos técnicos e à equipe administrativa.

Martín-Sempere, Rey-Rocha e Garzón-García (2002) verificaram que o número de pesquisadores com maior nível de senioridade dentro do grupo parece afetar mais significativamente a produtividade do que o tamanho da equipe. Bonaccorsi e Daraio (2003) encontraram que a idade média dos laboratórios reflete a atratividade e vitalidade científica dos laboratórios por meio do círculo virtuoso, ou seja, instituições de pesquisa com mais prestígio atraem maiores recursos e proporcionam mais cargos para jovens pesquisadores. Além disso, atraem mão de obra melhor qualificada e, por conseguinte, propicia a compra de materiais, equipamentos, etc. Mairesse e Turner (2005) encontraram que a produtividade dos colegas impacta fracamente na produção individual, porém a qualidade das publicações dos companheiros de laboratório afeta fortemente a qualidade média das publicações individuais.

Fatores ligados às características qualitativas dos centros de pesquisa – qualidade das universidades, prestígio e a reputação dos laboratórios – são citados como colaboradores na produtividade dos pesquisadores. O desempenho produtivo destes é maior em instituições de mais prestígio quando comparado com os que trabalham em instituições menos qualificadas pelo meio acadêmico (COLE; COLE, 1973). Allison e Longo (1990) analisaram 179 mudanças de emprego dos cientistas, a fim de capturar os efeitos em seus desempenhos. Quando houve transferência para centros mais conceituados do que o anterior foi observado aumentos nas taxas de publicação e de citações dos pesquisadores. Já aqueles que realizaram movimento inverso demonstraram diminuições substanciais na sua produtividade. Tal fato foi denominado como um efeito departamental. Em relação à consolidação dos centros de pesquisa, Martín-Sempere, Rey-Rocha e Garzón-García (2002) apontam que pesquisadores pertencentes a grupos consolidados são mais produtivos do que aqueles que se declararam fazer parte de grupos não consolidados. Esses, por sua vez, são mais produtivos que os pesquisadores sem um grupo de pesquisa específico ou que trabalham sozinhos.

Outra particularidade dos laboratórios apontada na literatura como fator determinante para a produtividade individual é o tamanho da infraestrutura. A principal questão está relacionada com o retorno de escala na produção do conhecimento. Bonaccorsi e Daraio (2003) concluíram que tamanho do laboratório e a produtividade não se correlacionam positivamente e, especificamente, nas áreas de química, ambiental e engenharia apresentou retornos decrescentes na produção científica. Ademais, é constatado que, em quase todos os campos da ciência, os laboratórios menores são mais produtivos. Entre os motivos que sustentam essa hipótese pode-se citar: baixos custos de coordenação, processos de decisão mais ágeis e menores volume de atividades administrativas. Mesmo resultado encontrado por Carayol e Matt (2006), os quais usaram o número de pesquisadores permanentes como *proxy*

para o tamanho do laboratório. Já Mairesse e Turner (2005) encontraram que o impacto do tamanho possui relação decrescente com a produção média de artigos, porém ligeiramente positiva no número de citações.

Por fim, as atividades de ensino e pesquisa desenvolvidas pelos cientistas possuem efeitos ambíguos, ou seja, são complementares (o fazer de uma reforça a outra) ou concorrentes, surgindo uma questão conflitante oriunda das expectativas e obrigações geradas no seu desenvolvimento. Desta forma, espera-se que quanto maior o tempo alocado nas atividades de pesquisa pelos cientistas, em detrimento da atividade de ensino, maior deve ser a sua produtividade acadêmica. Fox (1992) pesquisou quatro áreas da ciência (economia, ciência política, psicologia e sociologia), em diversos departamentos universitários dos EUA, a fim de analisar como a pesquisa e o ensino influenciam a produtividade dos pesquisadores/professores. As conclusões geradas foram que essas atividades são concorrentes, já que os pesquisadores mais produtivos se dedicam menos a atividade de ensino (preparação de cursos, por exemplo), além de as considerarem de menor importância quando comparada a atividade de pesquisa. Por outro ponto de vista, Carayol e Matt (2004) argumentam que centros de pesquisa que distribuem seu capital humano de forma igualitária entre as atividades de ensino e pesquisa tende a ser mais produtivos que os formados apenas por pesquisadores de tempo integral. A explicação deve-se ao fato dos professores manterem contatos com alunos de doutorado e direcioná-los, portanto, à pesquisa.

2.2.b Colaboração

Para pesquisadores a manutenção de altos níveis de produtividade dos pesquisadores é essencial para manter uma boa reputação no meio acadêmico e garantir financiamentos e motivação para projetos futuros. Contudo, as pessoas possuem capacidades cognitivas limitadas. Nesse sentido, a colaboração entre os pesquisadores aparece como solução para essa limitação individual por meio da divisão de tarefas, aproveitamento de habilidades específicas e compartilhamento de equipamentos, materiais e recursos financeiros. Conforme disposto em Li, Liao e Yen (2013, p.1515-1516): ““Such an interconnected chain of relationships constitutes a social network in which valuable resources are shared in the forms of information, understanding, and knowledge through the conduct of social interactions””.

Stephan (2012) argumenta que a colaboração na ciência ocorre com frequência em ambientes como os centros de pesquisa, os quais facilitam a troca de ideias e incentivam a especialização dos pesquisadores. A autora destaca como vantagens da colaboração entre os cientistas, tais itens:

1. Aumento da interdisciplinaridade na pesquisa;
2. Revolução da Tecnologia da Informação (TI) diminuindo os custos de comunicação;
3. Compartilhamento de equipamentos devido às suas crescentes complexidades, preços e especializações, tais como telescópios, embarcações, etc.;
4. Compartilhamento das inúmeras bases de dados; e
5. Compartilhamento dos riscos, inclusive os financeiros.

Os estudos como os de Price e Beaver (1966), Narin, Hamilton e Olivastro (1997) e Glänzel e Schubert (2001) encontram uma relação positiva entre a produtividade científica e tecnológica e a colaboração entre os pesquisadores. Outros resultados derivados desses estudos demonstram que a colaboração internacional entre universidades provocaram um efeito positivo na qualidade dos artigos. Gonzalez-Brambila, Veloso e Krackhardt (2013) analisaram um conjunto de hipóteses usando o banco de dados de publicações e citações de todos os artigos científicos que tiveram pelo menos um autor do México no período de 1981 a 2002. Tal estudo comprovou que houve maiores publicações tanto para os autores que formaram parcerias com cientistas de elite quanto para os que colaboraram com maior frequência com pesquisadores de outros campos do conhecimento. Desta forma, confirmando a teoria de que a cooperação externa desencadeia uma produtividade maior.

3 BASE DE DADOS

Esse trabalho reúne informações das características individuais e coletivas dos pesquisadores pertencentes às infraestruturas, a partir das seguintes fontes de dados:

1. Mapeamento das Infraestruturas de pesquisas do Brasil: realizado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) em parceria com o Ministério da Ciência, Tecnologia e Inovação (MCTI) e o Instituto de Pesquisa Econômica Aplicada (IPEA); e
2. Mapeamento dos pesquisadores: realizado por meio da extração de informações da Plataforma *Lattes* do CNPq.

Esses dois conjuntos de dados permitirão a realização do exercício proposto por este trabalho, o qual consiste na verificação do impacto das características coletivas sobre a produção científica dos coordenadores pertencentes às infraestruturas de pesquisa no Brasil. Desta forma, este capítulo está estruturado em duas subseções, as quais irão expor os procedimentos adotados para a coleta e construção das duas bases de dados, assim como os detalhes convenientes para o escopo deste trabalho.

3.1. Mapeamento das Infraestruturas

O Mapeamento das infraestruturas de pesquisas do Brasil foi realizado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) em 2013, com apoio do Ministério da Ciência, Tecnologia e Inovação (MCTI) e do Instituto de Pesquisa Econômica Aplicada (IPEA). Primeiramente, para fins de conceituação, “infraestrutura de pesquisa” é definida como o conjunto de instalações físicas e condições materiais de apoio – equipamentos, recursos e serviços – utilizados pelos pesquisadores para a realização de atividades de Pesquisa e Desenvolvimento. As configurações de infraestruturas de pesquisa presentes na base são: laboratórios, planta ou usina piloto, observatórios, estações ou rede de monitoramento e navio de pesquisa ou laboratório flutuante.

O processo de identificação de tais infraestruturas foi realizado pelo IPEA por meio de consultas aos *sites* de unidades de pesquisa (UPs), organizações sociais (OS) vinculadas ao MCTI e institutos de pesquisa vinculados à Comissão Nacional de Energia Nuclear (CNEN). A coleta de informações foi realizada por meio de um formulário informatizado (Anexo A), o qual foi enviado por *e-mail* para 4.961 infraestruturas de pesquisa. O preenchimento foi voluntário pelos coordenadores dos centros de pesquisa. Calcula-se como taxa de resposta, o percentual de 35,48%, entretanto há ausência de informações sobre o universo de infraestruturas de pesquisas que existem no Brasil, motivo pelo qual desencadeou tal levantamento pelos órgãos públicos envolvidos.

Portanto, o conjunto de dados recolhidos contém informações das características coletivas de 1.760 infraestruturas de pesquisa, as quais são compostas por 7.714 pessoas. Resumidamente, o questionário, que serviu como instrumento de coleta está dividido em quatro módulos:

- 1º módulo: Caracterização – possui as informações gerais, tais como: identificação, coordenação, descrição, áreas de atuação, linhas de pesquisa.
- 2º Módulo: Operação da Infraestrutura – informações sobre equipe técnico-científica da estrutura, utilização das instalações e equipamentos por usuários externos e quais as principais atividades e tipos de cooperação desenvolvidos;
- 3º Módulo: Custos Operacionais e Fontes de Receitas – informações sobre valores estimados da infraestrutura e as fontes de receitas e custos operacionais;
- 4º Módulo: Avaliação do Capital Humano e Físico – avaliação subjetiva do coordenador sobre as condições atuais da infraestrutura pesquisada – capacidade técnica, instalações físicas, equipamentos, manutenção, insumos de pesquisa e equipe técnica da estrutura⁸.

3.1.a Características Gerais

Iniciando a análise descritiva pelo módulo 1 do questionário, pode-se caracterizar as infraestruturas como, em maioria, de configurações do tipo de laboratório.

⁸ Foi utilizada a escala de Likert.

Tabela 1 – Infraestruturas por tipo de configuração.

Tipo da Infraestrutura	Quantidade
Estação ou Rede de Monitoramento	8
Laboratório	1.719
Navio de pesquisa ou Laboratório Flutuante	1
Observatório	12
Planta ou Usina piloto	20
Total	1.760

Fonte: Elaboração própria, a partir do Mapeamento do MCTI.

Tais infraestruturas estão distribuídas em todo o território brasileiro (Tabela 2), porém situam-se, em sua maioria, na região Sudeste do país – São Paulo (24,09%), Rio de Janeiro (19,43%) e Minas Gerais (12,9%).

Tabela 2 – Distribuição regional das infraestruturas.

Região	Quantidade
Centro-Oeste	113
Nordeste	170
Norte	54
Sudeste	1.004
Sul	419
Total	1.760

Fonte: Elaboração própria, a partir do Mapeamento do MCTI.

A idade média é de 15 anos, sendo 56,65% das infraestruturas criadas a partir dos anos 2000 (Tabela 3). Segundo Squeff e De Negri (2014), isso pode estar relacionado ao ciclo recente de investimentos do país em Ciência, Tecnologia & Inovação por meio de Leis e Decretos que buscam ““fortalecer a infraestrutura e serviços de apoio à pesquisa técnico-científica desenvolvida em instituições públicas de ensino superior e de pesquisa brasileiras””.

Tabela 3 – Ano de início de operação das infraestruturas.

Ano de início	Quantidade	%
Antes de 1970	50	2,84
1970-1979	110	6,25
1980-1989	193	10,96
1990-1999	410	23,30
2000-2009	654	37,15
2010-2012	343	19,50
Total	1.760	100

Fonte: Elaboração própria, a partir do Mapeamento do MCTI.

As 1.760 infraestruturas podem ser caracterizadas segundo as grandes áreas do conhecimento⁹ em que atuam. Conforme Tabela 4 estão, predominantemente, nas áreas de Engenharias (28%), Ciências Exatas e da Terra (21%), Ciências Biológicas (18%) e Ciências Agrárias (12%). Além disso, há uma parcela significativa (18%) de infraestruturas que atuam em mais de uma grande área do conhecimento, desta forma optou-se por agrupá-las e nomeá-las como *Multiárea*.

Tabela 4 – Distribuição das infraestruturas por grandes áreas do conhecimento das infraestruturas.

Grande Área	No. De Infraestrutura	%
Ciências Exatas e da Terra	363	21%
Ciências Biológicas	312	18%
Engenharias	485	28%
Ciências da Saúde	67	4%
Ciências Agrárias	206	12%
Ciências Sociais Aplicadas	2	0,0011%
Ciências Humanas	2	0,0011%
Multiárea	323	18%
TOTAL	1.760	100%

Fonte: Elaboração própria, a partir do Mapeamento do MCTI.

Diante do exposto pela Tabela, nota-se a presença de infraestruturas das áreas de Ciências Sociais Aplicadas e Ciências Humanas. Entretanto, decidiu-se excluí-las no momento das estimações dos modelos, pois não eram prioritárias para a pesquisa e, além disso, as questões dos questionários não eram adequadas para essas áreas do conhecimento.

A titulação da mão de obra é descrita na Tabela 5, na qual é possível observar que 89,21% dela possui pós-graduação (doutorado ou mestrado ou mestrado profissionalizante).

Tabela 5 – Mão de obra, por titulação.

Titulação Máxima	Qtd.	%
Doutorado	5.698	73,87%
Mestrado	1.163	15,08%
Mestrado Profissionalizante	20	0,26%
MBA	14	0,18%
Especialização	131	1,70%
Especialização Médica	2	0,03%
Graduados e Não-graduados	686	8,88%
Total	7.714	100,00%

Fonte: Elaboração própria, a partir do Mapeamento do MCTI.

⁹ Segundo o CNPq, define-se como grandes áreas do conhecimento: Ciências Exatas e da Terra; Ciências Biológicas; Engenharias; Ciências da Saúde; Ciências Agrárias; Ciências Sociais Aplicadas; Ciências Humanas; e Linguística, Letras e Artes.

A Tabela 6 apresenta a distribuição da mão de obra, segundo o tipo de vínculo com a instituição¹⁰. Como é possível notar a maioria da mão de obra possui vínculos formais com as instituições: 74,08% são servidores públicos ou celetistas, seguido pelos 16,49% como bolsistas.

Tabela 6 – Mão de obra, por tipo de vínculo.

Titulação de Vínculo	Qtd.	%
Servidor Público	5.447	60,26%
Celetista	1.248	13,81%
Bolsista	1.487	16,45%
Pesquisador visitante	255	2,82%
Outro	602	6,66%
Total	9.039	100,00%

Fonte: Elaboração própria, a partir do Mapeamento do MCTI.

As infraestruturas podem estar vinculadas ou às universidades (públicas/ privadas); ou aos institutos técnicos; ou às organizações independentes de caráter público ou privado. Desta maneira, a Tabela 7 demonstra como a mão de obra está distribuída entre esses espaços, sendo que contam com um total de 9.039 pessoas ocupadas, dos quais 917 pessoas trabalham em mais de uma infraestrutura, além da equipe técnica de apoio administrativo (11.830 pessoas).

Ademais, nota-se que as infraestruturas da amostra estão, em maioria, presentes em universidades (68,35%) e do total do capital humano, exceto equipe técnica, 64,79% são doutores e 10,23% são mestres. Desse modo, tem-se que 75,02% dos cientistas brasileiros possuem formação especializada em nível de pós-graduação *stricto sensu*, o que demonstra alto nível na formação da força de trabalho diretamente envolvida na pesquisa. Outro ponto importante a ser destacado, é que os estudantes, assim como os pesquisadores visitantes, estão, em maioria, nas universidades.

¹⁰ Como 917 da mão de obra atuam em mais de uma infraestrutura com diferentes vínculos possíveis, tem-se que o total da Tabela 6 excede o da Tabela 5.

Tabela 7 – Mão de obra das infraestruturas, por titulação e constituição/localização em 2012.

Infraestrutura	Qtd.	Pesquisadores com pós-graduação				Estudantes	Pesquisadores Visitantes	Equipe de Apoio
		Doutorado	Mestrado/ Mestrado Profissionalizante	MBA	Especializações			Técnico e administrativo
Público/Privado	412	1.423	351	7	51	269	80	5.379
Universidades	1.203	4.239	394	2	29	1.214	173	5.756
Inst. Técnicos	145	194	180	4	64	4	2	1.054
TOTAL	1.760	5.856	925	13	144	1.487	255	12.189

Fonte: Elaboração própria, a partir do Mapeamento do MCTI.

3.1.b Operação da Infraestrutura

No que diz respeito às atividades desenvolvidas pelas infraestruturas, têm-se cinco principais ocupações: ensino, pesquisa, desenvolvimento de tecnologias, extensão tecnológica e serviço tecnológico. A realização dessas atividades é ordenada de acordo com a intensidade que as desempenham, que são: contínuo; alguns dias da semana; alguns dias do mês; esporádico; e não desenvolve. Como exemplificação, torna-se interessante realizar uma interação dessas atividades na intensidade “*contínuo*” (Tabela 8), para avaliar como ocorre o desempenho dessas atividades nas infraestruturas.

Tabela 8 – Quantidade de infraestruturas, de acordo com a interação das atividades desenvolvidas na categoria “*contínuo*”.

Atividades	Ensino	Pesquisa	Desenv. Tecnologia	Extensão Tecnológica	Serviço Tecnológico
Ensino	692				
Pesquisa	628	1.421			
Desenv. Tecnologia	298	601	620		
Extensão Tecnológica	95	135	110	140	
Serviço Tecnológico	135	251	251	89	315

Fonte: Elaboração própria, a partir do Mapeamento do MCTI.

Conforme Tabela 8, o levantamento revelou que o principal foco e envolvimento das infraestruturas é a interação de atividades de pesquisa/ensino e pesquisa/desenvolvimento de tecnologia, dado que possuem as maiores concentrações 628 e 601 das infraestruturas respectivamente.

Ao realizar a análise por área do conhecimento da infraestrutura (Tabela 9), nota-se que as atividades de ensino e pesquisa são realizadas pela maioria delas. Quanto às atividades de desenvolvimento de tecnologia destacam-se as áreas de Engenharias (78,76%) e Multiárea (76,62%). Já as atividades de extensão tecnológica são praticadas, em maioria, pelas infraestruturas de Ciências Agrárias (55,83%) enquanto que as infraestruturas Multiárea (65,85%) são aquelas com maior participação na realização de Serviços Tecnológicos.

Tabela 9 – Atividades realizadas, segundo a grande área de atuação das infraestruturas.

Atividades	Grande Área do Conhecimento							
	Exatas e da Terra	Biológicas	Engenharias	Saúde	Agrárias	Sociais Aplicadas	Humanas	Multi-área
Ensino	77,41%	91,35%	83,71%	94,03%	91,75%	100%	100%	86,77%
Pesquisa	96,42%	98,40%	90,93%	94,03%	98,06%	100%	100%	96,62%
Desenv. Tecnologia	66,67%	51,92%	78,76%	56,72%	60,68%	100%	50%	76,62%
Ext. Tecnológica	41,60%	35,58%	46,60%	40,30%	55,83%	100%	50%	49,85%
Serv. Tecnológico	58,13%	43,91%	62,27%	53,73%	60,68%	50%	50%	65,85%
Outras Atividades	4,96%	5,45%	3,09%	4,48%	2,43%	0%	0%	6,15%

Fonte: Elaboração própria, a partir do Mapeamento do MCTI.

Outra informação significativa obtida no levantamento é que a ampla maioria das infraestruturas é multiusuária, ou seja, aberta à utilização para usuários externos¹¹. A quantidade de usuários externos no ano de 2012 totalizou 38.703, sendo 36.985 pesquisadores de outras instituições e empresas brasileiras e 1.718 pesquisadores oriundos de instituições e empresas do exterior.

Tabela 10 – Distribuição das infraestruturas, de acordo com a realização ou não de cooperação com outras infraestruturas e/ou empresas do Brasil ou do exterior.

Cooperação	No. De Infraestruturas	%
Com cooperação	1.560	88,64
Sem cooperação	200	11,36
Total	1.760	100

Fonte: Elaboração própria, a partir do Mapeamento do MCTI.

Do total de infraestruturas, 66,65% obtiveram presença de externos brasileiros e 17,73% contaram com a presença de usuários de outros países. Ainda no aspecto de relacionamento externo, no quesito de cooperação, 88,64% das infraestruturas realizaram algum tipo de interação entre centros de pesquisa e/ou empresas do Brasil ou do exterior (Tabela 10).

Quanto à cooperação segundo a área do conhecimento (Tabela 11), destacam-se as infraestruturas pertencentes às áreas de Ciências Humanas (100%), Ciências Biológicas

¹¹ Usuário externo é definido como todos os pesquisadores e alunos que não fazem parte da equipe permanente, incluindo aqueles vinculados à mesma instituição que abriga a infraestrutura.

(94,55%) e Multiárea (94,12%). Por outro lado, as infraestruturas pertencentes às áreas de Ciências Sociais Aplicadas (50%) apresentam as menores taxas de relações de cooperação.

Tabela 11 – Distribuição (%) das infraestruturas que realização cooperação segundo área do conhecimento de atuação.

Grande Área	%
Ciências Exatas e da Terra	89,81
Ciências Biológicas	94,55
Engenharias	83,09
Ciências da Saúde	77,61
Ciências Agrárias	85,92
Ciências Sociais	50,00
Ciências Humanas	100,00
Multiáreas	94,12

Fonte: Elaboração própria, a partir do Mapeamento do MCTI.

3.1.c Custos Operacionais e Fontes de Receitas

O terceiro módulo do questionário trata dos valores estimados do capital físico (estrutura e equipamentos) e as principais fontes de financiamento. Desta maneira, a Tabela 12 discrimina a distribuição das infraestruturas segundo as faixas de valores estimados das estruturas e dos equipamentos de pesquisa, em valores de 2012¹². Como se pode notar, há uma forte correlação entre o valor da infraestrutura¹³ e o valor dos equipamentos disponíveis, ou seja, 60% e 63% das instituições de pesquisas estão localizadas na estimativa até R\$500 mil reais no valor total de suas infraestruturas e no valor médio de seus equipamentos de pesquisa, respectivamente.

¹² É importante deixar claro que não são valores precisos, dado que é resultado de estimativas realizadas pelos coordenadores de cada infraestrutura, e não de um levantamento contábil de todos os equipamentos.

¹³ O valor das infraestruturas incluem os valores estimados das instalações físicas, dos equipamentos, do mobiliário, entre outros.

Tabela 12 – Distribuição das infraestruturas pesquisadas segundo faixa de valor estimado das infraestruturas e dos equipamentos de pesquisa (em R\$ de 2012).

Faixa do valor total estimada da infraestrutura e do valor médio dos equipamentos de pesquisa (em R\$, 2012).	No. De Infraestrutura por valor total estimado da infraestrutura	No. De Infraestrutura por valor médio dos equipamentos de pesquisa
Até R\$ 500 mil	1.054	1.110
Acima de R\$ 500 mil até R\$ 1 milhão	301	250
Acima de R\$ 1 milhão até R\$ 3 milhões	222	260
Acima de R\$ 3 milhões até R\$ 5 milhões	73	51
Acima de R\$ 5 milhões até R\$ 10 milhões	55	51
Acima de R\$ 10 milhões até R\$ 20 milhões	33	19
Acima de R\$ 20 milhões até R\$ 30 milhões	11	3
Acima de R\$ 30 milhões até R\$ 50 milhões	2	2
Acima de R\$ 50 milhões	NA	3
Acima de R\$ 50 milhões até R\$ 100 milhões	2	NA
Acima de R\$ 100 milhões até R\$ 200 milhões	4	NA
Acima de R\$ 200 milhões	2	NA
Não informado	1	1
TOTAL	1.760	1.760

Fonte: Elaboração própria, a partir do Mapeamento do MCTI.

Em relação às grandes áreas do conhecimento (Tabela 13), nota-se a concentração da maioria das infraestruturas na menor faixa de valor que vai até R\$ 500 mil. Por outro lado, na estimativa de valores mais elevados (acima de R\$ 200 milhões) encontram-se apenas duas infraestruturas, uma da área de Ciências Exatas e da Terra (0,28%) e outra de Multiárea (0,31%). As infraestruturas de Engenharias e Multiárea são mais dispersas entre as faixas de valores estimados.

A Tabela 14 descreve as estimativas dos custos operacionais e das receitas anuais das infraestruturas de pesquisa analisadas por faixa de valores, em R\$ de 2012. Quanto aos custos operacionais anuais, pode-se observar que 38,69% dos coordenadores dos centros de pesquisa responderam que seus custos operacionais estão na faixa de até R\$ 50 mil/ano, enquanto que 5,40% das infraestruturas estão nas faixas acima de R\$ 1 milhão/ano. Além disso, 13,92% não souberam ou não responderam a questão. Uma explicação deve-se ao fato das infraestruturas serem tratadas de forma agregada nas instituições as quais pertencem. No que diz respeito às estimativas das receitas anuais, 24,49% estão na faixa de até R\$ 50 mil e apenas 7,38% encontram-se nas faixas superiores de acima de R\$ 1 milhão. Nesse quesito, 34,21% não souberam ou não responderam.

Tabela 13 – Estimativa do valor das infraestruturas, segundo a área de conhecimento em que atuam (em R\$ de 2012).

Estimativas dos valores das infraestruturas	Ciências Exatas e da Terra	Ciências Biológicas	Engenharias	Ciências da Saúde	Ciências Agrárias	Ciências Sociais Aplicadas	Ciências Humanas	Multiárea
Até R\$ 500 mil	230	200	276	43	131	2	2	170
Acima de R\$ 500 mil até R\$ 1 milhão	60	63	64	10	47	-	-	57
Acima de R\$ 1 milhão até R\$ 3 milhões	40	31	74	8	24	-	-	45
Acima de R\$ 3 milhões até R\$ 5 milhões	14	10	30	2	1	-	-	16
Acima de R\$ 5 milhões até R\$ 10 milhões	12	4	19	3	1	-	-	16
Acima de R\$ 10 milhões até R\$ 20 milhões	6	2	14	1	1	-	-	9
Acima de R\$ 20 milhões até R\$ 30 milhões	0	2	4	0	0	-	-	5
Acima de R\$ 30 milhões até R\$ 50 milhões	0	0	1	0	0	-	-	1
Acima de R\$ 50 milhões até R\$ 100 milhões	0	0	2	0	0	-	-	0
Acima de R\$ 100 milhões até R\$ 200 milhões	0	0	1	0	0	-	-	3
Acima de R\$ 200 milhões	1	0	0	0	0	-	-	1
Não respondido	-	-	-	-	1	-	-	
Total	363	312	485	67	206	2	2	323

Fonte: Elaboração própria, a partir do Mapeamento do MCTI.

Tabela 14 – Estimativa dos custos operacionais e receitas anuais das infraestruturas (em R\$ de 2012).

Faixa do valor total estimada (em R\$, 2012).	No. De Infraestrutura por estimativa dos custos operacionais anuais		No. De Infraestrutura por estimativa das receitas anuais	
		%		%
Até R\$ 50 mil	681	38,69	431	24,49
Acima de R\$ 50 mil até R\$ 100 mil	268	15,23	184	10,45
Acima de R\$ 100 mil até R\$ 150 mil	106	6,02	82	4,66
Acima de R\$ 150 mil até R\$ 200 mil	75	4,26	67	3,81
Acima de R\$ 200 mil até R\$ 300 mil	82	4,66	77	4,38
Acima de R\$ 300 mil até R\$ 500 mil	95	5,40	90	5,11
Acima de R\$ 500 mil até R\$ 750 mil	64	3,64	46	2,61
Acima de R\$ 750 mil até R\$ 1 milhão	49	2,78	51	2,90
Acima de R\$ 1 milhão até R\$ 2 milhões	57	3,24	59	3,35
Acima de R\$ 2 milhões até R\$ 5 milhões	27	1,53	46	2,61
Acima de R\$ 5 milhões	11	0,63	25	1,42
Não é possível estimar	244	13,86	601	34,15
Não respondido	1	0,06	1	0,06
Total	1.760	100	1.760	100

Fonte: Elaboração própria, a partir do Mapeamento do MCTI.

O levantamento sobre financiamento também pode ser feito de acordo com as grandes áreas e as suas principais fontes financiadoras. Em valores de 2012, os dados revelam que a Petrobras é a maior fonte de receitas médias anuais, seguido pelo orçamento das próprias instituições, pela Financiadora de Estudos e Projetos (FINEP), as Fundações Estaduais de Amparo à Pesquisa (FAPs), o CNPq, as empresas privadas e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Tabela 15 – Financiamento recebido, por entidade financiadora (em R\$ médio de 2012).

Áreas de conhecimento	Fonte de Receita (valor médio em R\$)						Empresas privadas
	Orçamento próprio	CNPq	CAPES	FINEP	FAPs	Petrobras	
Ciências Exatas e da Terra	133.427,80	75.209,37	19.585,05	102.412,80	90.395,94	238.737,20	22.269,97
Ciências Biológicas	34.847,17	76.141,48	28.764,78	23.044,87	151.934,70	29.966,08	8.317,94
Engenharias	40.393,48	67.266,45	13.917,48	136.513,70	66.629,02	389.413,40	154.924,40
Ciências da Saúde	129.633,10	127.902,60	21.409,81	26.608,15	192.764,00	2.985,07	26.656,72
Ciências Agrárias	25.907,77	28.682,32	17.839,51	32.075,19	49.658,15	23.341,75	14.344,00
Ciências Sociais Aplicadas	4.800,00	25.000,00	2.500,00	0,00	0,00	0,00	0,00
Ciências Humanas	0,00	6.000,00	0,00	0,00	0,00	0,00	0,00
Multiárea	538.240,90	97.842,09	32.073,67	186.908,90	132.413,50	120.626,90	45.208,11
Média Geral	151.920,10	73.896,58	21.812,73	102.128,00	101.621,20	187.270,70	59.886,45

Fonte: Elaboração própria, a partir do Mapeamento do MCTI.

Por área do conhecimento, Ciências Exatas e da Terra e Engenharias possuem os maiores financiamentos advindos da Petrobras. Posteriormente, Ciências Biológicas, Ciências da Saúde e Ciências Agrárias têm suas receitas oriundas, principalmente, das Fundações Estaduais de Amparo à Pesquisa (FAPs). Já as áreas de Ciências Sociais Aplicadas e Ciências Humanas são financiadas quase que exclusivamente pelo CNPq. Por fim, as infraestruturas que se dedicam a mais de uma grande área (Multiárea) têm como receita os repasses das próprias instituições onde estão inseridas.

3.1.d Avaliação do Capital Humano e Físico

Por fim, no quarto módulo, os coordenadores avaliaram de forma subjetiva tanto seu capital físico como humano em uma escala de cinco níveis: *muito bom, bom, regular, ruim e não se aplica/não sei*. Em relação à qualidade das infraestruturas, 64,54% dos coordenadores declararam suas instalações físicas como “*ruim*” e “*regular*”. Na avaliação da qualidade dos equipamentos, 53,46% os classificaram como “*bom*” e “*muito bom*”, enquanto que 45,29% classificaram de “*ruim*” a “*regular*”. Em vista disso, os dados revelam que as condições das instalações físicas e de equipamentos das infraestruturas de pesquisa brasileiras podem ser consideradas de qualidade intermediária. Uma análise importante é saber qual a avaliação das infraestruturas separadas pelas áreas do conhecimento.

Tabela 16 – Distribuição (%) dos coordenadores, os quais avaliaram as condições físicas das infraestruturas como “*muito boa*” e “*boa*”, por a área do conhecimento.

Área de conhecimento	Instalações físicas	Equipamentos	Manutenção	Insumos de pesquisa
Ciências Exatas e da Terra	36,09%	52,89%	74,66%	82,37%
Ciências Biológicas	26,28%	46,47%	69,55%	85,26%
Engenharias	38,56%	56,91%	75,67%	81,86%
Ciências da Saúde	28,36%	53,73%	76,12%	86,57%
Ciências Agrárias	32,68%	50,73%	73,17%	84,88%
Ciências Sociais Aplicadas	50%	50%	100%	100%
Ciências Humanas	0%	0%	50%	50%
Multiárea	36,22%	57,89%	78,64%	86,38%

Fonte: Elaboração própria, a partir do Mapeamento do MCTI.

Desta maneira, a Tabela 16 aponta que as instalações físicas das áreas de Ciências Sociais Aplicadas (50%), Engenharias (38,56%), Multiárea (36,22%) e Ciências Exatas e da Terra (36,09%) estão mais bem avaliadas. Em relação à qualidade dos equipamentos e à manutenção, as infraestruturas Multiárea (57, 89% e 78,64%, respectivamente) foram mais bem avaliadas. Já em relação aos insumos, todos os centros foram bem avaliados, destaque para os centros de pesquisa das áreas de Ciências Sociais Aplicadas, Ciências da Saúde e Multiárea.

Quanto às capacidades das equipes técnicas, os coordenadores deveriam avaliar em relação ao padrão observado em outras infraestruturas do gênero no Brasil e no exterior. Desta forma, 82,02% consideravam suas situações de “*adequados aos padrões brasileiros*” a “*avançados e compatíveis com o exterior*”.

Ademais, houve questionamentos quanto ao número de capital humano disponível e suas formações, os quais deveriam avaliar com base em cinco níveis: *adequado*, *inadequado*, *pouco adequado* e *não se aplica/não sei*. Da Tabela 17, pode-se observar que quanto ao número de pesquisadores, exceto Ciências Humanas e Ciências da Saúde, menos de um quarto das infraestruturas consideram “*adequado*”. Por outro lado, nota-se que a maioria dos centros de pesquisa parece estar satisfeito com o número de técnicos que compõe a sua equipe.

Em relação à formação dos pesquisadores, 50% das infraestruturas de Ciências da Saúde, Ciências Sociais Aplicadas e Ciências Humanas consideram “*adequado*”, enquanto que nas demais tal cenário ocorre em menos da metade. Já a formação dos técnicos, uma grande parcela dos coordenadores avaliou com situação *adequado*.

Tabela 17 – Distribuição (%) dos coordenadores, os quais avaliaram os recursos humanos das infraestruturas como “*adequado*”, por a área do conhecimento.

Área de conhecimento	Nº de pesquisadores	Formação dos pesquisadores	Nº de técnicos	Formação dos técnicos
Ciências Exatas e da Terra	24,41%	44,48%	69,57%	69,23%
Ciências Biológicas	24,49%	48,16%	69,80%	74,29%
Engenharias	24,76%	45,63%	69,90%	72,09%
Ciências da Saúde	28,33%	50,00%	65,00%	75,00%
Ciências Agrárias	24,54%	48,47%	67,48%	72,39%
Ciências Sociais Aplicadas	0,00%	50,00%	100,00%	50,00%
Ciências Humanas	50,00%	50,00%	50,00%	100,00%
Multiárea	21,29%	40,30%	67,68%	72,24%

Fonte: Elaboração própria, a partir do Mapeamento do MCTI.

Por fim, pode-se dizer que a avaliação do capital humano é considerada ruim no sentido de número de pesquisadores e intermediário na formação deles, porém adequado para o número e formação dos técnicos que integram a mão de obra das infraestruturas.

3.2. Mapeamento dos Pesquisadores

A segunda base foi construída, exclusivamente, para o presente trabalho por meio de um robô de busca eletrônica. Tal mecanismo extraiu informações individuais (do período de 2000 a 2013) existentes no sistema de Currículos Lattes, parte integrante da Plataforma *Lattes* do CNPq. Essa plataforma é mantida pelo governo brasileiro a fim de gerenciar as informações dos pesquisadores, das instituições e das atividades de pesquisa do país. Desta forma, torna-se possível realizar uma análise descritiva das características coletivas e individuais dos 7.714 pesquisadores que pertencem às 1.760 infraestruturas da base do MCTI.

A versão atual da plataforma consiste em um formulário eletrônico que é preenchido pelo próprio usuário. Devido a essa autonomia no registro manual das próprias informações, pode-se surgir erros de preenchimento de diversas espécies, desta forma ocasionando inconsistências nas análises e validações de informações. Porém, processos de aperfeiçoamentos contínuos são realizados por parte do CNPq em busca de novos mecanismos de controle que promovam o direcionamento do preenchimento das informações de tal forma a minimizar a necessidade de digitação manual de nomes por parte dos usuários.

Apesar de a base conter informações sobre produções de livros/capítulos, patentes, publicações em jornais, entre outras, no presente trabalho será explorado apenas a produção de artigos publicados para representar a produtividade científica individual. Ademais, serão utilizadas as informações sobre características curriculares dos docentes, tais como: universidade de vínculo; local e ano de titulação de doutorado; tipos e quantidade de orientações.

Em relação às características de ocupação no trabalho dos pesquisadores, pode-se dizer que dos 7.714 cientistas brasileiros distintos, 88,11% realizam atividades em apenas uma infraestrutura e 29,09% dos pesquisadores exercem a função de coordenador (Tabela 18). Um ponto interessante a ser destacado é o alto percentual de pesquisadores trabalhando em uma única infraestrutura, porém um pouco mais da metade se dizem trabalhar em regime de

dedicação exclusiva (57,32%). Uma explicação para tal fato é que os pesquisadores exercem outras atividades que fogem ao escopo das realizadas pelas infraestruturas.

Tabela 18 – Distribuição da mão de obra e de artigos publicados no quinquênio (2009-2013), segundo as características de ocupação nas infraestruturas.

	Qtd. de Cientistas	Qtd. de artigos no quinquênio
Trabalha em mais de 1 infraestrutura	6.797 (88,11%)	58.458 (81,89%)
Trabalha mais de 30h semanais	4.422 (57,32%)	38.516 (53,95%)
Exerce função de coordenador	2.244 (29,09%)	32.816 (45,97%)

Fonte: Elaboração própria, a partir do Mapeamento do MCTI e da Plataforma *Lattes*.

Quanto à produtividade no quinquênio 2009-2013, a Tabela 18 indica que os pesquisadores que trabalham em mais de uma infraestrutura são mais produtivos quando comparados com os que trabalham em apenas uma infraestrutura, pois produziram 81,89% dos artigos do período. No mesmo sentido, os pesquisadores que trabalham mais de 30 horas semanais na mesma infraestrutura, percebe-se que são menos produtivos, dado que são 57,32% da amostra e produzem um pouco mais da metade do total de artigos.

Ainda que haja poucos exercendo a função de coordenador (29,09%), nota-se que eles participam com quase metade do total de publicações no quinquênio analisado. Portanto, torna-se interesse investigar suas características individuais. A Tabela 19 descreve quais são as titulações máximas dos 2.244 coordenadores pertencentes às 1.760 infraestruturas. Com isso, observa-se que 98,27% possuem títulos de pós-graduação.

Tabela 19 – Coordenadores, por titulação máxima.

Titulação Máxima	Quantidade	%
Doutorado	1.998	89,04
Mestrado	170	7,58
Mestrado Profissionalizante	5	0,22
MBA	4	0,18
Especialização	28	1,25
Graduados e Não-graduados	39	1,73
Total	2.244	100

Fonte: Elaboração própria, a partir do Mapeamento do MCTI.

Quando observa-se a Tabela 20, o tipo de vínculo de 87,30% dos coordenadores são regulares (servidor público ou celetista).

Tabela 20 – Coordenadores, por tipo de vínculo.

Tipo de Vínculo	Quantidade	%
Servidor Público	1.666	74,24
Celetista	293	13,06
Outros	96	4,28
Bolsista	161	7,17
Pesquisador Visitante	27	1,21
Null	1	0,04
Total	2.244	100

Fonte: Elaboração própria, a partir do Mapeamento do MCTI.

Pode-se, também, observa como os coordenadores estão divididos entre os anos de início de operação das infraestruturas. Portanto, conforme Tabela 21, 81,06% dos coordenadores estão presente em infraestruturas com três anos ou mais de funcionamento.

Tabela 21 – Coordenadores, por ano de início de operação das infraestruturas.

Ano de início	Quantidade	%
Antes de 1970	68	3,02
1970-1979	130	5,81
1980-1989	286	12,75
1990-1999	556	24,76
2000-2009	797	35,52
2010-2012	407	18,14
Total	2.244	100

Fonte: Elaboração própria, a partir do Mapeamento do MCTI.

A quantidade total de artigos produzidos pelos 7.714 pessoas pertencentes às 1.760 infraestruturas está exposto no Gráfico 1 abaixo. Como é possível observar, há uma produtividade bastante desigual, tal como apontado pela literatura. No período de cinco anos, as 7.714 pessoas publicaram 71.386 artigos, com média de publicação de 9,25 e desvio padrão de 15,43. O número máximo de publicação foi de 391 artigos no período acumulado e o mínimo de zero. Há concentração de 1.772 pessoas com produção zero no período, representando 22,97% da amostra. Além disso, há centralização em pequenos grupos de valores, ou seja, 69,50% do total de pessoas publicaram até a quantidade média de artigos da amostra (9,25) e 9,07% publicaram acima de 25 artigos no quinquênio de análise.

Gráfico 1 – Histograma da quantidade de artigos publicados entre 2009 a 2013 pelo total da mão de obra pertencentes às infraestruturas.

Fonte: Elaboração própria, a partir da Plataforma *Lattes*.

O levantamento da Plataforma *Lattes*, também, disponibiliza a área do conhecimento de atuação do pesquisador. Desta forma, a Tabela 22 descreve como a quantidade total de artigos publicados no quinquênio está distribuída de acordo com as grandes áreas do conhecimento dos pesquisadores. Aqui, também, optou-se pelo agrupamento na opção Multiárea dos pesquisadores que atuaram em mais de um campo do conhecimento.

Tabela 22 – Quantidade de artigos publicados de 2009 a 2013 segundo a grande área do conhecimento de atuação.

Grande Área	Qtd. total (A)	Qtd. de artigos no quinquênio (B)	B/A (%)
Ciências Exatas e da Terra	818	5.472	6,69
Ciências Biológicas	525	3.522	6,71
Engenharias	946	2.570	2,72
Ciências da Saúde	151	1.651	10,93
Ciências Agrárias	225	1.781	7,92
Ciências Humanas	10	11	1,10
Ciências Sociais Aplicada	9	19	2,11
Multiáreas	5.030	56.360	11,20
TOTAL	7.714	71.386	9,25

Fonte: Elaboração própria, a partir da Plataforma *Lattes*.

Conforme descrita na Tabela 22, as pessoas estão concentradas no campo *Multiárea* e, portanto, nota-se nessa área uma maior quantidade de artigos publicados no período de 2009 a 2013. No mesmo sentido, quando se analisa a quantidade média de artigos publicados

nesse período, os pertencentes à *Multiárea*, também tem a maior produtividade, seguida pelas áreas da Ciência da Saúde (10,93) e Ciências Agrárias (7,92).

Quanto às regiões do Brasil (Gráfico 2), os coordenadores com maiores quantidade de artigos publicados no quinquênio (2009-2013), encontram-se na região Sudeste, responsáveis por 62% (44.216 artigos publicados) da produção total, assim como a maioria das infraestruturas da amostra, conforme análise da base de dados do MCTI. Por outro lado, os pesquisadores situados na região Norte apresentaram as menores taxas 2% (1.301 artigos publicados).

Gráfico 2 – Distribuição da quantidade de artigos publicados de 2009 a 2013 pelas infraestruturas, segundo as regiões do Brasil de atuação da mão de obra das infraestruturas.

Fonte: Elaboração própria, a partir do Mapeamento do MCTI e da Plataforma *Lattes*.

Como estratégia para determinar a geração do cientista, utilizou-se a diferença entre o ano de 2013 e a variável que determina o ano de obtenção do título de doutorado. A partir disso, possibilitou-se a criação de quatro faixas de idade de pesquisa, conforme discriminado na Tabela 23.

Tabela 23 – Quantidade de artigos publicados no período de 2009-2013 de acordo com a Faixa de Idade de obtenção do título de doutorado dos pesquisadores.

Faixas de Idade	Qtd. de Coordenadores (A)	Qtd. de artigos no quinquênio (B)	B/A
34-51 anos	147	2.850	19,39
24-33 anos	498	10.228	20,54
14-23 anos	1.476	23.871	16,17
04-13 anos	2.566	26.007	10,14
sem doutorado	3.027	8.430	2,78
Total	7.714	71.386	9,25

Fonte: Elaboração própria, a partir da Plataforma *Lattes*.

A análise por categorização das faixas de idade de pesquisa evidencia que os pesquisadores entre 14 e 23 anos de doutorado, ou seja, aqueles que obtiveram o título no período de 1990 e 1999 possuem a maior quantidade de artigos publicados no quinquênio de análise, enquanto que os mais velhos— doutorado obtido entre 1962 e 1979 – são os que apresentaram as menores quantidades. Já os doutores que obtiveram tal título entre 1980 e 1989 possuem maior produtividade média do período (B/A).

4 ANÁLISE DE CLUSTER: DEFINIÇÃO DE CATEGORIAS DAS INFRAESTRUTURAS

A fim de analisar qual perfil de infraestrutura de pesquisa que propicia uma maior produção científica no Brasil, será adotado como estratégia, a construção de tipologias de infraestruturas que permitem sua caracterização e interpretação. Com isso, cria-se uma ferramenta para analisar seus padrões de organização de pesquisa.

A análise de Cluster é um procedimento da Estatística Multivariada cujo objetivo principal é, por meio de uma análise de relação entre um conjunto de variáveis, agrupar cada observação em subgrupos relativamente homogêneos denominados *clusters*, de tal forma que dentro de cada grupo haja grande similaridade, mas que sejam distintos entre si (RENCHE, 2002). Essa análise também pode ser referida como uma classificação por meio de reconhecimento de padrões, que provavelmente não seria possível serem detectadas sem o referido procedimento. Para Quintal (2006, p.6) “a ideia chave é que os *clusters* tenham significado e sejam interpretáveis”.

Na análise de clusters, os dados são alocados em uma matriz multivariada $n \times p$, $X = [X_{ij}]$, $i = 1, \dots, n$; $j = 1, \dots, p$, a qual contém valores de cada variável observada em cada objeto, sendo X_{ij} o valor da variável j para o objeto i . As variáveis podem ser contínuas ou categorizadas. Aos valores desta matriz multivariada são aplicadas as medidas de proximidade, transformando a matriz inicial X em uma matriz $n \times n$. A essa matriz transformada são aplicados os métodos para a construção de *clusters*. De maneira geral, a análise de *cluster* segue os seguintes passos: seleção de variáveis; transformação de variáveis; seleção da medida de semelhança/dessemelhança; escolha do método de formação de *clusters*; e discussão e apresentação dos resultados (QUINTAL, 2006). Desta forma, será utilizada a criação de *clusters* para classificar as infraestruturas de C, T&I em grupos similares internamente.

A vantagem dessa análise é a não necessidade de informação *a priori* dos critérios de inclusão nos grupos. Suas técnicas são amplamente usadas em diversos campos do conhecimento, tais como: medicina, sociologia, engenharia e economia. Existem duas abordagens comuns de agrupamento das observações: hierárquico e não-hierárquico. O primeiro, inicialmente, determina n grupos (um para cada observação) culminando com um

único. Já o segundo, utiliza um método de classificação ótima de um *centroide* inicial e, então, as observações são realocadas em K *cluster* mutuamente exclusivos.

A semelhança mede o grau de proximidade entre as variáveis, por conseguinte a dessemelhança mede o grau de diferença entre elas. A medida de similaridade convenientemente usada na literatura é a distância Euclidiana entre duas observações, ou seja, consiste na minimização das distâncias entre cada par de observações em relação ao *centroide* determinado inicialmente. Dada um conjunto de variáveis, a semelhança entre i e j é dada por s_{ij} que satisfaz as seguintes propriedades:

1. $0 \leq s_{ij} \leq 1, \forall i, j$: os valores entre 0 e 1 indicam os graus de semelhança, sendo $s_{ij} = 0$ significando máxima dessemelhança e $s_{ij} = 1$ máxima semelhança;
2. $s_{ij} = s_{ji}, \forall i, j \in I$ (*Simetria*); e
3. $s_{ij} = 1 \forall i \in I$ (*Identidade*).

4.1. Construção da base de dados

A fim de tipificar as infraestruturas em *cluster*, utilizou-se todas as informações presentes no banco de dados do MCTI, exceto às pertencentes às áreas de Ciências Sociais Aplicadas e Ciências Humanas. Desta forma, totaliza-se 1.756 infraestruturas e 7.694 pesquisadores distintos. A relação das variáveis utilizadas na análise de *cluster* está exposta no Quadro 1:

Insumos:

* Avaliação do Capital Humano:

- 1 > Percentual de Pesquisadores de dedicação exclusiva (mais de 30 horas semanais);
- 2 > Percentual de Pesquisadores com doutorado;
- 3 > Percentual de Pesquisadores com mestrado;
- 4 > Média de Estudantes de doutorado;
- 5 > Percentual de Pesquisadores Visitantes;

* Avaliação do Capital Físico:

- 6 > Log da Relação Capital-Trabalho: logaritmo da média do valor estimado da infraestrutura dividido pela quantidade de pessoal ocupado.

Intervenientes:

- > Grande área do conhecimento;
- > Dummy de realização ou não de cooperação.

Tipos de Produto:

- > Tipos de atividades desenvolvidas:

- 10. Ensino;
- 11. Pesquisa;
- 12. Desenvolvimento de Tecnologias;
- 13. Extensão Tecnológica;
- 14. Serviço Tecnológico; e
- 15. Outras.

Quadro 1 – Conjunto de variáveis de determinação dos *clusters*.

Fonte: Elaboração própria.

No entanto, com o intuito de padronizar a formatação das variáveis para estimações dos *clusters*, foram necessárias algumas modificações em seu conjunto inicial. Primeiramente, as variáveis foram separadas de acordo com a estrutura de uma Função de Produção, ou seja, em variáveis de insumo (capital físico e humano), intervenientes e variáveis de resultados (tipos de produtos).

A manipulação nas variáveis de insumos de capital humano ocorreu por meio da transformação em percentual e média das originais. Desta forma, a Tabela 24 descreve como está distribuído o capital humano, sendo as primeiras cinco variáveis representando o percentual médio dos pesquisadores presentes nas 1.756 infraestruturas. Então, pode-se notar que as infraestruturas, em média, contam com mais da metade da sua mão de obra trabalhando mais de 30 horas semanais e 92% com a titulação de doutorado ou mestrado. Além disso, a média de estudantes de doutorado por infraestrutura é de 3,49 e o percentual de pesquisadores estrangeiros e visitantes é 6% e 2%, respectivamente, por infraestrutura.

Tabela 24 – Descrição média das variáveis de Capital Humano presente por infraestrutura.

No. Variável	Variáveis	Média
1	Percentual de pesquisadores com Dedicação Exclusiva	53%
2	Percentual de pesquisadores com Doutorado	80%
3	Percentual de pesquisadores com Mestrado	12%
4	Percentual de pesquisadores Visitantes	2%
5	Percentual de pesquisadores Estrangeiros	6%
6	Média de estudantes com Doutorado	3,49

Fonte: Elaboração própria.

Quanto à variável insumo de capital físico, criou-se uma relação capital-trabalho por infraestrutura. Dado que a variável de valor do capital físico encontrava-se no questionário em

faixas de valores estimados da infraestrutura, primeiramente tirou-se a média de cada categoria, depois dividiu-se pelo total de pessoal ocupado (PO). Como total de pessoal ocupado inclui-se: coordenadores, pesquisadores e equipe de apoio (doutores, mestres, especialistas, graduados, nível médio, servidores e prestadores de serviços terceirizados).

O segundo grupo de variáveis, denominadas intervenientes, dizem respeito à realização ou não de cooperação por elas e às áreas do conhecimento da infraestrutura. A primeira variável será mantida tal qual a original, enquanto que a nova configuração de grandes áreas é apresentada na Tabela 25.

Tabela 25 – Descrição da variável grande área do conhecimento usada na criação dos *clusters*.

Grande Área	No. De Infraestruturas	%
Ciências Exatas e da Terra	363	20,67
Ciências Biológicas	312	17,77
Engenharias	485	27,62
Ciências da Saúde	67	3,82
Ciências Agrárias	206	11,73
Engenharias + Ciências Exatas e da Terra	115	6,55
Engenharias + Ciências Biológicas	14	0,80
Ciências Biológicas + Ciências da Saúde	41	2,33
Ciências Biológicas + Ciências Agrárias	38	2,16
Multiáreas	115	6,55
Total	1.756	100.00

Fonte: Elaboração própria.

Como pode ser observado, além dos 5 principais campos do conhecimento, optou-se por desmembrar do item Multiárea, as com maiores representatividades, tais como a combinação de Engenharias com Ciências Exatas e da Terra e as interações de Ciências Biológicas com Engenharias, Ciências da Saúde e Ciências Agrárias, totalizando 6,55% e 5,29% das infraestruturas respectivamente.

Por fim, o último grupo de variáveis trata dos tipos de produtos das infraestruturas, os quais são caracterizados como as atividades que desenvolvem (Tabela 26). Essas variáveis foram redefinidas com apenas 4 categóricas, agora variando de 0 a 3 conforme o grau de realização de cada uma. Sendo 0 não desenvolve; 1 contínuo; 2 alguns dias da semana e alguns dias do mês; e 3 esporádico.

Tabela 26 – Descrição da distribuição das variáveis de atividades, por intensidade de realização.

Atividades/Categorias	0	1	2	3	Total
Ensino	14,24%	39,35%	31,38%	15,03%	100,00%
Pesquisa	4,50%	80,69%	12,02%	2,79%	100,00%
Desenv. de Tecnologia	31,78%	35,31%	15,83%	17,08%	100,00%
Extensão Tecnológica	54,90%	7,86%	10,54%	26,71%	100,00%
Serviço Tecnológico	41,63%	17,94%	13,15%	27,28%	100,00%

Fonte: Elaboração própria.

Na Tabela 26, observa-se que a maioria das infraestruturas executam as atividades de ensino com intensidades que variam entre 1 (contínuo) e 2 (alguns dias da semana e alguns dias do mês), já o desenvolvimento de tecnologia está dividido entre as categorias 1 (contínuo) e 0 (não desenvolve), enquanto que as atividades de pesquisa concentram-se, preponderantemente, na categoria *contínuo*. Por outro lado, as atividades de extensão tecnológica e serviço tecnológico não são realizadas pela maioria das infraestruturas (54,9% e 41,63% respectivamente). Desta forma, deve-se esperar que os *clusters* formados apresentassem diferenças na realização das três primeiras atividades da Tabela e, conseqüentemente, não haja dessemelhanças na execução das duas últimas.

4.2. Cluster Não-Hierárquico

A realização da construção de um *cluster* não-hierárquico tem como objetivo encontrar grupos de infraestruturas que possuem características semelhantes. As principais vantagens desse método são: bom desempenho em grande quantidade de observações e capacidade de reagrupar as observações em *clusters* diferentes daqueles em que foram alocados inicialmente. Sua desvantagem está na necessidade de definir o número de grupos *a priori*. Os passos para a construção de um *cluster* não-hierárquico são os seguintes:

1. Determina-se o número inicial de grupos das n observações em k *cluster*;
2. Calculam-se os centroides para cada um dos k *clusters*;
3. Agrupa-se as observações aos *clusters* cujos centroides se encontram mais próximos, depois volta-se ao passo (2) até não ocorrer variações significativas na distância mínima de cada observação.

Neste trabalho optou-se por utilizar o método *K-medians*¹⁴, já que proporciona uma maior estabilidade no tratamento de valores extremos presentes na amostra. Tal metodologia usa a mediana de todos os pontos do conjunto de observações para determinar o *centroide* de cada grupo. Além disso, a distância será minimizada por meio dos quadrados das distâncias Euclidianas (L2squared).

Para definir quantos grupos serão escolhidos, realizou-se uma análise de sensibilidade a fim de testar a validade dos *clusters* gerados pelos diversos processos de inicialização de escolha de *K*¹⁵. Os grupos formados estão descritos na Tabela 27.

Tabela 27 – Divisão das infraestruturas de acordo com o número de grupos e processos de inicialização de escolha de K.

No. De Clusters	random K (rk)	last K (lk)	first K (fk)	every Kth (eKth)	segments K (sk)
1	1.433	323	433	932	824
2	323	1.433	1.323	824	932
Total	1.756	1.756	1.756	1.756	1.756
1	533	1.058	275	568	791
2	568	277	605	1.012	608
3	655	421	876	176	357
Total	1.756	1.756	1.756	1.756	1.756
1	274	394	197	348	516
2	353	1.039	584	516	417
3	439	143	840	405	431
4	690	180	135	487	392
Total	1.756	1.756	1.756	1.756	1.756
1	110	539	90	409	454
2	178	249	436	467	239
3	557	680	690	436	284
4	528	110	194	332	324
5	383	178	343	112	455
Total	1.756	1.756	1.756	1.756	1.756

Fonte: Elaboração própria.

O processo de validação recomendado pela literatura determina que se defina uma divisão aleatória dos dados em dois subgrupos, aqui denominados de A e B. Após tal divisão, deve-se aplicar o mesmo método de *cluster* em cada um desses subgrupos. Dos resultados gerados, é conferida validação se os *clusters* formados em cada subgrupo são similares. Desta maneira, no presente trabalho optou-se por aleatorizar os dados por meio de uma distribuição

¹⁴ Comando disponível no *software* Stata.

¹⁵ Entre os métodos disponíveis no *software* Stata estão: *krandom*, *firstk*, *lastk*, *everykth*, *segmentsK*

uniforme e as observações foram divididas entre aquelas que se encontravam acima da média – grupo A – e as abaixo da média – grupo B.

Desta forma, foram realizados os processos de validação com todos os grupos apresentados na Tabela 27, conforme pode ser visto no Apêndice A. Das análises de sensibilidade executadas entre os *clusters* com as observações iniciais e as subdivididas, percebe-se que o *cluster* com três grupos (K=3) utilizando o método de inicialização *firstK* atingiu a melhor correspondência. Assim, das 1.756 infraestruturas, os grupo A e B agruparam 878 infraestruturas cada.

	fk3A					fk3B			
fk3	1	2	3	TOTAL	fk3	1	2	3	TOTAL
1	144			144	1	131			131
2		289		289	2	2	245	69	316
3			445	445	3			431	431
TOTAL	144	289	445	878	TOTAL	133	245	500	878

Quadro 2: Correspondência entre o *cluster* global e cada subgrupo.

Fonte: Elaboração própria.

Como pode ser percebidos na parte direita do Quadro 2, todo o subgrupo A está 100% dentro dos *clusters* do fk3. Na parte esquerda, o grupo 2 do subgrupo B é construído completamente pelas observações do *cluster* 2 do fk3, enquanto que os grupos 1 e 3 possuem, respectivamente, 98,5% e 86,2% dos grupos 1 e 3 do fk3. Desta forma, pode-se concluir que o *cluster* oriundo do método de inicialização *firstK* com K=3 possui a divisão mais similar das infraestruturas quando comparado com o *cluster* global.

Diante disso, foi realizado um teste MANOVA (*Multivariate analysis-of-variance*) como um teste de diferenças de médias entre as variáveis utilizadas para a formação do *cluster*. Na análise do Quadro 3, vê-se que foram utilizadas as 1.756 infraestruturas e o resultado indica que os testes multivariados rejeitam a hipótese nula, o que indica algum tipo de diferença de média entre todas as variáveis utilizadas na formação do *cluster* fk3. Desta forma, pode-se concluir que o mesmo realizou uma boa divisão das infraestruturas, dado que esses grupos gerados são diferentes entre si.

Number of obs = 1756										
W = Wilks' lambda					L = Lawley-Hotelling trace					
P = Pillai's trace					R = Roy's largest root					
Source		Statistic		df	F(df1,	df2) =	F	Prob>F		
fk3	W	0.1594		2	30.0	3478.0	174.48	0.0000	e	
	P	1.1391			30.0	3480.0	153.47	0.0000	a	
	L	3.4024			30.0	3476.0	195.11	0.0000	a	
	R	2.7119			15.0	1740.0	314.58	0.0000	u	
Residual				1753						
Total				1755						
e = exact, a = approximate, u = upper bound on F										

Quadro 3: Resultado do teste MANOVA (*Multivariate analysis-of-variance*) aplicado sobre o *cluster* fk3 gerado.

Fonte: Elaboração própria, a partir do *Software* Stata.

4.3. Resultados por tipologia das Infraestruturas

Os resultados gerados por meio dos centroides de cada variável utilizada no método de *cluster* estão descritos na Tabela 28. Aliado a isso, tem-se o teste *mvreg*, exposto no Apêndice B, que apresenta como se comportam as diferenças de médias das variáveis entre os *clusters*.

Tabela 28 – Centroides das variáveis, por grupos do *cluster* fk3.

Variável/Clusters		fk3=1	fk3=2	fk3=3
1.	Perc. de Dedicção Exclusiva	58,68%	57,70%	48,10%
2.	Perc. de pessoal com Doutorado	80,50%	91,00%	71,80%
3.	Perc. de pessoal com Mestrado	13,30%	6,23%	16,10%
4.	Média de estudantes de Doutorado	2,26	8	0,73
5.	Perc. de pesq. visitantes	3,55%	2,78%	1,70%
6.	Perc. de pesq. estrangeiros	6,46%	7,00%	4,54%
7.	Log do Capital-Trabalho	9,22	9,11	93,40%
8.	Principais áreas do Conhecimento	Multiárea	Engenharia e Ciências Biológicas	Engenharia e Ciências Exatas e da Terra

(Continua.)

(Conclusão.)

Variável/Clusters		fk3=1	fk3=2	fk3=3
9.	Realiza cooperação	93,45%	97,52%	81,05%
10.	Atividade de Ensino (<i>contínuo</i>)	35,27%	56,03%	29,11%
11.	Atividade de Pesquisa (<i>contínuo</i>)	85,82%	96,36%	68,26%
12.	Atividade de Desenv. De Tecnologia (<i>contínuo</i>)	42,55%	46,78%	25,11%
13.	Atividade de Extensão Tecnológica (<i>contínuo</i>)	12,00%	9,92%	5,14%
14.	Atividade de Serviço Tecnológico (<i>contínuo</i>)	23,27%	17,02%	16,89%
15.	Outras atividades (<i>contínuo</i>)	4,73%	2,48%	2,28%

Fonte: Elaboração própria.

Analizando os três *clusters não-hierárquicos* pelos centroides das variáveis e tomando o *cluster* 1 como referência no teste de diferenças de médias (Apêndice B) pode-se esboçar uma caracterização dos *clusters* por meio dos subconjuntos de variáveis formados no *cluster hierárquico* (Apêndice C). No primeiro grupo de variáveis (2, 4, 7 e 9), o percentual de pessoal com doutorado (variável 2) e média de estudantes de doutorado (variável 4) obtiveram associações lineares positivas e estatísticas significativas entre os *clusters*, servindo de grande utilidade na categorização. Já a variável 7 (log do capital-trabalho) não apresentou diferenças entre os *clusters*, apenas quando associada com a variável 9 (realização de cooperação) pode-se perceber uma relação linear negativa.

O subconjunto de variáveis formado pelos produtos das infraestruturas (atividades) geraram diferenças de médias entre os *clusters* apenas nas atividades de ensino e pesquisa, porém dentro de cada um percebe-se que as infraestruturas possuem distintas intensidades de realizações das atividades, o que deve ajudar a caracterizá-los. E como esperado, as variáveis 13 e 14 (Extensão e Serviço tecnológico) obtiveram o mesmo comportamento e não exerceram influência na categorização dos *clusters*. No subconjunto de variáveis de cooperação externa, as variáveis de percentual de pesquisadores estrangeiros (6) e visitante (5) apresentaram diferenças estatísticas entre o *cluster* 3 em comparação ao 1 e seus movimentos internos possuem associação negativa, tal fato, também, ocorre na variável 1 (percentual de dedicação exclusiva). Por último, a variável 8 é estatisticamente significativa entre todos os *clusters*, desta forma auxiliando na definição das grandes áreas do conhecimento de atuação do conhecimento de cada *cluster*.

Com isso, a seguir é proposta uma caracterização de cada grupo formado, tanto com as variáveis utilizadas para a formação dos *clusters* quanto às demais que compõem o banco de dados, as quais auxiliarão na análise das semelhanças e diferenças entre cada um deles.

- **Grupo 1 (*laboratórios intensivos em pesquisa e desenvolvimento de tecnologia*):** em comparação, pode ser caracterizado como um grupo intermediário no percentual de pesquisadores com doutorado e mestrado (80,5% e 13,3%, respectivamente) e na média da quantidade de estudantes de doutorado (2,26). Porém, possui o maior percentual de pesquisadores visitantes (3,55%). Atuam, em maioria, na categoria Multiárea do conhecimento e, entre as atividades realizadas de forma *contínua*, a pesquisa e o desenvolvimento de tecnologias são as principais, seguidas pelas de ensino, serviço tecnológico e extensão tecnológica. Estão presentes em universidades públicas/privada (63,64%) e em laboratórios públicos (30,18%), além de possuírem o maior tamanho em área física, 402,6 m².
- **Grupo 2 (*laboratórios multiatividades*):** pode ser caracterizado por possuir o maior percentual de pesquisadores com doutorado e pesquisadores estrangeiros (91% e 7%, respectivamente), além disso possui a maior média da presença de estudantes de doutorado (8,0). Atuam, principalmente, nas áreas de Engenharia (27,44%) e Ciências Biológicas (26,61%). As atividades de pesquisa, ensino e desenvolvimento de tecnologia são as principais e estão presentes, na grande maioria, nas universidades públicas/privada (85,45%). As atividades de serviço tecnológico e extensão tecnológica são praticadas continuamente por menos da metade das infraestruturas. São laboratórios mais antigos – média de 22 anos – e possuem uma estrutura organizacional piramidal, ou seja, poucos coordenadores em relação à quantidade de pesquisadores.
- **Grupo 3 (*laboratórios intensivos em pesquisa*):** pode ser caracterizado por possuir o menor percentual de pesquisadores com dedicação exclusiva e com doutorado (48,1% e 71,8%, respectivamente) e a média mais baixa de estudantes de doutorado (0,73). As principais áreas do conhecimento que atuam são Engenharia (36,33%) e Ciências Exatas e da Terra (27,11%). A atividade de pesquisa é realizada de forma *contínua* pela maior parte das infraestruturas, enquanto que a atividade de ensino e desenvolvimento de tecnologia aparecem em segundo plano. Estão presentes em universidades públicas/privada (57,88%), laboratórios públicos (25,91%) e instituições de ensino técnico (14,5%). Possui o menor tamanho em área física, 316,4m² e a maior relação capital-trabalho. A estrutura organizacional é mais horizontal, ou seja, quase metade dos pesquisadores possui o cargo de coordenadores.

Das variáveis presentes no banco de dados das infraestruturas que não promoveram distinção significativa, ou seja, diferenças de médias pode-se definir o seguinte perfil médio: maioria são estruturas de laboratório; presentes na região Sudeste do Brasil (São Paulo, Rio de Janeiro e Minas Gerais); considerando *regular* a situação de suas instalações físicas e *boa* a situação de seus equipamentos. Ademais, seus financiamentos são oriundos, em maior parcela, de fontes externas.

Ainda no banco de dados do MCTI, pode-se dividir por tipologia de infraestrutura como ocorre a participação da mão de obra segundo as suas características de execução do trabalho (Tabela 29).

Tabela 29 – Discriminação das características de ocupação dos pesquisadores, de acordo com as tipologias de *cluster não-hierárquico*.

	Grupo 1	Grupo 2	Grupo 3
Trabalham somente em 1 laboratório	18,81%	45,19%	36,00%
Trabalha mais de 30 horas semanais	16,97%	53,73%	29,30%
Percentual de coordenadores	61,84%	58,54%	69,33%

Fonte: Elaboração própria.

Como resultado, tem-se que o grupo 2 possui o maior percentual de pesquisadores trabalhando em um único laboratório e mais de 30 horas semanais. Cenário oposto acontece no grupo 1, enquanto que o grupo 3 se apresenta como um *cluster* intermediário nesses quesitos. Quanto ao cargo de coordenador com título de doutorado até 2009 e empregado com servidor público, celetista ou outros, os maiores percentuais em relação à quantidade de mão de obra total por infraestrutura ocorre no grupo 3, seguindo pelo grupo 1 e 2.

5 METODOLOGIA ECONOMÉTRICA

Segundo Bonaccorsi e Daraio (2004), a abordagem de Função de Produção na medição da produtividade científica também pode ser expressa por meio de uma relação funcional que cruza dados observados de *inputs* e *outputs*. Segundo Stephan (2012), os recursos compreendidos como *inputs* do processo de produção do conhecimento são compostos por: fatores humanos constituídos pelas características individuais dos pesquisadores e da equipe de trabalho em que está inserido; fatores de capital físico composto pelo valor dos equipamentos; e características gerais das infraestruturas. Todos estes viabilizando a realização das pesquisas e, em diversos casos, ditando o avanço do conhecimento. Já os recursos gerados pelo sistema – *outputs* – são considerados, por exemplo, como a quantidade de artigos publicados em um determinado período.

Desta maneira, neste capítulo serão discutidas a especificação para a Função de Produção do conhecimento e os procedimentos econométricos adotados, a fim de explicar a produtividade científica dos pesquisadores pertencentes aos centros de pesquisas do Brasil por meio de dados em *cross-section*. Com isso, será avaliado a contribuição das características individuais e coletivas sobre a produtividade por meio da publicação de artigos desses pesquisadores brasileiros.

5.1. Modelo com Dados em *Cross-Section* para Determinantes da Produtividade Individual

A Função de Produção (FP) expressa uma relação tecnológica existente entre os fatores empregados no processo de produção (*inputs*) e a quantidade de produtos decorrentes de tal processo (*outputs*). Portanto, a análise da FP começa pela construção de uma relação funcional, que aqui assume-se ser do tipo Cobb-Douglas:

$$Y = A K^{\alpha} L^{\beta} \quad (1)$$

Em que Y são os *outputs* do processo produtivo; A são os *inputs fixos*, incluindo o efeito do progresso tecnológico; e K e L representam os fatores de produção – *capital* e *trabalho*; sendo A , α e β positivos.

Tomando como base a divisão da mão de obra dentro das infraestruturas de pesquisa realizada, como, por exemplo, em Carayol e Matt (2004). Neste trabalho separou-se o pessoal total ocupado (PO) da infraestrutura nas 3 categorias propostas, de acordo com o grau de escolaridade e vínculos empregatícios:

1. Pesquisadores Permanentes (L_1): pesquisadores com pós-graduação e graduação, possuem vínculo empregatício como servidor público, celetista ou outros, adicionado aos pesquisadores que exercem a função de coordenador;
2. Pesquisadores Não-Permanentes (L_2): pesquisadores que recebem bolsas de estudo, independente da titulação e não exercem a função de coordenador;
3. Não-Pesquisadores (L_3): mão de obra com ensino médio ou inferior, possuem vínculo empregatício como servidor público, celetista ou outros e não exercem a função de coordenador. Somado à equipe denominada apoio administrativo independente da titulação.

Portanto, o *input* trabalho da função de produção pode ser descrito como a interação entre os 3 tipos de mão de obra, tal como segue:

$$L = L_1 + L_2 + L_3 \quad (2)$$

Desta maneira, substituindo (2) na equação (1), tem-se que a equação da Função de Produção do conhecimento pode ser escrito da seguinte forma:

$$Y_{in} = A_{in} K^\alpha L_1^{\beta_1} L_2^{\beta_2} L_3^{\beta_3} \exp(u_i) \quad (3)$$

Em termos operacionais, a função de produção Cobb-Douglas possui o atrativo no processo de estimação ao ter a propriedade de se tornar linear quando reescrita em termos do ajuste logarítmico de suas variáveis (Baumol, 1977). Sendo assim, logaritizando a equação (3) tem-se:

$$\ln y = \ln A_{in} + \alpha \ln K + \beta_1 \ln L_1 + \beta_2 \ln L_2 + \beta_3 \ln L_3 + u_i \quad (4)$$

A literatura argumenta que o coordenador da infraestrutura é o detentor da liderança acadêmica e intelectual no ambiente de produção de pesquisa, ou seja, sua participação é de extrema importância na transmissão dos conhecimentos para os demais membros. Ademais, o pesquisador líder é responsável pela coordenação e planejamento dos trabalhos de pesquisa do grupo, aglutinando os esforços dos demais pesquisadores e apontando novos horizontes e áreas de atuação dos trabalhos (CNPq, 2015). Desta maneira, a especificação econométrica do presente trabalho terá como objetivo determinar a função de produção científica do pesquisador que exerce a função de chefia da infraestrutura (coordenador) com vínculo empregatício de servidor público, celetista ou outros e com título de doutorado obtido até 2009. Dado isso, serão selecionadas apenas as infraestruturas que tiveram ano de início de operação até 2009, a qual totalizou 1.652 cargos de coordenação, sendo que engloba 1.445 coordenadores distintos.

A quantidade de artigos publicados no período de 2009 a 2013, do coordenador i pertencente à infraestrutura do tipo n , será utilizada como medida de produtividade científica, porém este montante pode incluir dupla de contagem. Com isso, dividiu-se cada ocorrência pelo número de coautores, a fim de obter a contribuição científica efetiva de cada autor. Outra forma de construção é ponderar a produtividade científica pelo Fator de Impacto (FI)¹⁶ das revistas indexadas pelo *Institute for Scientific Information* (ISI).

Para isso, utilizou-se as informações qualitativas contidas na base de dados de tal instituição que, em 2013, continha informações de 11.519 periódicos internacionais, sendo 107 provenientes do Brasil. O uso das informações do ISI é amplamente difundido pelos trabalhos em áreas de bibliometria e cienciometria. Contudo, algumas ressalvas são necessárias. Um dos aspectos negativos ocorre quando os fatores de impacto das revistas de diferentes áreas do conhecimento são comparados. Tal avaliação não deve ser realizada, pois há diferenças nas densidades das citações (número de referências citadas por artigo) nos periódicos e no ritmo de obsolescência entre as áreas, favorecendo as revistas onde a renovação do conhecimento é mais rápida. Além disso, pode haver correlação de um alto FI com um grande número de citações em um único artigo de revista ou com uma pequena disponibilidade de periódicos em uma área científica. Outro problema são as citações de

¹⁶ O Fator de Impacto é uma medida, calculado pelo ISI, que reflete a importância de um dado periódico em sua área do conhecimento, sendo aqueles com maiores valores considerados mais importantes do que aqueles com menores valores. Em termos matemáticos, o FI do ano corrente é obtido pela razão entre o número de citações feitas a trabalhos publicados nos últimos dois anos e o número total de artigos publicados pelo periódico nos mesmos dois anos (JCR, 2015).

artigos não pela sua qualidade, mas por ser contestado devido a algum erro nos métodos empregados ou na interpretação dos resultados experimentais.

Por esses motivos, há uma necessidade de equalizar os fatores de impacto do ISI para utilizar como instrumento de avaliação da qualidade da produtividade científica. Sombatsompop e Markpin (2005) propôs o *Impact Factor Point Average* (IFPA), o qual busca normalizar as diferenças entre os fatores de impacto e da área do conhecimento ponderado pela relação entre o *ranking* do periódico e o número de revistas existentes em cada diferente área do conhecimento. De tal forma como segue:

$$IFPA = \left[\frac{I_j}{I_A} \right] \cdot \left[1 - \frac{R}{N+1} \right] \cdot [n] \quad (5)$$

Onde I_j e I_A correspondem ao fator de impacto do periódico e ao fator de impacto agregado médio de todos os periódicos da mesma categoria científica¹⁷, respectivamente; R é o *ranking* do periódico em sua categoria; N relaciona o número de periódicos existentes na mesma categoria; e, n compreende o número de artigos publicados pelo pesquisador.

Ainda segundo Sombatsompop e Markpin (2005), a relação $\left[\frac{I_j}{I_A} \right]$ normaliza os diferentes fatores de impacto dos periódicos em cada categoria do conhecimento, isto é, quanto maior o I_A , menor será o IFPA. Já o termo $\left[1 - \frac{R}{N+1} \right]$ normaliza os diferentes números de revistas em cada categoria do conhecimento, portanto quanto maior o R e menor o N , menor será o IFPA; e $[n]$ indica a qualidade específica dos pesquisadores por meio da quantidade de artigos publicados nos periódicos.

Escrevendo os *inputs* de capital e trabalho da Função de Produção do conhecimento por pessoal total ocupado (PO) na infraestrutura, a equação (4) será reescrita como:

$$\ln y_{in} = \ln A_{in} + \alpha \ln K + \beta_1 \ln L_1 + \beta_2 \ln L_2 + \beta_3 \ln L_3 + \ln L + u_i$$

Se $\ln L = (\alpha + \beta_1 + \beta_2 + \beta_3 + (1 - \alpha - \beta_1 - \beta_2 - \beta_3)) \ln L$, então:

$$\ln y_{in} = \ln A_{in} + \alpha \ln \frac{K}{L} + \beta_1 \ln \frac{L_1}{L} + \beta_2 \ln \frac{L_2}{L} + \beta_3 \ln \frac{L_3}{L} + (\alpha + \beta_1 + \beta_2 + \beta_3) \ln L + u_{1i} \quad (6)$$

¹⁷ Todo periódico deve pertencer a uma categoria de área do conhecimento, podendo estar incluída em até outras três categorias diferentes a depender da abrangência de campos acadêmicos a que faz parte.

Mais detalhadamente, o parâmetro A significa a produtividade total dos fatores, ou seja, indica a eficiência estimada do processo. Aqui, tal variável pode representar as características do ambiente institucional da infraestrutura em que o pesquisador está inserido. De tal forma que pode ser definido como:

$$A_{in} = c \cdot idade^{\gamma_1} \cdot idade^{2\gamma_2} \cdot D_n \cdot y_{L_1} \quad (7)$$

Com base na equação (7), a produtividade total dos fatores pode ser caracterizada por variáveis invariantes e variantes no tempo. Sendo tais variáveis construídas da seguinte maneira:

- Produção dos pesquisadores permanentes: log da quantidade de artigos publicados no período de 2009 a 2013, dos pesquisadores permanentes, dividida pelo número de coautores, menos a produção dos coordenadores.

Assim, busca-se captar as externalidades geradas pelo desempenho dos pesquisadores permanentes sobre a produtividade científica do coordenador;

- Cluster: dummies D_n que são correspondentes das infraestruturas de pesquisa do tipo n , definidas segundo análise de *cluster* não-hierárquico, as quais assumem valor 1 quando *cluster* do tipo 2.

Desta forma, pode-se inferir a importância do conjunto de características formado pelas infraestruturas pertencentes aos *clusters* na produtividade científica do coordenador;

- Idade do laboratório: log da diferença entre ano de criação da infraestrutura menos o ano de 2013;
- Idade do laboratório ao quadrado: log do quadrado da diferença entre ano de criação da infraestrutura menos o ano de 2013;
- Dummy de pesquisador visitante: *dummy* D_v , a qual assume valor 1 ao identificar a presença de pesquisadores visitantes na infraestrutura e 0 caso contrário.

Quanto à relação capital-trabalho da infraestrutura, $\frac{K}{L}$, foi construída pelo log da média de cada faixa de valor da infraestrutura dividida pelo log do PO. Desta forma, busca-se captar como a interação entre o valor dos equipamentos/materiais e mão de obra pode explicar os diferentes níveis produtividade dos pesquisadores.

Em relação às variáveis dos fatores de qualificação do trabalho da infraestrutura foram construídas da seguinte maneira:

- Percentual de Pesquisadores Permanentes, $\frac{L_1}{L}$: log do percentual dos pesquisadores permanentes total dividido pelo PO por infraestrutura;
- Percentual de Pesquisadores Não-Permanentes, $\frac{L_2}{L}$: log do percentual dos pesquisadores não-permanentes total dividido pelo PO por infraestrutura;
- Percentual de Não-Pesquisadores, $\frac{L_3}{L}$: log do percentual dos não-pesquisadores dividido PO por infraestrutura.

E, u_i identifica os fatores não observáveis do pesquisador i , n e segue uma distribuição normal bivariada $\sim (0, \rho)$.

Para tentar minimizar o problema de viés de variável omitida, dado que não se pode negligenciar a característica de trajetória dependente que a variável de interesse (produção científica) tende a apresentar, torna-se importante a inclusão da variável dependente defasada em um período. Tal variável foi construída pela soma da quantidade total de artigos publicados no período de 2004 a 2008 pelo coordenador i da infraestrutura n divididos pelo número de coautores. Portanto, será incluída na equação (6) da seguinte maneira:

$$\ln y_{in} = \ln A + \alpha \ln \frac{K}{L} + \beta_1 \ln \frac{L_1}{L} + \beta_2 \ln \frac{L_2}{L} + \beta_3 \ln \frac{L_3}{L} + \theta_1 \ln L + \theta_2 \ln y_{in-1} + u_{2i} \quad (8)$$

Por fim, incluindo variáveis de características individuais como controles na equação (8), a especificação a ser estimada é:

$$\ln y_{in} = \ln A + \alpha \ln \frac{K}{L} + \beta_1 \ln \frac{L_1}{L} + \beta_2 \ln \frac{L_2}{L} + \beta_3 \ln \frac{L_3}{L} + \theta_1 \ln L + \theta_2 \ln y_{in-1} + \theta_3 z + u_{3i} \quad (9)$$

Sendo os instrumentos de controles da equação econométrica (9), pertencentes ao vetor z , construídos da seguinte maneira:

- Geração: variável idade de pesquisa representa coortes no tempo que foram construídas a partir da diferença entre: o ano de 2013 e o ano de obtenção do título de doutorado de cada coordenador. Posteriormente, agrupou-se os resultados em 4 *dummies*: 34 – 51 anos de doutorado; 24 – 33 anos de doutorado; 14 – 23 anos de doutorado; e 04 – 13 anos de doutorado (grupo de referência).

Assim, tenta-se captar os impactos do modelo de Ciclo de Vida descrito pela literatura por meio dos efeitos de geração;

- Área de Pesquisa: variáveis *dummies* de áreas de pesquisa do coordenador, que são: Ciências Exatas e da Terra; Ciências Biológicas; Engenharias; Ciências da Saúde; Ciências Agrárias; e Multiárea (grupo de referência).

Com isso, tenta-se obter os efeitos de variação de publicação entre as grandes áreas do conhecimento do cientista; e

- Número de Infraestruturas que trabalha: variável *dummy* que determina em quantas infraestruturas o cientista realiza atividades, a qual assume valor 1 quando o coordenador trabalhar em mais de uma infraestrutura.

5.1.1 Tobit

Em amostra com censura, para valores maiores que zero, existem informações completas – tanto das variáveis dependentes quanto das independentes – e para valores censurados as informações são incompletas, observando-se apenas as variáveis independentes. Dado isso, tem-se que estimadores por Mínimos Quadrados Ordinários podem gerar resultados não consistentes, portanto, o modelo também será estimado por meio do método Tobit. Tal método foi proposto originalmente por Tobin (1956) e fornece estimativas consistentes ao permitir acomodar situações em que há acumulação de observações num único ponto de massa na fronteira do domínio. Portanto, é uma estratégia de estimação para amostras com presença de zeros ao tratar o problema de censura na variável dependente.

Como a medida de produtividade do coordenador é dado pela quantidade de artigos publicados, há indivíduos que produziram zero no período de análise. Nesse caso, o $\ln y_{in}$ não

pode ser definida para esse subconjunto da população. Para resolver tal problema, somou-se 0,5 às variáveis dependentes, a fim de defini-las para todos os coordenadores da amostra.

O modelo Tobit do presente trabalho, segundo Wooldridge (2002), é o seguinte:

$$\ln y_{in} = X\beta + u \quad (10)$$

e

$$\ln y_{in} = \max(0, \ln y_{in}^*) \quad (11)$$

Sendo o termo de erro $u|X \sim N(0, \sigma^2)$, X vetor de variáveis explicativas e β vetor de parâmetros.

A esperança condicional de $\ln y_{in}$ é não negativa e independente da distribuição e forma funcional de $\ln y_{in}^*$. Além disso, se u for não dependente de X e normalmente distribuído, pode-se definir a esperança condicional de $\ln y_{in}$ explicitamente:

$$\ln y_{in} = E(\ln y_{in} | X, \ln y_{in} > 0) + u_{in} \quad (12)$$

Se x_j é contínuo e não relacionado a outros regressores, pode-se definir o Efeito Marginal Condicional de $\ln y_{in} > 0$ da seguinte maneira:

$$\begin{aligned} \frac{\partial E(\ln y_{in} | X, \ln y_{in} > 0)}{\partial x_j} &= \beta_j \left\{ 1 - \lambda \left(\frac{x\beta}{\sigma} \right) + \left[\frac{x\beta}{\sigma} + \lambda \left(\frac{x\beta}{\sigma} \right) \right] \right\} \\ \frac{\partial E(\ln y_{in} | X, \ln y_{in} > 0)}{\partial x_j} &= \beta_j \Theta \left(\frac{x\beta}{\sigma} \right) \end{aligned} \quad (13)$$

Onde $0 < \Theta \left(\frac{x\beta}{\sigma} \right) < 1$.

A equação (13) mostra que o efeito marginal de x_j no $E(\ln y_{in} | X, \ln y_{in} > 0)$ não é determinado totalmente apenas por β_j , mas há um fator de ajustamento o multiplicando que depende de x por meio de $x\beta/\sigma$. Para as variáveis binárias, o efeito marginal condicional é obtido pela diferença do valor esperado, $E(\ln y_{in} | X, \ln y_{in} > 0)$, entre $x = 1$ e $x = 0$.

Já para a subamostra censurada, o Efeito Marginal não condicional é:

$$\frac{\partial E(\ln y_{in} | X)}{\partial x_j} = \beta_j \Phi \left(\frac{x\beta}{\sigma} \right) \quad (14)$$

Sendo: $P(\ln Y_{in} > 0 | X) = \Phi \left(\frac{x\beta}{\sigma} \right)$.

Portanto, sendo x_j contínuo e não relacionado a outros regressores, tem-se um efeito marginal não condicional resultante da ponderação de β_j pela probabilidade de se obter uma resposta positiva dado X , $P(\ln Y_{in} > 0 | X)$.

6 RESULTADOS

O objetivo dessa seção é apresentar os resultados obtidos por meio dos exercícios econométricos propostos neste trabalho, sempre comparando os resultados com os estimados por Mínimos Quadrados Ordinários (OLS). Portanto, nas subseções posteriores serão apresentados os resultados dos efeitos marginais da produtividade científica do coordenador com e sem as ponderações pelo *Impact Factor Point Average* (IFPA).

6.1. Estatísticas Descritivas

Primeiramente, é importante apresentar uma estatística descritiva sobre o conjunto das variáveis que medem a produtividade científica dos 1.445 coordenadores distintos e dos 4.954 parceiros de pesquisa, ou seja, os pesquisadores permanentes. A Tabela 30 descreve como ocorrem as performances dos coordenadores e dos demais pesquisadores permanentes durante o quinquênio de análise, por meio de suas estatísticas descritivas.

Tabela 30 – Estatísticas descritivas das variáveis da produtividade científica dos coordenadores e pesquisadores permanentes.

Variável	Qtd. de observações	Quantidade	Média	Desvio Padrão	Mínimo	Máximo
Artigos publicados no quinquênio pelos coordenadores	1.445	24.901	17,23	20,82	0(4,57%)	391
Trabalha em apenas 1 infra	1.337	22.926	17,15	21,15	0	391
Por tempo de obtenção do doutorado:						
34 a 51 anos	51	1.230	24,12	18,70	0	79
24 a 33 anos	207	5.024	24,27	26,51	0	255
14 a 23 anos	588	11.197	19,04	23,21	0	391
04 a 13 anos	599	7.450	12,44	14,10	0	151
Por grande área do conhecimento:						
Ciências Exatas e da Terra	83	1.401	16,88	44,95	0	391
Ciências Biológicas	54	908	16,82	13,74	2	58
Engenharias	81	557	6,88	7,55	0	38
Ciências da Saúde	11	240	21,82	25,40	0	79
Ciências Agrárias	29	529	18,24	13,40	0	52
Multiáreas	1.187	21.266	17,92	18,92	0	255

(Continua).

(Conclusão).

Variável	Qtd. de observações	Quantidade	Média	Desvio Padrão	Mínimo	Máximo
Artigos publicados/ coautores no quinquênio pelos coordenadores	1.445	5.364,38	3,71	4,50	0 (4,57%)	102,90
Trabalha em 1 infra	1.377	4.894,45	3,66	4,55	0	102,90
Por tempo de obtenção do doutorado:						
34 a 51 anos	51	240,30	4,71	3,31	0	12,20
24 a 33 anos	207	1.083,07	5,23	4,83	0	36,38
14 a 23 anos	588	2.470,81	4,20	5,53	0	102,90
04 a 13 anos	599	1.570,2	2,62	2,62	0	31,80
Por grande área do conhecimento:						
Ciências Exatas e da Terra	83	291,36	3,51	11,27	0	102,90
Ciências Biológicas	54	173,87	3,22	2,66	0,29	12,21
Engenharias	81	166,96	2,06	2,21	0	13,62
Ciências da Saúde	11	45,10	4,10	5,19	0	18,32
Ciências Agrárias	29	120,24	4,15	3,32	0	14,21
Multiáreas	1.187	4.566,85	3,85	3,85	0	36,38
Artigos publicados/ coautores no quinquênio pelos coordenadores ponderados pelo IFPA	1.445	2.630,40	1,82	4,03	0 (11,21%)	119,71
Por tempo de obtenção do doutorado:						
34 a 51 anos	51	137,86	2,70	2,47	0	9,42
24 a 33 anos	207	540,51	2,61	3,73	0	32,36
14 a 23 anos	588	1.255,70	2,14	5,55	0	119,71
04 a 13 anos	599	696,33	1,16	1,70	0	18,37
Por grande área do conhecimento:						
Ciências Exatas e da Terra	83	157,19	1,89	3,25	0	23,57
Ciências Biológicas	54	102,04	1,89	2,17	0,039	9,15
Engenharias	81	92,10	1,14	1,66	0	10,29
Ciências da Saúde	11	31,53	2,87	3,05	0	7,73
Ciências Agrárias	29	48,19	1,66	3,07	0	16,50
Multiáreas	1.187	2.199,34	1,85	4,28	0	119,71
Média dos artigos publicados/coautores no quinquênio pelos pesquisadores permanentes	4.954		1,53	2,01	0	14,86
Média dos artigos publicados/coautores no quinquênio pelos pesquisadores permanentes ponderado pelo IFPA	4.954		0,54	1,02	0	8,62

Fonte: Elaboração própria, a partir da Plataforma *Lattes* e ISI.

Ao analisar a quantidade de artigos publicados pelos coordenadores no quinquênio (2009-2013), observa-se uma média de 17,23 artigos, sendo que 4,57% da amostra não publicaram. O total de coordenadores que trabalham em apenas uma infraestrutura é de 1.337 e produziram, em média, 17,15 artigos no somatório dos cinco anos. A geração dos coordenadores com maior média de produção (24,27 artigos no quinquênio) foi a que possui de 24 a 33 anos de doutorado, ou seja, obtiveram o título entre 1980 e 1989, por outro lado, os de menor média foram os mais novos (4 a 13 anos de doutorado). Entre as grandes áreas do conhecimento, os que têm formação na área de Ciências da Saúde obtiveram maiores médias de publicação (21,82 artigos no período), enquanto que os pertencentes às Engenharias obtiveram as menores médias (6,88 artigos no quinquênio).

Em se tratando da quantidade de artigos publicados entre 2009 e 2013 dividida pelo número de coautores, nota-se que a média dos pesquisadores que trabalham em apenas uma infraestrutura, 3,66 artigos no período, permanece bastante semelhante com a do total de coordenadores (3,71 artigos no quinquênio). Por tempo de obtenção do doutorado, as maiores médias se mantêm entre os coordenadores de 24 a 33 anos de doutorado (5,23 artigos no período) seguindo pelos de 34 a 51 anos (4,71 artigos no período). No mesmo sentido, os de formação em Ciências da Saúde ainda possuem as maiores médias, 4,15 artigos no quinquênio e os da Engenharias com menores médias, 2,06 artigos nos cinco anos.

Quando avalia-se a quantidade de artigos publicados pelos coordenadores ponderada pelo IFPA, tem-se um aumento no percentual de coordenadores que publicaram zero no quinquênio (11,21%). Isso ocorre devido à queda da quantidade de artigos publicados que são indexados internacionalmente. As estatísticas descritivas demonstram que os coordenadores com maior tempo de doutorado, ou seja, aqueles que obtiveram o título de doutorado obtido entre 1962 e 1979, apresentaram as maiores médias (2,7 artigos no quinquênio), enquanto que os mais novos, título obtido entre 2000 e 2009, registram as menores médias (1,16 artigos no quinquênio). Em relação às grandes áreas de formação dos coordenadores, aqueles que pertencem às Ciências da Saúde permanecem com médias mais elevadas (2,87 artigos nos cinco anos) e os das Engenharias com menores médias, 1,14 artigos nos cinco anos.

Por fim, em relação à média de produção dos demais pesquisadores permanentes por infraestrutura, tem-se que os mesmos publicaram 1,53 artigos/quinquênio e quando pondera-se pelo indexador IFPA tal produção cai para 0,54 artigos/quinquênio.

Ao analisar a quantidade de artigos publicados pelos coordenadores por agrupamentos de *cluster* (Tabela 31), tem-se que, em geral, os coordenadores do *cluster* 2 possuem as maiores médias produtivas no quinquênio e resultados superiores de desigualdades dado pelas

estatísticas de assimetria e curtose, independente da medida de produtividade científica analisada. Como grupo intermediário, apresenta-se os coordenadores do *cluster* 1, exceto quando pondera-se a produtividade científica pelo IFPA, pois passa a apresentar menores índices de desigualdade. E, por fim, têm-se os coordenadores do *cluster* 3, com os menores resultados de produtividade científica média e desigualdades.

Tabela 31 – Estatísticas descritivas das variáveis da produtividade científica dos coordenadores, por *clusters*.

Variável	Qtd. de observações	Quantidade	Média	Desvio Padrão	Assimetria	Curtose
Artigos publicados no quinquênio pelos coordenadores	1.337	22.926	17,15	21,15	6,71	92,51
Por cluster:						
1	207	3.333	16,10	15,79	3,00	21,94
2	647	13.980	21,61	26,53	6,53	75,67
3	483	5.613	11,62	12,52	2,36	10,11
Artigos publicados no quinquênio/coautores pelos coordenadores	1.337	4.894,45	3,66	4,55	9,26	177,08
Por cluster:						
1	207	687,27	3,32	3,26	2,92	19,14
2	647	2.932,06	4,53	5,63	9,34	149,58
3	483	1.275,12	2,64	2,88	2,70	14,32
Artigos publicados no quinquênio/coautores pelos coordenadores ponderados pelo IFPA	1.337	2.134,40	1,60	2,32	4,71	42,38
Por cluster:						
1	207	278,25	1,34	1,65	1,74	6,16
2	647	1.370,37	2,12	2,88	4,37	33,01
3	483	485,76	1,01	1,38	2,50	11,61

Fonte: Elaboração própria, a partir da Plataforma *Lattes* e ISI.

Ao realizar a análise de como a mão de obra está dividida por infraestrutura (Tabela 32), observa-se que as infraestruturas são compostas, principalmente, pelos pesquisadores denominados permanentes (média de 54,97%), seguido pelos não-pesquisadores (média de 38,90%) e pelos pesquisadores não-permanentes (média de 6,13%). Quando analisa-se como se comporta o percentual de coordenadores por infraestrutura, nota-se que, em média, tem-se 21,48% do pessoal total ocupado exercendo a função de coordenador.

Tabela 32 – Descritiva das variáveis de mão de obra na infraestrutura e por *cluster*.

Variável	Média	Desvio Padrão	Mínimo	Máximo
Perc. de pesquisadores permanentes	0,5497	0,32	0,21	1,00
<i>Cluster 1</i>	0,5116	0,31	0,03	1,00
<i>Cluster 2</i>	0,5899	0,31	0,03	1,00
<i>Cluster 3</i>	0,5769	0,32	0,03	1,00
Perc. de pesquisadores não-permanentes	0,0613	0,15	0	0,8000
<i>Cluster 1</i>	0,0672	0,13	0	0,5714
<i>Cluster 2</i>	0,0729	0,16	0	0,8000
<i>Cluster 3</i>	0,0480	0,13	0	0,8000
Perc. de não-pesquisadores	0,3890	0,32	0	0,9787
<i>Cluster 1</i>	0,4212	0,31	0	0,9677
<i>Cluster 2</i>	0,3372	0,30	0	0,9787
<i>Cluster 3</i>	0,3751	0,32	0	0,9600
Perc. de coordenadores	0,2148	0,23	0,002	1,00
<i>Cluster 1</i>	0,1885	0,22	0,008	1,00
<i>Cluster 2</i>	0,1675	0,16	0,002	1,00
<i>Cluster 3</i>	0,2521	0,24	0,009	1,00

Fonte: Elaboração própria, a partir do Mapeamento do MCTI.

Ao realizar a análise por *cluster*, tem-se que o grupo 2 tem o maior percentual médio de pesquisadores permanentes e de não-permanentes, por outro lado possui o menor percentual de não-pesquisadores e de coordenadores em sua composição, quando comparado aos demais *clusters*. Já o grupo 1 tem, em média, o mais baixo percentual de pesquisadores permanentes, mais alta média de percentual de não-pesquisadores e uma média intermediária de percentual de pesquisadores não-permanentes e de coordenadores. Enfim, o grupo 3 tem uma média intermediária do percentual de pesquisadores permanentes e não-pesquisadores, além disso possui a mais baixa média percentual de pesquisadores não-permanentes e a mais alta no percentual de coordenadores.

Adicionalmente, a relação capital-trabalho, apresentada na Tabela 33, descreve que por pessoal total ocupado (PO), o valor médio do capital das infraestruturas é de R\$ 129.063,70. Por *clusters*, o grupo 1 tem a maior média da relação capital trabalho, enquanto que o grupo 3 apresenta os menores valores.

Tabela 33 – Descritiva da variável relação capital-trabalho nas infraestruturas e por *cluster*.

Variável	Média	Desvio Padrão	Mínimo	Máximo
Capital-Trabalho	129.063,70	248.151,40	0	4.000.000
<i>Cluster 1</i>	179.825,10	392.622,00	2.083,33	4.000.000
<i>Cluster 2</i>	151.451,30	268.435,30	943,40	2.333.333
<i>Cluster 3</i>	113.788,60	203.845,70	0	1.785.714

Fonte: Elaboração própria, a partir do Mapeamento do MCTI.

6.2. Modelo de Produtividade Científica Individual

As Tabelas 34, 35 e 36 apresentam os efeitos marginais dos valores esperados não condicional da produtividade científica dos coordenadores gerados pelo procedimento econométrico Tobit, de acordo com as equações econométricas (6, 8 e 9). Adicionalmente, serão expostos os resultados estimados por Mínimos Quadrados Ordinários dos determinantes coletivos e individuais, para efeitos de comparação. Também foram realizadas as mesmas estimações das variáveis que compõem a Função de Produção, porém em números absolutos (Apêndice D), ou seja, sem dividir pelo pessoal total ocupado (PO).

As estimações são realizadas, sempre, seguindo a ordem das especificações das equações econométricas 6, 8 e 9, a fim de mostrar qual o comportamento das variáveis coletivas das infraestruturas, à medida que são incluídas variáveis de controle. Na Tabela 34 são apresentados os parâmetros estimados por OLS e os efeitos marginais não-condicional do Tobit para toda a amostra (1.652 cargos de coordenação). Porém, sabe-se da existência de 108 coordenadores que exercem tal função em mais de uma infraestrutura. Portanto, para resolver tal problema usa-se um peso amostral como o número de observações na população que cada observação representa, que aqui é o inverso da quantidade de vezes que o coordenador aparece na amostra.

Nota-se, em um primeiro momento, que com a inclusão das variáveis de controle (características individuais dos coordenadores), ocorre uma diminuição na magnitude de todos os coeficientes estimados. Mais detalhadamente, os parâmetros associados à variável de produção média dos pesquisadores permanentes, acumulado no período de 2009 a 2013,

foram estatisticamente significativos a 1%, com elasticidades positivas em todos os modelos estimados. Ou seja, pode ser devido às externalidades positivas, apontadas por Mairesse e Turner (2002) e Carayol e Matt (2003), da atividade de publicação dos pesquisadores dentro da infraestrutura.

Em relação à composição do capital humano, nota-se a grande importância dos estudantes bolsistas nas infraestruturas de pesquisa, com elasticidades positivas, em média, de 0,774. Ou seja, uma maior participação dos pesquisadores não-permanentes contribui com maiores produtividades científicas dos coordenadores. Quanto às variáveis *dummies* de *cluster*, tanto os coordenadores pertencentes ao grupo 1 (laboratórios intensivos em pesquisa e desenvolvimento de tecnologia) quanto 3 (laboratórios intensivos em pesquisa), possuem produtividades científicas menores, em torno de -24,00% e -35,79%, respectivamente, quando comparados ao grupo 2 (laboratórios de multiatividades). Exceto quando estima-se com a inclusão da variável dependente defasada. Portanto, em geral, pode-se argumentar que infraestruturas que se dedicam as atividades conjuntas de pesquisa, ensino e desenvolvimento de tecnologia impactam positivamente a produtividade dos coordenadores brasileiros.

Entre as variáveis que apresentaram significâncias apenas no modelo com variáveis coletivas estão: pessoal total ocupado, percentual de não-pesquisadores e relação capital-trabalho. Aqui, nota-se um efeito negativo, porém baixo, dos parâmetros estimados da variável pessoal ocupado na infraestrutura, o que evidencia rendimentos decrescentes de escala. Em relação à presença de não-pesquisadores, notam-se elasticidades positivas e, portanto, impacta em maior produtividade dos coordenadores. Quanto à relação capital-trabalho, observa-se elasticidades positivas e estatisticamente significantes ao nível de 5%, o que ratifica a importância dos investimentos em recursos físicos nas infraestruturas na determinação dos diferenciais produtivos dos coordenadores.

Com a inclusão de variáveis de controle, ou seja, das características individuais dos coordenadores, observam-se resultados dos efeitos marginais bem semelhantes nas estimações por OLS e Tobit. O regressor representativo da produção científica do coordenador no quinquênio anterior foi incluído a fim de captar as características de trajetória dependente da produção científica. O parâmetro estimado de tal variável mostra-se estatisticamente significativo em um nível de 1%, com elasticidades em torno de 0,705. Isso pode estar relacionado à “vantagem cumulativa” citada pela literatura, favorecendo os coordenadores que produziram no passado a produzir mais no presente. Nota-se, também, que não houve mudanças na importância da participação dos pesquisadores não-permanentes e da produção média dos pesquisadores permanentes na produtividade científica do coordenador.

A variável dummy que determina se o coordenador trabalha em uma ou mais infraestruturas apresentou coeficientes significantes e positivos, independente do método analisado. Portanto, infere-se que os mesmos são menos produtivos, em média, 17,64% quando realizam atividades em apenas uma infraestrutura. Em relação aos efeitos geracionais, em geral, os coordenadores de formação mais recente (1990 a 1999 e 2000 a 2009) são mais produtivos que as gerações mais antigas (1962 a 1979 e 1980 a 1989). Da mesma maneira, nota-se que os coordenadores que atuam em Multiárea são mais produtivos que os atuantes na área de Ciências Exatas e da Terra.

As Tabelas 35 e 36 apresentam os resultados discriminados por agrupamentos de infraestruturas, conforme definidos pelo método de *cluster* não hierárquico. É importante notar que, nesse momento, estimou-se com os coordenadores que trabalham em apenas uma infraestrutura, por isso a diminuição do número de observações para 1.337 coordenadores.

A análise da Tabela 35 demonstra que, o *cluster* 1 (laboratórios intensivos em pesquisa e desenvolvimento de tecnologia), independente do método analisado, apresentou coeficientes da produção média dos pesquisadores permanentes maiores que os demais grupos, com elasticidades positivas em torno de 0,219. Quanto ao capital humano, no primeiro momento, nota-se a importância dos pesquisadores permanentes, não-permanentes e dos não-pesquisadores com elasticidades médias de 1,971, 1,728 e 1,381, respectivamente. No que diz respeito à idade da infraestrutura, foram obtidos coeficientes conforme preditos pela literatura, isto é, trajetória de retornos marginais decrescentes. Porém, com a inclusão do fator autorregressivo, há uma perda na significância dos parâmetros estimados associados à mão de obra da infraestrutura e nota-se impactos gerados, apenas, pelos coeficientes da produção dos pesquisadores permanentes no quinquênio de análise, com elasticidades médias de 0,088, e do fator autorregressivo, os quais apresentam elasticidades médias de 0,683.

No *cluster* 2 (laboratórios de multiatividades), nota-se uma elasticidade positiva da produção dos pesquisadores permanentes, em torno de 0,193. Não observou-se efeitos marginais significativos da mão de obra da infraestrutura. Por outro lado, observa-se que a presença de pesquisadores visitantes contribui para a elevação média de 22,26% na produtividade científica do coordenador. Por fim, a relação capital-trabalho também apresentou impactos positivos, com elasticidades em torno de 0,056. Com a inclusão da variável defasada, o qual apresentou coeficientes em torno de 0,706, não houve diferenças de sinal e significância dos parâmetros estimados, apenas diminuição das magnitudes. Exceto pela queda da relevância da relação capital-trabalho.

Por fim, os resultados para o *cluster* 3 (laboratórios intensivos em pesquisa) demonstram que somente os parâmetros estimados das variáveis do fator autorregressivo (elasticidades médias de 0,72), produção média dos pesquisadores permanentes e o percentual de estudantes bolsistas (pesquisadores não-permanentes) contribuem positivamente para a produtividade científica do coordenador.

A Tabela 36 demonstra as mesmas análises, porém controladas pelas demais características individuais do coordenador. No geral, quanto às características coletivas, os *clusters* apresentaram os mesmo resultados da tabela anterior, exceto o *cluster* 1 que não apresentou coeficientes significativos da produção média dos pesquisadores permanentes. Quanto aos *dummies* de geração, os coordenadores do *cluster* 2 e da geração mais nova (doutorado obtido entre 2000 e 2009) apresentaram-se mais produtivos, em relação aos mais velhos (doutorado obtido entre 1962-1979). Por outro lado, os coordenadores do *cluster* 1 e 3 não apresentaram essas diferenças estatísticas. Quanto às grandes áreas do conhecimento dos coordenadores, sendo Multiárea como referência, nota-se que no *cluster* 1, os coordenadores das áreas de Ciências Biológicas (-29,71%), Ciências da Saúde (-23,97%) e Ciências Agrárias (-48,65%) são menos produtivos, enquanto que os pertencentes às Engenharias são 40,21% mais produtivos. No *cluster* 2, coordenadores das áreas de Ciências Exatas e da Terra e Engenharias apresentaram produtividades menores em média de -18,82% e -15,30%, respectivamente. Por fim, o *cluster* 3 apresentou efeitos marginais negativos sobre a produtividade dos coordenadores pertencentes às áreas de Ciências Exatas e da Terra (-21,51%) e efeitos positivos se pertencentes à área de Engenharias (23,06%).

Tabela 34 – Resultados das Regressões – Efeitos Marginais da Produtividade Científica dos coordenadores contra as características coletivas e individuais.

	y_{in}/coautores					
	OLS	Tobit	OLS	Tobit	OLS	Tobit
Prod. Cient. do quinquênio anterior			0.695*** (0.0171)	0.714*** (0.019)	0.701*** (0.0196)	0.720*** (0.021)
Prod. dos Pesquisadores	0.190*** (0.0210)	0.197*** (0.022)	0.0692*** (0.0148)	0.072*** (0.015)	0.0672*** (0.0147)	0.070*** (0.015)
Permanentes no quinquênio	-0.0573* (0.0331)	-0.062* (0.035)	0.0212 (0.0228)	0.020 (0.024)	0.0315 (0.0228)	0.031 (0.024)
Pessoal Ocupado	0.312 (0.276)	0.345 (0.288)	0.225 (0.181)	0.254 (0.189)	0.205 (0.183)	0.232 (0.190)
Pesquisadores Permanentes	0.750*** (0.195)	0.798*** (0.202)	0.329** (0.130)	0.359*** (0.135)	0.319** (0.131)	0.347** (0.135)
Pesquisadores Não-Permanentes	0.371* (0.213)	0.402* (0.222)	0.131 (0.140)	0.154 (0.146)	0.0876 (0.141)	0.109 (0.147)
Não-Pesquisadores	0.0611 (0.0775)	0.066 (0.080)	0.0220 (0.0577)	0.026 (0.059)	0.0432 (0.0569)	0.048 (0.058)
Dummy de Pesq. Visitantes	0.0347** (0.0141)	0.037** (0.015)	0.00813 (0.00984)	0.009 (0.010)	0.0131 (0.00991)	0.014 (0.010)
Relação Capital-Trabalho	0.181 (0.214)	0.174 (0.222)	-0.0199 (0.145)	-0.031 (0.150)	0.0357 (0.149)	0.036 (0.154)
Idade do Laboratório	-0.0239 (0.0387)	-0.023 (0.040)	-0.000883 (0.0262)	0.001 (0.027)	-0.00893 (0.0268)	-0.009 (0.028)
Idade do Laboratório ao Quadrado	-0.269*** (0.0610)	-0.280*** (0.064)	-0.0608 (0.0432)	-0.063 (0.045)	-0.0744* (0.0432)	-0.078* (0.045)
Cluster 1						
Cluster 3	-0.435*** (0.0469)	-0.451*** (0.049)	-0.133*** (0.0340)	-0.138*** (0.035)	-0.132*** (0.0340)	-0.137*** (0.035)
Trabalha em mais de 1 Infraest.					0.159*** (0.0351)	0.166*** (0.036)
Geração: 1962-1979					-0.187** (0.0740)	-0.200*** (0.075)
Geração: 1980-1989					-0.0784 (0.0481)	-0.088* (0.050)
Geração: 1990-1999					-0.00692 (0.0349)	-0.014 (0.037)
Pesquisador da área de Ciências Exatas e da Terra					-0.198*** (0.0665)	-0.196*** (0.068)
Pesquisador da área de Engenharias					0.0155 (0.0671)	0.026 (0.067)
Pesquisador da área de Ciências Biológicas					-0.0451 (0.0644)	-0.072 (0.071)
Pesquisador da área de Ciências da Saúde					0.0104 (0.105)	-0.012 (0.118)
Pesquisador da área de Ciências Agrárias					0.0220 (0.107)	0.021 (0.110)
Constante	1.030*** (0.371)		0.528** (0.248)		0.362 (0.255)	
N	1.652	1.652	1.652	1.652	1.652	1652
R2	0.145		0.590		0.598	
Pseudo R2		0.0599		0.3367		0.3445

Fonte: Elaboração própria.

Tabela 35 – Resultados das Regressões – Efeitos Marginais da Produtividade Científica dos coordenadores contra as características coletivas por *Cluster*.

	$y_{in}/coautores$											
	OLS			Tobit			OLS			Tobit		
	<i>cluster 1</i>	<i>cluster 2</i>	<i>cluster 3</i>	<i>cluster 1</i>	<i>cluster 2</i>	<i>cluster 3</i>	<i>cluster 1</i>	<i>cluster 2</i>	<i>cluster 3</i>	<i>cluster 1</i>	<i>cluster 2</i>	<i>cluster 3</i>
Prod. Cient. do quinquênio anterior							0.673*** (0.0500)	0.701*** (0.0289)	0.695*** (0.0272)	0.692*** (0.052)	0.711*** (0.030)	0.729*** (0.031)
Prod. dos Pesquisadores Permanentes no quinquênio	0.215*** (0.0587)	0.192*** (0.0353)	0.173*** (0.0331)	0.222*** (0.060)	0.194*** (0.036)	0.186*** (0.036)	0.0875* (0.0484)	0.063*** (0.0226)	0.068*** (0.0245)	0.089* (0.049)	0.062*** (0.023)	0.075*** (0.026)
Pessoal Ocupado	-0.0300 (0.0851)	-0.0682 (0.0494)	-0.0283 (0.0661)	-0.033 (0.088)	-0.071 (0.051)	-0.039 (0.071)	0.0821 (0.0557)	-0.0106 (0.0341)	0.0407 (0.0441)	0.083 (0.058)	-0.012 (0.035)	0.035 (0.048)
Pesquisadores Permanentes	1.913** (0.919)	-0.350 (0.399)	0.479 (0.524)	2.028** (0.933)	-0.356 (0.404)	0.544 (0.561)	0.862 (0.592)	-0.113 (0.276)	0.516 (0.330)	0.942 (0.597)	-0.115 (0.279)	0.584 (0.355)
Pesquisadores Não-Permanentes	1.664*** (0.565)	0.142 (0.281)	1.227*** (0.369)	1.791*** (0.579)	0.148 (0.284)	1.325*** (0.392)	0.548 (0.389)	0.0629 (0.194)	0.655*** (0.235)	0.625 (0.393)	0.065 (0.196)	0.719*** (0.249)
Não-Pesquisadores	1.327* (0.713)	0.00688 (0.304)	0.410 (0.397)	1.435** (0.725)	0.008 (0.307)	0.467 (0.428)	0.427 (0.469)	-0.0512 (0.209)	0.342 (0.249)	0.497 (0.474)	-0.052 (0.210)	0.400 (0.271)
Dummy de Pesq. Visitantes	-0.174 (0.219)	0.198** (0.0957)	-0.0991 (0.170)	-0.191 (0.225)	0.204** (0.095)	-0.084 (0.177)	-0.116 (0.155)	0.128* (0.0715)	-0.0839 (0.133)	-0.123 (0.159)	0.132* (0.071)	-0.071 (0.137)
Relação Capital-Trabalho	0.0589 (0.0379)	0.0551** (0.0214)	0.0191 (0.0243)	0.062 (0.040)	0.057*** (0.022)	0.021 (0.026)	0.00415 (0.0282)	0.00844 (0.0155)	0.0111 (0.0165)	0.005 (0.030)	0.009 (0.016)	0.012 (0.017)
Idade do Laboratório	1.776** (0.731)	-0.249 (0.308)	0.418 (0.407)	1.915** (0.776)	-0.257 (0.311)	0.401 (0.432)	0.654 (0.485)	-0.295 (0.209)	0.198 (0.262)	0.756 (0.513)	-0.308 (0.210)	0.178 (0.278)
Idade do Laboratório ao Quadrado	-0.331** (0.139)	0.0575 (0.0519)	-0.0693 (0.0775)	-0.359** (0.149)	0.059 (0.052)	-0.066 (0.082)	-0.140 (0.0915)	0.0547 (0.0352)	-0.0449 (0.0490)	-0.162* (0.097)	0.057 (0.035)	-0.042 (0.052)
Constante	-0.804 (1.196)	0.908* (0.551)	0.729 (0.686)				-0.172 (0.791)	0.708* (0.369)	0.307 (0.448)			
N	207	647	483	207	647	483	207	647	483	207	647	483
R2	0.157	0.081	0.093				0.560	0.570	0.560			
Pseudo R2				0.0652	0.0347	0.0370				0.3041	0.3462	0.3047

Fonte: Elaboração Própria.

Tabela 36 – Resultados das Regressões – Efeitos Marginais da Produtividade Científica dos coordenadores contra as características coletivas e individuais por *Cluster*.

	yin/coautores					
	OLS			Tobit		
	cluster 1	cluster 2	cluster 3	cluster 1	cluster 2	cluster 3
Prod. Cient. do quinquênio anterior	0.714*** (0.0536)	0.717*** (0.0329)	0.689*** (0.0316)	0.734*** (0.055)	0.728*** (0.034)	0.721*** (0.034)
Prod. dos Pesquisadores	0.0612 (0.0480)	0.0561** (0.0227)	0.0686*** (0.0249)	0.061 (0.048)	0.055** (0.023)	0.075*** (0.026)
Permanentes no quinquênio	0.0898 (0.0579)	-0.00256 (0.0346)	0.0491 (0.0454)	0.090 (0.059)	-0.004 (0.035)	0.042 (0.049)
Pessoal Ocupado	0.772 (0.665)	-0.152 (0.278)	0.467 (0.335)	0.857 (0.656)	-0.156 (0.280)	0.537 (0.356)
Pesquisadores Permanentes	0.483 (0.422)	0.0283 (0.196)	0.617*** (0.237)	0.565 (0.419)	0.029 (0.197)	0.678*** (0.248)
Pesquisadores Não-Permanentes	0.351 (0.519)	-0.0800 (0.211)	0.268 (0.253)	0.424 (0.513)	-0.081 (0.212)	0.330 (0.271)
Não-Pesquisadores	-0.115 (0.159)	0.128* (0.0708)	-0.0406 (0.130)	-0.123 (0.159)	0.133* (0.070)	-0.027 (0.135)
Dummy de Pesq. Visitantes	0.00382 (0.0299)	0.00839 (0.0159)	0.0167 (0.0166)	0.004 (0.031)	0.009 (0.016)	0.018 (0.017)
Relação Capital-Trabalho	0.718 (0.499)	-0.288 (0.205)	0.226 (0.278)	0.817 (0.515)	-0.292 (0.205)	0.224 (0.292)
Idade do Laboratório	-0.148 (0.0934)	0.0540 (0.0343)	-0.0495 (0.0517)	-0.169* (0.097)	0.055 (0.034)	-0.049 (0.054)
Idade do Laboratório ao Quadrado	-0.0593 (0.148)	-0.269** (0.109)	-0.0219 (0.129)	-0.065 (0.150)	-0.285** (0.111)	-0.033 (0.129)
Geração: 1962-1979	-0.243 (0.165)	-0.0823 (0.0670)	-0.0168 (0.100)	-0.246 (0.167)	-0.089 (0.068)	-0.028 (0.104)
Geração: 1980-1989	-0.0995 (0.101)	0.00522 (0.0529)	-0.00354 (0.0638)	-0.104 (0.102)	-0.001 (0.054)	-0.017 (0.068)
Geração: 1990-1999	0.219 (0.158)	-0.212* (0.109)	-0.238** (0.0975)	0.247 (0.155)	-0.205* (0.107)	-0.241** (0.101)
Pesquisador de Ciências Exatas e da Terra	0.329* (0.173)	-0.168* (0.0929)	0.198** (0.0936)	0.347** (0.171)	-0.164* (0.092)	0.217** (0.093)
Pesquisador de Engenharias	-0.370** (0.175)	0.0595 (0.0991)	-0.0981 (0.0955)	-0.335* (0.180)	0.056 (0.101)	-0.168 (0.115)
Pesquisador de Ciências Biológicas	-0.274* (0.159)	0.0407 (0.154)	-0.172 (0.223)	-0.253 (0.160)	0.043 (0.151)	-0.308 (0.329)
Pesquisador de Ciências da Saúde	-0.674*** (0.193)	0.0388 (0.0997)	0.0671 (0.235)	-0.659*** (0.184)	0.040 (0.099)	0.052 (0.243)
Pesquisador de Ciências Agrárias	-0.291 (0.845)	0.682* (0.374)	0.175 (0.470)			
Constante						
N	207	647	483	207	647	483
R2	0.579	0.581	0.568			
Pseudo R2				0.3192	0.3565	0.3133

Fonte: Elaboração Própria.

6.3. Modelo de Produtividade Científica Individual ponderada pelo IFPA

As Tabelas 37, 38 e 39 apresentam os efeitos marginais da produtividade científica dos coordenadores ponderada pelo IFPA gerados pelos procedimentos econométricos de acordo com as equações econométricas 6, 8 e 9. Tal estratégia tem o objetivo de avaliar, de forma igualitária, a produtividade científica dos coordenadores de acordo com a qualidade dos artigos publicados em diferentes áreas do conhecimento. Realizou-se, também, as estimações ponderadas somente pelo IFPA com as variáveis absolutas da Função de Produção (Apêndice E). Dos resultados, nota-se que os resultados, apesar de similares, apresentam diferenças quanto aos coeficientes estimados da mão de obra.

Na Tabela 37, os resultados das estimações, somente com características coletivas, demonstram, independente do método econométrico utilizado, elasticidades positivas da produção média dos pesquisadores permanentes ponderada pela qualidade; dos pesquisadores não-permanentes; e da relação capital-trabalho sobre a produtividade científica individual do coordenador ponderada pelo IFPA. Quanto às *dummies* de *cluster*, nota-se que há uma produtividade maior entre os coordenadores pertencentes ao grupo 2, em relação ao 1 e 3. Com a inclusão do fator autorregressivo, nota-se que tal variável apresentou elasticidades médias de 0,728, além disso dentre as características coletivas, os determinantes da quantidade de artigos publicados ponderada pela qualidade são semelhantes ao modelo anterior, porém nota-se menores magnitudes dos coeficientes estimados. Da mesma maneira, ocorreu com o acréscimo das variáveis de características individuais dos coordenadores. Quanto ao coordenador que trabalha em mais de uma infraestrutura são, em média, 16,88% mais produtivos em termos de qualidade de publicação no período, quando comparados com os que exercem atividades em apenas uma infraestrutura.

Nas Tabelas 38 e 39, as análises são separadas por tipo de *cluster*. No primeiro grupo, observa-se a importância dos parâmetros estimados da produção média ponderada pelo IFPA dos pesquisadores permanentes no quinquênio de análise para maior produtividade dos seus coordenadores, independente do método econométrico utilizado e mantido com a inclusão da variável dependente defasada. Da mesma maneira ocorre com a mão de obra, de tal forma que impactam de forma significativa e positiva em todos os níveis (pesquisadores permanentes, pesquisadores não-permanentes e não-pesquisadores). Além disso, os coeficientes associados

à relação capital-trabalho, em geral, geram efeitos positivos e significativos sobre a produtividade científica do coordenador. Nota-se, também, nesse grupo retornos marginais decrescentes da idade da infraestrutura, porém tornam-se não significativos a partir da inclusão do fator autorregressivo. Para o segundo agrupamento, os determinantes dos diferenciais de produtividade dos coordenadores ponderada pela qualidade são: produção média dos pesquisadores permanentes ponderada pelo IFPA e a relação capital-trabalho. Com a inclusão da variável de produção dos coordenadores no quinquênio anterior (elasticidade média de 0,704), nota-se a permanência dos sinais e significâncias anteriores, além disso quanto à variável de idade da infraestrutura, há um retorno marginal crescente sobre a variável dependente. No terceiro agrupamento, nota-se que a produtividade científica do coordenador é impactada positivamente pelos coeficientes estimados da produção média dos pesquisadores permanentes indexada pelo índice IFPA, do percentual de pesquisadores não-permanentes e da relação capital-trabalho. Por outro lado, sofre efeitos marginais negativos com a presença de pesquisadores visitantes. Com o acréscimo do fator autorregressivo, não ocorrem alterações e tal variável apresenta-se positiva e significativa, com elasticidades média de 0,711.

Com a inclusão das variáveis individuais (Tabela 39), nota-se que os determinantes da produtividade científica dos coordenadores pertencentes ao *cluster* 1 são: própria produção no quinquênio anterior (elasticidade média de 0,795); produção dos pesquisadores permanentes no quinquênio atual (elasticidade média de 0,108); percentual de pesquisadores permanentes, não-permanentes e não-pesquisadores (elasticidade média de 1,635; 1,00; 1,312, respectivamente). A produtividade científica dos coordenadores do *cluster* 2, em geral, é determinada pelo fator autorregressivo, pela produção dos pesquisadores permanentes no quinquênio atual e da relação capital-trabalho, os quais possuem elasticidades médias de 0,709, 0,077 e 0,039, respectivamente. Os coeficientes das variáveis idade e idade ao quadrado da infraestrutura produzem impactos com retornos marginais crescentes. Quanto aos efeitos geracionais, os coordenadores de formação recente são mais produtivos quando comparados com os da geração de 1962-1979 (-18,45%). Os coordenadores do *cluster* 3 têm sua produtividade determinada pelos parâmetros estimados do fator autorregressivo, da produção dos pesquisadores permanentes no quinquênio atual e da relação capital-trabalho, com elasticidades médias de 0,691; 0,069; e 0,046, respectivamente. Além disso, sofre efeitos marginais negativos pela presença de pesquisadores visitantes (-23,51%). O efeito geracional é presente na *coorte* mais antiga (1962-1979), os quais são mais produtivos, em 36,89%, quando comparados com a *coorte* mais nova (2000-2009). Quanto às variáveis de grande área

do conhecimento, em todos os *clusters* analisados, nota-se que, em geral, os coordenadores pertencentes à Multiárea são mais produtivos em comparação com as demais áreas.

Tabela 37 – Resultados das Regressões – Efeitos Marginais da Produtividade Científica dos coordenadores contra as características coletivas e individuais ponderados pelo IFPA.

	y_{in}/coautores*IFPA					
	OLS	Tobit	OLS	Tobit	OLS	Tobit
Prod. Cient. do quinquênio anterior			0.718*** (0.0163)	0.738*** (0.018)	0.708*** (0.0179)	0.728*** (0.019)
Prod. dos Pesquisadores	0.195*** (0.0200)	0.194*** (0.020)	0.0818*** (0.0136)	0.088*** (0.015)	0.0831*** (0.0138)	0.090*** (0.015)
Permanentemente no quinquênio						
Pessoal Ocupado	0.000463 (0.0313)	-0.006 (0.031)	0.0127 (0.0217)	0.008 (0.023)	0.0195 (0.0217)	0.017 (0.023)
Pesquisadores Permanentemente	0.172 (0.256)	0.217 (0.256)	0.164 (0.176)	0.217 (0.189)	0.153 (0.177)	0.196 (0.190)
Pesquisadores Não-Permanentemente	0.593*** (0.183)	0.646*** (0.181)	0.296** (0.124)	0.370*** (0.132)	0.300** (0.124)	0.368*** (0.132)
Não-Pesquisadores	0.227 (0.197)	0.271 (0.197)	0.157 (0.137)	0.209 (0.147)	0.128 (0.139)	0.169 (0.149)
Dummy de Pesq. Visitantes	-0.0470 (0.0723)	-0.064 (0.073)	-0.0364 (0.0588)	-0.052 (0.063)	-0.0305 (0.0586)	-0.044 (0.063)
Relação Capital-Trabalho	0.102*** (0.0140)	0.105*** (0.014)	0.0318*** (0.00984)	0.038*** (0.011)	0.0341*** (0.00997)	0.042*** (0.011)
Idade do Laboratório	-0.0793 (0.263)	-0.095 (0.251)	-0.264 (0.231)	-0.293 (0.234)	-0.260 (0.241)	-0.272 (0.244)
Idade do Laboratório ao Quadrado	0.0113 (0.0488)	0.014 (0.046)	0.0461 (0.0436)	0.051 (0.044)	0.0449 (0.0454)	0.047 (0.046)
Cluster 1	-0.300*** (0.0592)	-0.293*** (0.055)	-0.0955** (0.0385)	-0.110*** (0.041)	-0.0983** (0.0389)	-0.114*** (0.041)
Cluster 3	-0.395*** (0.0450)	-0.397*** (0.043)	-0.113*** (0.0327)	-0.134*** (0.035)	-0.113*** (0.0328)	-0.134*** (0.035)
Trabalha em mais de 1 Infraest.					0.150*** (0.0436)	0.162*** (0.044)
Geração: 1962-1979					-0.0401 (0.0677)	-0.062 (0.070)
Geração: 1980-1989					0.00684 (0.0465)	-0.001 (0.049)
Geração: 1990-1999					0.0586* (0.0328)	0.052 (0.035)
Pesquisador da área de Ciências					-0.0658 (0.0631)	-0.082 (0.067)
Pesquisador da área de Engenharias					0.0350 (0.0673)	0.064 (0.066)
Pesquisador da área de Ciências Biológicas					0.0106 (0.0669)	-0.034 (0.077)
Pesquisador da área de Ciências da Saúde					0.157 (0.157)	0.152 (0.158)
Pesquisador da área de Ciências Agrárias					0.0745 (0.0773)	0.106 (0.078)
Constante	-0.0589 (0.421)		0.410 (0.344)		0.327 (0.357)	
N	1.652	1.652	1.652	1.652	1.652	1.652
R2	0.172		0.602		0.606	
Pseudo R2		0.0739		0.3368		0.3416

Fonte: Elaboração Própria.

Tabela 38 – Resultados das Regressões – Efeitos Marginais da Produtividade Científica dos coordenadores contra as características coletivas por *Cluster* ponderados pelo IFPA.

	$y_{in}/\text{coautores*IFPA}$											
	OLS			Tobit			OLS			Tobit		
	<i>cluster 1</i>	<i>cluster 2</i>	<i>cluster 3</i>	<i>cluster 1</i>	<i>cluster 2</i>	<i>cluster 3</i>	<i>cluster 1</i>	<i>cluster 2</i>	<i>cluster 3</i>	<i>cluster 1</i>	<i>cluster 2</i>	<i>cluster 3</i>
Prod. Cient. do quinquênio anterior							0.763*** (0.0371)	0.695*** (0.0264)	0.707*** (0.0335)	0.792*** (0.041)	0.713*** (0.028)	0.715*** (0.036)
Prod. dos Pesquisadores	0.231*** (0.0519)	0.191*** (0.0361)	0.173*** (0.0301)	0.220*** (0.049)	0.186*** (0.036)	0.169*** (0.029)	0.102*** (0.0372)	0.073*** (0.0235)	0.064*** (0.0209)	0.103*** (0.039)	0.072*** (0.024)	0.072*** (0.022)
Permanentes no quinquênio	-0.0569 (0.0790)	0.00790 (0.0502)	0.0389 (0.0521)	-0.069 (0.075)	0.008 (0.050)	0.021 (0.051)	0.0205 (0.0446)	0.00797 (0.0350)	0.0201 (0.0372)	0.010 (0.047)	0.008 (0.036)	0.008 (0.040)
Pessoal Ocupado	1.952** (0.831)	-0.310 (0.407)	0.331 (0.402)	1.929** (0.775)	-0.302 (0.403)	0.364 (0.403)	1.492*** (0.453)	-0.0698 (0.284)	0.115 (0.283)	1.621*** (0.466)	-0.069 (0.293)	0.176 (0.308)
Pesquisadores Permanentes	1.612*** (0.530)	0.173 (0.292)	0.967*** (0.288)	1.641*** (0.499)	0.198 (0.288)	0.969*** (0.280)	0.866*** (0.300)	0.177 (0.197)	0.329 (0.207)	0.987*** (0.313)	0.205 (0.203)	0.394* (0.219)
Pesquisadores Não-Permanentes	1.557** (0.646)	-0.0326 (0.308)	0.123 (0.306)	1.590*** (0.602)	-0.031 (0.305)	0.182 (0.309)	1.174*** (0.366)	-0.0373 (0.218)	0.0937 (0.215)	1.315*** (0.377)	-0.038 (0.225)	0.169 (0.238)
Não-Pesquisadores	-0.245 (0.163)	0.0299 (0.103)	-0.233* (0.137)	-0.246 (0.155)	0.030 (0.101)	-0.261** (0.133)	-0.151 (0.118)	0.0702 (0.0896)	-0.246** (0.114)	-0.175 (0.127)	0.073 (0.091)	-0.278** (0.117)
Dummy de Pesq. Visitantes	0.108*** (0.0365)	0.112*** (0.0227)	0.101*** (0.0229)	0.120*** (0.037)	0.113*** (0.023)	0.099*** (0.023)	8.99e-05 (0.0244)	0.0368** (0.0159)	0.0433** (0.0173)	0.012 (0.027)	0.039** (0.016)	0.047** (0.019)
Relação Capital-Trabalho	1.195* (0.640)	-0.553 (0.446)	0.267 (0.334)	1.194* (0.609)	-0.504 (0.436)	0.166 (0.325)	0.355 (0.364)	-0.682* (0.398)	-0.0154 (0.224)	0.419 (0.382)	-0.674* (0.401)	-0.094 (0.238)
Idade do Laboratório	-0.217* (0.123)	0.0911 (0.0795)	-0.0539 (0.0617)	-0.220* (0.118)	0.084 (0.077)	-0.034 (0.060)	-0.0737 (0.0677)	0.121* (0.0729)	-0.00507 (0.0416)	-0.089 (0.072)	0.121* (0.073)	0.010 (0.044)
Idade do Laboratório ao Quadrado	-1.156 (1.104)	0.141 (0.706)	-0.775 (0.577)				0.406 (0.640)	0.803 (0.585)	-0.107 (0.398)			
Constante												
N	207	647	483	207	647	483	207	647	483	207	647	483
R2	0.194	0.102	0.168				0.698	0.546	0.571			
Pseudo R2				0.0879	0.0423	0.0753				0.4440	0.3097	0.3259

Fonte: Elaboração Própria.

Tabela 39 – Resultados das Regressões – Efeitos Marginais da Produtividade Científica dos coordenadores contra as características coletivas e individuais por *Cluster* ponderados pelo IFPA.

	yin/coautores					
	OLS			Tobit		
	<i>cluster 1</i>	<i>cluster 2</i>	<i>cluster 3</i>	<i>cluster 1</i>	<i>cluster 2</i>	<i>cluster 3</i>
Prod. Cient. do quinquênio anterior	0.777*** (0.0411)	0.699*** (0.0281)	0.687*** (0.0374)	0.812*** (0.045)	0.718*** (0.030)	0.695*** (0.039)
Prod. dos Pesquisadores	0.107*** (0.0382)	0.0767*** (0.0236)	0.0644*** (0.0218)	0.109*** (0.039)	0.077*** (0.024)	0.073*** (0.023)
Permanentemente no quinquênio	0.0114 (0.0468)	0.0111 (0.0353)	0.0321 (0.0379)	0.002 (0.049)	0.013 (0.036)	0.022 (0.041)
Pessoal Ocupado	1.572*** (0.501)	-0.0877 (0.284)	0.0561 (0.287)	1.697*** (0.509)	-0.096 (0.292)	0.106 (0.311)
Pesquisadores Permanentemente	0.938*** (0.321)	0.162 (0.195)	0.276 (0.210)	1.061*** (0.331)	0.183 (0.201)	0.334 (0.221)
Pesquisadores Não-Permanentemente	1.246*** (0.402)	-0.0624 (0.220)	0.0193 (0.222)	1.378*** (0.407)	-0.072 (0.226)	0.080 (0.244)
Não-Pesquisadores	-0.174 (0.119)	0.0668 (0.0877)	-0.253** (0.114)	-0.197 (0.124)	0.069 (0.089)	-0.283** (0.118)
Dummy de Pesq. Visitantes	-0.00554 (0.0255)	0.0356** (0.0161)	0.0433** (0.0175)	0.007 (0.027)	0.039** (0.016)	0.048** (0.019)
Relação Capital-Trabalho	0.406 (0.364)	-0.719* (0.405)	-0.0477 (0.224)	0.478 (0.375)	-0.700* (0.406)	-0.102 (0.238)
Idade do Laboratório	-0.0782 (0.0674)	0.127* (0.0739)	-0.00252 (0.0414)	-0.095 (0.070)	0.125* (0.074)	0.008 (0.044)
Idade do Laboratório ao Quadrado	-0.132 (0.205)	-0.189* (0.0975)	0.325*** (0.112)	-0.162 (0.193)	-0.219** (0.100)	0.303** (0.119)
Geração: 1962-1979	-0.166 (0.132)	0.0216 (0.0697)	0.0440 (0.0832)	-0.199 (0.136)	0.009 (0.071)	0.055 (0.086)
Geração: 1980-1989	-0.0377 (0.0772)	0.0455 (0.0544)	0.0768 (0.0530)	-0.060 (0.084)	0.036 (0.056)	0.074 (0.057)
Geração: 1990-1999	0.201 (0.150)	-0.174* (0.0995)	-0.0361 (0.0920)	0.218 (0.142)	-0.185* (0.103)	-0.059 (0.097)
Pesquisador de Ciências Exatas e da Terra	-0.0412 (0.192)	-0.0208 (0.103)	0.146 (0.0908)	-0.003 (0.190)	-0.008 (0.101)	0.186** (0.087)
Pesquisador de Engenharias	0.0801 (0.304)	0.0647 (0.104)	-0.0721 (0.0865)	0.062 (0.363)	0.041 (0.112)	-0.137 (0.102)
Pesquisador de Ciências Biológicas	-0.204* (0.120)	0.0983 (0.265)	0.00505 (0.153)	-0.108 (0.130)	0.108 (0.257)	-0.073 (0.189)
Pesquisador de Ciências da Saúde	-0.615*** (0.133)	0.0196 (0.0899)	0.119 (0.141)	-0.523*** (0.121)	0.037 (0.087)	0.155 (0.141)
Pesquisador de Ciências Agrárias	0.480 (0.659)	0.839 (0.585)	-0.153 (0.410)			
Constante						
N	207	647	483	207	647	483
R2	0.709	0.553	0.581			
Pseudo R2				0.4564	0.3154	0.3356

Fonte: Elaboração Própria.

7 CONCLUSÃO

Com uma infraestrutura de pesquisa quase integralmente criada no século XX, os recentes avanços da produção científica brasileira demonstram o potencial produtivo da infraestrutura de pesquisa nacional¹⁸, levando a uma melhora da posição brasileira nos rankings internacionais, principalmente, no que tange à quantidade artigos publicados em periódicos indexados internacionalmente. No entanto, pouco se tem sido investigado sobre o papel dessa infraestrutura de pesquisa sobre a produtividade científica, bem como sua inter-relação com a produção de outros serviços e desenvolvimento de tecnologias. Portanto, a fim de preencher essa lacuna, o presente trabalho teve como objetivo de pesquisa analisar os determinantes coletivos na produção dos cientistas brasileiros, especificamente entre aqueles que exercem a função de coordenador na infraestrutura. Além disso, buscou-se conhecer qual a configuração de laboratório de pesquisa seria a mais estimulante à produção científica no Brasil. Para isso, foram utilizadas duas bases de dados, uma oriunda do MCTI e outra extraída da Plataforma *Lattes*, as quais derivaram em 1.760 infraestruturas de pesquisa composta por 7.714 pesquisadores distintos. Sendo o público-alvo de 1.445 coordenadores ocupando 1.652 vagas de coordenação em 1.156 infraestruturas.

O objetivo de realizar-se estimações separadamente por características observáveis dos pesquisadores e dos laboratórios foi investigar, individualmente, os principais efeitos diferenciais das variáveis sobre a produtividade científica individual, além disso identificar quais os impactos específicos sobre as tipologias pré-definidas dos laboratórios. Os resultados econométrico obtidos por OLS e Tobit oriundos das equações (6, 8 e 9), indicam, de forma geral, que pertencer ao *cluster* tipo 2 (laboratórios de multiatividades) aumenta a produtividade dos coordenadores, enquanto que os pertencentes aos laboratórios do tipo 1 (laboratórios intensivos em pesquisa e desenvolvimento de tecnologia) e 3 (laboratórios intensivos em pesquisa) são menos produtivos.

¹⁸ Segundo CNPq por infraestrutura de pesquisa entende-se o conjunto de instalações físicas e condições materiais de apoio (equipamentos e recursos) utilizados pelos pesquisadores para a realização de atividades de P&D. São exemplos de infraestruturas de pesquisa: grandes instalações de pesquisa, laboratórios, redes integradas de instalações de P&D, plantas piloto, biotérios, salas limpas, redes de informática de alto desempenho, bases de dados, coleções, bibliotecas especializadas, observatórios, telescópios, navios de pesquisa, reservas e estações experimentais, entre outras.

Ademais, o caso brasileiro é fundamentado na produção discente, ou seja, há uma grande importância da participação dos pesquisadores não-permanentes (estudantes bolsistas) no aumento produtivo dos coordenadores. Os parâmetros estimados da contribuição da produção dos outros pesquisadores permanentes, também, produziram efeitos marginais positivos e significativos sobre a produtividade dos coordenadores. Quanto às características individuais, tem-se que a produtividade é altamente dependente da produção científica do coordenador no quinquênio anterior; da sua participação em mais de uma infraestrutura – inferindo que os cientistas mais produtivos são os que realizam mais atividades em mais de uma infraestrutura.

As principais diferenças das estimações quando pondera-se a produção dos coordenadores pelo índice IFPA é no que diz respeito aos efeitos geracionais e a relação capital-trabalho. Nas primeiras estimações percebe-se que as gerações com formação mais recente (4 a 13 anos de doutorado), em geral, são mais produtivas que as gerações mais antigas, porém quando avaliada pela qualidade no método OLS, nota-se uma maior produtividade dos coordenadores pertencentes à geração de 1990-1999. Ademais, apesar de pequeno, tem-se que aumentos nos parâmetros estimados da relação capital-trabalho contribuem de forma mais consistentes para elevações da produtividade dos coordenadores no que diz respeito às publicações internacionalmente reconhecidas.

Na avaliação por *cluster*, como resultados semelhantes, tem-se que, em geral, as produções médias dos pesquisadores permanentes presentes nas infraestruturas e a produção do coordenador no quinquênio anterior apresentam elasticidades positivas e significativas em todos os modelos e métodos estimados.

Mais especificamente, nos laboratórios intensivos em pesquisa e desenvolvimento de tecnologia (*cluster* 1), destaca-se o impacto expressivo da participação da mão de obra em todos os níveis – pesquisadores permanentes, não-permanentes e não-pesquisadores, principalmente quando se trata da produção indexada internacionalmente. Nota-se, também diferenças produtivas quanto às grandes áreas do conhecimento.

Já para os coordenadores pertencentes aos laboratórios de multiatividades (*cluster* 2), também observa-se os efeitos positivos oriundos dos coeficientes da variável *dummy* de presença de pesquisadores visitantes e de aumentos nos coeficientes da variável relação capital-trabalho. Porém, quando avalia-se a variável dependente pelo IFPA, os determinantes para a produção dos coordenadores desse grupo restringem-se apenas aos parâmetros estimados da relação capital-trabalho. Aqui, pode-se observar retornos marginais crescentes oriundos pelos coeficientes estimados da variável idade e idade ao quadrado da infraestrutura,

após a inclusão da variável dependente defasada. Quanto às *dummies* de gerações, nota-se que os coordenadores mais novos são mais produtivos que os demais. Observa-se, também, diferenças produtivas entre as áreas do conhecimento.

Quanto aos laboratórios intensivos em pesquisa (*cluster* 3), apresentam elasticidades positivas da participação expressiva de percentual de estudantes (pesquisadores não-permanentes) na produtividade científica dos coordenadores, além de um decréscimo com a presença de pesquisadores visitantes. Quando avalia-se as publicações indexadas internacionalmente, nota-se uma diminuição da produtividade com a presença de pesquisadores visitantes e elevação causada pela elasticidade da relação capital-trabalho, além das diferenças na produção em relação às grandes áreas do conhecimento. Quanto às *coortes* de geração, aqui os coordenadores de formação mais antiga possuem maiores níveis de produtividade.

Por fim, através dos exercícios propostos neste trabalho, foi possível estabelecer relações sobre os determinantes coletivos e individuais sobre a produtividade científica dos coordenadores pertencentes aos centros de pesquisa do Brasil, além de caracterizar tais centros em tipos de infraestruturas. Desta forma, foi possível estabelecer que as estruturas com maiores participações de pesquisadores não-permanentes (estudantes bolsistas), com estrutura organizacional mais piramidal, que possuem alta produtividade dos seus pesquisadores permanentes – àqueles ligados diretamente às atividades principais das infraestruturas – além de maiores relações capital-trabalho propiciam maiores níveis de produtividade dos seus pesquisadores chefes, ou seja, os coordenadores.

REFERÊNCIAS

ALLISON, Paul D.; LONG, J. Scott. Departmental effects on scientific productivity. **American sociological review**, p. 469-478, 1990.

BAYER, Alan E.; DUTTON, Jeffrey E. Career age and research-professional activities of academic scientists: Tests of alternative nonlinear models and some implications for higher education faculty policies. **The Journal of Higher Education**, p. 259-282, 1977.

BECKER, Gary. Human Capital: a theoretical and empirical analysis, with special reference to education. Nova Iorque: NBER/Columbia University Press, 1962.

BONACCORSI, A., DARAIO, C., The organization of science. Size, agglomeration and age effects in scientific productivity. In: **Proceedings of the Conference “Rethinking Science Policy: Analytical Frameworks for Evidence-based Policy”**, 2003.

CARAYOL, Nicolas; MATT, Mireille. Does research organization influence academic production?: Laboratory level evidence from a large European university. **Research Policy**, v. 33, n. 8, p. 1081-1102, 2004.

_____. Individual and collective determinants of academic scientists' productivity. **Information Economics and Policy**, v. 18, n. 1, p. 55-72, 2006.

COLE, Jonathan R.; COLE, Stephen. **Social stratification in science**. Chicago: University of Chicago Press, 1973. p. 298.

COLE, Stephen. Age and scientific performance. **American Journal of Sociology**, p. 958-977, 1979.

CNPq. **Conselho Nacional de Desenvolvimento Científico e Tecnológico**. Disponível em: <<http://www.cnpq.br/>>. Acesso em: 01 Fev 2015.

DIAMOND, Arthur M. The life-cycle research productivity of mathematicians and scientists. **Journal of Gerontology**, v. 41, n. 4, p. 520-525, 1986.

FOX, Mary Frank. Publication productivity among scientists: A critical review. **Social Studies of Science**, v. 13, n. 2, p. 285-305, 1983.

_____. Research, Teaching, and Publication Productivity: Mutuality Versus Competition in Academia. **Sociology of education**, v. 65, n. 4, p. 293-305, 1992.

_____. Gender, family characteristics, and publication productivity among scientists. **Social Studies of Science**, v. 35, n. 1, p. 131-150, 2005.

GLÄNZEL, W., Schubert, A. Double effort = double impact? A critical view at international co-authorship in chemistry. **Scientometrics**, 50, 199-214, 2001.

GONZALEZ-BRAMBILA, Claudia; VELOSO, Francisco M. The determinants of research output and impact: A study of Mexican researchers. **Research Policy**, v. 36, n. 7, p. 1035-1051, 2007.

GONZALEZ-BRAMBILA, Claudia N.; VELOSO, Francisco M.; KRACKHARDT, David. The impact of network embeddedness on research output. **Research Policy**, v. 42, n. 9, p. 1555-1567, 2013.

HORTA, Hugo; LACY, T. Austin. How does size matter for science? Exploring the effects of research unit size on academics' scientific productivity and information exchange behaviors. **Science and Public Policy**, v. 38, n. 6, p. 449-460, 2011.

JCR. **Jornal Citation Reports**. Disponível em: <<http://login.webofknowledge.com/>>. Acesso em: 01 Dez 2014.

KANNEBLEY Jr., Sérgio; CAROLO, Murilo Damião; DE NEGRI, Fernanda. Impacto dos Fundos Setoriais sobre a Produtividade Acadêmica de Cientistas Universitários. **Estudos Econômicos (São Paulo)**, v. 43, n. 4, 2014.

LEVIN, Sharon G.; STEPHAN, Paula E. Research productivity over the life cycle: Evidence for academic scientists. **The American Economic Review**, p. 114-132, 1991.

LI, Eldon Y.; LIAO, Chien Hsiang; YEN, Hsiuju Rebecca. Co-authorship networks and research impact: A social capital perspective. **Research Policy**, v. 42, n. 9, p. 1515-1530, 2013.

LOTKA, Alfred James. The frequency distribution of scientific productivity. **Journal of Washington Academy Sciences**, 1926.

MAIRESSE, Jacques; TURNER, Laure. **Measurement and explanation of the intensity of co-publication in scientific research: an analysis at the laboratory level**. National Bureau of Economic Research, 2005.

SOMBATSOMPOP, Narongrit; MARKPIN, Teerasak. Making an equality of ISI impact factors for different subject fields. **Journal of the American Society for Information Science and Technology**, v. 56, n. 7, p. 676-683, 2005.

MARTÍN-SEMPERE, María José; REY-ROCHA, Jesús; GARZÓN-GARCÍA, Belén. The effect of team consolidation on research collaboration and performance of scientists. Case study of Spanish university researchers in Geology. **Scientometrics**, v. 55, n. 3, p. 377-394, 2002.

MERTON, Robert King (Ed.). **Social theory and social structure**. Simon and Schuster, 1968.

NARIN, Francis; HAMILTON, Kimberly S.; OLIVASTRO, Dominic. The increasing linkage between US technology and public science. **Research Policy**, v. 26, n. 3, p. 317-330, 1997.

PRICE, Derek J. De Solla.; BEAVER, Donald. Collaboration in an invisible college. **American psychologist**, v. 21, n. 11, p. 1011, 1966.

PELZ, Donald Campbell; ANDREWS, Frank Meredith. **Scientists in organization: Productive climates for research and development**. J. Wiley, 1976.

QUINTAL, Guida Maria da Conceição Caldeira. **Análise de clusters aplicada ao sucesso/insucesso em matemática**. Universidade da Madeira, 2006.

RENCER, Alvin C. **Methods of multivariate analysis**. John Wiley & Sons, 2003.

ROACH, Michael; SAUERMAN, Henry. A taste for science? PhD scientists' academic orientation and self-selection into research careers in industry. **Research Policy**, v. 39, n. 3, p. 422-434, 2010.

SCHULTZ, Theodore W. Investment in human capital. *The American Economic Review*, v. LI, n. 1, p. 1-17, march.1961.

SQUEFF, FHS; DE NEGRI, F. Infraestrutura científica e tecnológica no Brasil: análises preliminares. **Brasília: Ipea, jun, 2014.**

STEPHAN, Paula E. The economics of science. **Journal of Economic literature**, p. 1199-1235, 1996.

_____. The economics of science. **Handbook of the economics of innovation**, v. 1, p. 217-274, 2010.

_____. **How economics shapes science.** Harvard University Press, 2012.

STEPHAN, Paula E. et al. Doctoral Education and Economic Development: The Flow of New PhDs to Industry. **Economic Development Quarterly**, v. 18, n. 2, p. 151-167, 2004.

STEPHAN, Paula E. et al. Who's patenting in the university? Evidence from the survey of doctorate recipients. **Econ. Innov. New Techn.**, v. 16, n. 2, p. 71-99, 2007.

TURNER, Laure; MAIRESSE, Jacques. Individual Productivity Differences in Public Research: How important are non-individual determinants? An Econometric Study of French Physicists' publications and citations (1986-1997). **Centre National de la Recherche Scientifique**, 2005.

WOOLDRIDGE, J. M. **Econometric analysis of cross section and panel data.** Cambridge: The MIT Press, 2002. 735p

WEISS, Yhoram, LILLARD, Lee. Output variability, academic labor contracts and waiting times for promotion. **Research in Labor Economics**, v. 5, p. 157-188, 1982.

XIE, Yu; SHAUMAN, Kimberlee A. Sex differences in research productivity: New evidence about an old puzzle. **American Sociological Review**, p. 847-870, 1998.

_____. **Women in science: Career processes and outcomes.** Cambridge, MA: Harvard University Press, 2003.

Apêndice A – Processos de validação dos clusters

rk2a			
rk2	1	2	Total
1		545	545
2	236	97	333
Total	236	642	878

rk2b			
rk2	1	2	Total
1	475	47	522
2		356	356
Total	475	403	878

rk3a				
rk3	1	2	3	Total
1		275	13	288
2	107		140	247
3		147	196	343
Total	107	422	349	878

rk3b				
rk3	1	2	3	Total
1	38	215	22	275
2	37		213	250
3	71	185	97	353
Total	146	400	332	878

rk4a					
rk4	1	2	3	4	Total
1	32	199		29	260
2	129			13	142
3	112	132		34	278
4			58	140	198
Total	273	331	58	216	878

rk4b					
rk4	1	2	3	4	Total
1	201		50	1	252
2	68	15	5	47	135
3	76		23	189	288
4		114	79	10	203
Total	345	129	157	247	878

rk5a						
rk5	1	2	3	4	5	Total
1	88		2		145	235
2		126		1		127
3		2		140		142
4	110	60	1		8	179
5	54	3	136		2	195
Total	252	191	139	141	155	878

rk5b						
rk5	1	2	3	4	5	Total
1			179	44	2	225
2	12	114				126
3	133					133
4		114	7	80	1	202
5		13	2	19	158	192
Total	145	241	188	143	161	878

lk2a			
lk2	1	2	Total
1	172		172
2		706	706
Total	172	706	878

lk2b			
lk2	1	2	Total
1	45	106	151
2	227	500	727
Total	272	606	878

lk3a				
lk3	1	2	3	Total
1	451		77	528
2		142		142
3		2	206	208
Total	451	144	283	878

lk3b				
lk3	1	2	3	Total
1		72	458	530
2		36	99	135
3	17	196		213
Total	17	304	557	878

lk4a					
lk4	1	2	3	4	Total
1	112			83	195
2	96	424			520
3			71		71
4	16		71	5	92
Total	224	424	142	88	878

lk4b					
lk4	1	2	3	4	Total
1	2	76	121		199
2	179		85	255	519
3	44		1	27	72
4	15	13	34	26	88
Total	240	89	241	308	878

lk5a						
lk5	1	2	3	4	5	Total
1	16	2	250			268
2		38			86	124
3	153	165	19			337
4				54		54
5		4		89	2	95
Total	169	209	269	143	88	878

lk5b						
lk5	1	2	3	4	5	Total
1	201	20			50	271
2			14	111		125
3	2	146		44	151	343
4	15	33			8	56
5	18	11		24	30	83
Total	236	210	14	179	239	878

fk2a			
fk2	1	2	Total
1	166	61	227
2	6	645	651
Total	172	706	878

fk2b			
fk2	1	2	Total
1	136	70	206
2	339	333	672
Total	475	403	878

fk3a				
fk3	1	2	3	Total
1	144			144
2		289		289
3			445	445
Total	144	289	445	878

fk3b				
fk3	1	2	3	Total
1	131			131
2	2	245	69	316
3			431	431
Total	133	245	500	878

fk4a					
fk4	1	2	3	4	Total
1	68	7	20	4	99
2		141		139	280
3		88	337		425
4	74				74
Total	142	236	357	143	878

fk4b					
fk4	1	2	3	4	Total
1	58	5	2	33	98
2	2	244	56	2	304
3			337	78	415
4	61				61
Total	121	249	395	113	878

fk5a						
fk5	1	2	3	4	5	Total
1	47			2		49
2		145	12	53		210
3		11	340	3		354
4	94			2	1	97
5		1		110	57	168
Total	141	157	352	170	58	878

fk5b						
fk5	1	2	3	4	5	Total
1	41					41
2		11	9	3	206	229
3			234	68	34	336
4	82	1		14		97
5		159			16	175
Total	123	171	243	85	256	878

ek2a			
ek2	1	2	Total
1	162	298	460
2	171	247	418
Total	333	545	878

ek2b			
ek2	1	2	Total
1	247	225	472
2	203	203	406
Total	450	428	878

ek3a				
ek3	1	2	3	Total
1	226	39		265
2	60		460	520
3		93		93
Total	286	132	460	878

ek3b				
ek3	1	2	3	Total
1	5	15	283	303
2	346	130	16	492
3			83	83
Total	351	145	382	878

ek4a					
ek4	1	2	3	4	Total
1	84		5	80	169
2	211		31	18	260
3		142	2	59	203
4	1		212	33	246
Total	296	142	250	190	878

ek4b					
ek4	1	2	3	4	Total
1	31	10	132	6	179
2	4	189	47	16	256
3	202				202
4	47		5	189	241
Total	284	199	184	211	878

ek5a						
ek5	1	2	3	4	5	Total
1		12			198	210
2		192	35			227
3		41			176	217
4		20			146	166
5	2		49	7		58
Total	2	265	84	7	520	878

ek5b						
ek5	1	2	3	4	5	Total
1	9		114	76		199
2	45	169			26	240
3	67	25	127			219
4	34	2	110	20		166
5					54	54
Total	155	196	351	96	80	878

sk2a			
sk2	1	2	Total
1	176	242	418
2	154	306	460
Total	330	548	878

sk2b			
sk2	1	2	Total
1	203	203	406
2	247	225	472
Total	450	428	878

sk3a				
sk3	1	2	3	Total
1	338	21	28	387
2		298		298
3	11		182	193
Total	349	319	210	878

sk3b				
sk3	1	2	3	Total
1	173	188	43	404
2	114	103	93	310
3	31	11	122	164
Total	318	302	258	878

sk4a					
sk4	1	2	3	4	Total
1		99	21	138	258
2	131			77	208
3		2	199	16	217
4		187	1	7	195
Total	131	288	221	238	878

sk4b					
sk4	1	2	3	4	Total
1	43	117	94	4	258
2	10	18		181	209
3	186	1	27		214
4		24	173		197
Total	239	160	294	185	878

sk5a						
sk5	1	2	3	4	5	Total
1	196			12	4	212
2	3	118				121
3	1	3	141			145
4	8			10	152	170
5	3			223	4	230
Total	211	121	141	245	160	878

sk5b						
sk5	1	2	3	4	5	Total
1	35		61	11	135	242
2		6			112	118
3		139				139
4	140		3	11		154
5	1		86	138		225
Total	176	145	150	160	247	878

Apêndice B – Teste MVREG.

Equation	Obs	Parms	RMSE	"R-sq"	F	P
perc_dedicacao_exclusiva	1.756	3	0,4323	0,0131	11,6366	0,0000
perc_doutorado	1.756	3	0,3027	0,0761	72,1910	0,0000
perc_mestrado	1.756	3	0,2102	0,0435	39,8858	0,0000
perc_estud_doutorado	1.756	3	4,4759	0,3586	490,0773	0,0000
perc_pesq_visit	1.756	3	0,0996	0,0052	4,6175	0,0100
perc_pesq_estrang	1.756	3	0,1402	0,0070	6,1399	0,0022
IUCF	1.756	3	1,6915	0,0038	3,3084	0,0368
grande_area_	1.756	3	1,7223	0,7126	2.173,2630	0,0000
sem_cooperacao	1.756	3	0,3077	0,0593	55,2053	0,0000
ativ_ensino_	1.756	3	0,9117	0,0058	5,1198	0,0061
ativ_pesq_	1.756	3	0,5004	0,0370	33,7186	0,0000
ativ_tecno_	1.756	3	1,0615	0,0013	1,1642	0,3124
ativ_ext_tecno_	1.756	3	1,3097	0,0012	1,0224	0,3600
ativ_serv_tecno_	1.756	3	1,2517	0,0044	3,8643	0,0212
ativ_outra_	1.756	3	0,3501	0,0006	0,5224	0,5932

		Coef.	Std. Err.	t	P> t	[95% Interv. Conf.]	
perc_dedicacao_exclusiva	2	-0,0099	0,0314	-0,3100	0,7530	-0,0715	0,0518
	3	-0,1061	0,0299	-3,5500	0,0000	-0,1647	-0,0475
	_cons	0,5868	0,0261	22,5100	0,0000	0,5356	0,6379
perc_doutorado	2	0,1050	0,0220	4,7700	0,0000	0,0619	0,1482
	3	-0,0871	0,0209	-4,1600	0,0000	-0,1281	-0,0461
	_cons	0,8051	0,0183	44,1100	0,0000	0,7693	0,8409
perc_mestrado	2	-0,0710	0,0153	-4,6400	0,0000	-0,1010	-0,0410
	3	0,0277	0,0145	1,9100	0,0560	-0,0008	0,0562
	_cons	0,1333	0,0127	10,5200	0,0000	0,1084	0,1581
perc_estud_doutorado	2	5,7881	0,3255	17,7800	0,0000	5,1497	6,4265
	3	-1,5253	0,3094	-4,9300	0,0000	-2,1321	-0,9185
	_cons	2,2582	0,2699	8,3700	0,0000	1,7288	2,7876
perc_pesq_visit	2	-0,0078	0,0072	-1,0700	0,2840	-0,0220	0,0064
	3	-0,0189	0,0069	-2,7400	0,0060	-0,0324	-0,0054
	_cons	0,0356	0,0060	5,9200	0,0000	0,0238	0,0474
perc_pesq_estrang	2	0,0057	0,0102	0,5600	0,5780	-0,0143	0,0257
	3	-0,0193	0,0097	-1,9900	0,0470	-0,0383	-0,0002
	_cons	0,0646	0,0085	7,6400	0,0000	0,0480	0,0812
IUCF	2	-0,1097	0,1230	-0,8900	0,3730	-0,3510	0,1316
	3	0,1194	0,1169	1,0200	0,3080	-0,1100	0,3487
	_cons	9,2213	0,1020	90,4000	0,0000	9,0213	9,4214

(Continua.)

(Conclusão)

		Coef.	Std. Err.	t	P> t	[95% Interv. Conf.]	
grande_area_	2	-7,1468	0,1253	-57,0600	0,0000	-7,3925	-6,9011
	3	-7,6288	0,1191	-64,0800	0,0000	-7,8623	-7,3953
	_cons	10,2509	0,1039	98,7000	0,0000	10,0472	10,4546
sem_cooperacao	2	-0,0407	0,0224	-1,8200	0,0690	-0,0846	0,0032
	3	0,1240	0,0213	5,8300	0,0000	0,0823	0,1658
	_cons	0,0655	0,0186	3,5300	0,0000	0,0291	0,1018
ativ_ensino_	2	-0,1650	0,0663	-2,4900	0,0130	-0,2950	-0,0349
	3	-0,0258	0,0630	-0,4100	0,6820	-0,1494	0,0978
	_cons	1,5418	0,0550	28,0400	0,0000	1,4340	1,6496
ativ_pesq_	fk3						
	2	-0,0883	0,0364	-2,4300	0,0150	-0,1596	-0,0169
	3	0,1267	0,0346	3,6600	0,0000	0,0589	0,1945
	_cons	1,0982	0,0302	36,3900	0,0000	1,0390	1,1574
ativ_tecno_	fk3						
	2	-0,1157	0,0772	-1,5000	0,1340	-0,2671	0,0357
	3	-0,0650	0,0734	-0,8900	0,3750	-0,2089	0,0789
	_cons	1,2545	0,0640	19,6000	0,0000	1,1290	1,3801
ativ_ext_tecno_	fk3						
	2	-0,0516	0,0953	-0,5400	0,5880	-0,2384	0,1353
	3	-0,1182	0,0905	-1,3100	0,1920	-0,2957	0,0594
	_cons	1,1673	0,0790	14,7800	0,0000	1,0124	1,3222
ativ_serv_tecno_	2	-0,0648	0,0910	-0,7100	0,4770	-0,2433	0,1137
	3	-0,2051	0,0865	-2,3700	0,0180	-0,3748	-0,0354
	_cons	1,3855	0,0755	18,3600	0,0000	1,2374	1,5335
ativ_outra_	2	-0,0212	0,0255	-0,8300	0,4060	-0,0711	0,0288
	3	-0,0245	0,0242	-1,0100	0,3120	-0,0719	0,0230
	_cons	0,0873	0,0211	4,1300	0,0000	0,0459	0,1287

Quadro 4 – Teste MVREG.

Fonte: *Software Stata*.

Apêndice C – *Cluster* Hierárquico

A construção de um *cluster* hierárquico aglomerativo¹⁹ tem como vantagem a não necessidade de um conhecimento *a priori* do número de *clusters*. Entretanto, sua desvantagem ocorre que uma vez feita a alocação de um objeto a determinado grupo, o mesmo não pode mais ser realocado para outro. A escolha da distância entre dois grupos foi pelo *Simple Linkage* (Ligação Simples ou Critério do vizinho mais próximo), de tal maneira que usa como ponto de partida uma matriz de dessemelhança e define como distância mais próxima, entre quaisquer dois grupos A e B:

$$d_{AB} = \min \{d_{ij} : i \in A, j \in B\}$$

A partir de então, foi construído uma matriz de similaridade $\mathbf{R} = (r_{ij})$ de tal forma que se utilizou as correlações entre cada par de variáveis como medidas de proximidade. As variáveis com correlações altamente negativas são consideradas como muito dissimilares e as altamente positivas como bastantes similares (Tabela 23).

O dendrograma é uma representação matemática do procedimento de construção de um *cluster* completo. Seu desenho demonstra a árvore e o valor da medida de proximidade em cada passo do processo hierárquico, de tal maneira que os nós representam os *clusters* e o tamanho do tronco (altura) representa a distância que os mesmos se unem. O algoritmo utilizado para a construção do dendrograma exige que a matriz de dados seja uma matriz de dessemelhança, portanto se faz necessário a conversão da matriz de similaridade (correlação) $\mathbf{R} = (r_{ij})$ para a matriz de dissimilaridade $\mathbf{D} = d_{ij}$ de tal forma que:

$$d_{ij} = 1 - |r_{ij}|$$

¹⁹ Nos algoritmos aglomerativos são os métodos hierárquicos mais usados, o qual parte de n grupos com um objetivo em cada e estes vão sendo agrupados sucessivamente até culminar com a construção de um grupo que inclui a totalidade dos objetos (processo de fusão sucessiva de grupos).

Gráfico 3 – Ligação Simples para dissimilaridade entre as 15 variáveis utilizadas na construção do *cluster*.

Fonte: Elaboração própria.

A análise derivada do dendrograma, por meio das relações de dissimilaridade, revela que existem três principais agrupamentos de variáveis. O primeiro subconjunto (2, 4, 7 e 9) reúne as variáveis de insumos (capital humano e físico), o segundo (14, 13, 12, 11, 10 e 3) é composto pelas variáveis que compõem os produtos desenvolvidos pelas infraestruturas (atividades) e o terceiro (5 e 6) é caracterizado por variáveis de cooperação externa (pesquisadores estrangeiros e visitantes). Por esses agrupamentos não é possível realizar uma associação entre os três subconjuntos, ou seja, não espera-se que existam associações significativas entre eles, mas que as tipologias dos *clusters não-hierárquicos* formados pelas infraestruturas sejam definidas pelas diferenças dentro de cada subconjunto de variáveis.

De maneira analítica, tem-se que as variáveis 13 (Atividade de Extensão Tecnológica) e 14 (Atividades de Serviço Técnico) possuem a maior similaridade e estão relacionadas – com maiores dissimilaridades – com as variáveis 12 (Atividades de Desenvolvimento de Tecnologia), 11 (Atividades de Pesquisa) e 10 (Atividades de Ensino). Tal fato indica um encadeamento lógico entre as atividades realizadas pelas infraestruturas. Neste caso, deve-se esperar na formação dos *clusters* que as interações das intensidades do desenvolvimento

dessas atividades possam melhor caracterizá-los. A ligação com o segundo subconjunto de variáveis ocorre por meio da variável 3 (Percentual de Pesquisadores com Mestrado) com a variável 9 (Realização de Cooperação). Nesse subconjunto, a maior similaridade ocorre entre as variáveis 2 (Percentual de Pesquisadores com Doutorado) e 4 (Percentual de Estudantes de Doutorado), que demonstra uma provável relação de orientador-orientando. Assim, mais uma vez deve-se esperar que essas variáveis apresentem comportamentos associativos dentro dos grupos de variáveis, porém varie entre os *clusters*.

Os dois primeiros subconjuntos se unem com o terceiro – variáveis 5 (Percentual de Pesquisadores Visitantes) e 6 (Percentual de Pesquisadores Estrangeiros) – por meio da variável de realização de cooperação (9), o que demonstra uma ligação coerente. Por fim, os três subconjuntos de variáveis se interligam com menor associação com as variáveis 1 (Percentual de Pesquisadores de Dedicação Exclusiva), 8 (Grande Área do conhecimento) e 15 (Outras Atividades), as quais não guardam relação com os demais subconjuntos de variáveis, podendo ainda variar entre *clusters*.

Tabela 40 – Matriz de similaridade usando como medida de proximidade a correlação entre os pares de variáveis utilizadas na criação do *cluster*.

No. Variável	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	1,000														
2	-0,075	1,000													
3	0,055	-0,808	1,000												
4	0,041	0,225	-0,182	1,000											
5	-0,063	0,069	-0,049	0,026	1,000										
6	0,037	0,087	-0,065	0,070	0,131	1,000									
7	0,034	0,188	-0,169	-0,038	-0,072	0,024	1,000								
8	0,036	0,038	-0,015	0,033	0,028	-0,028	-0,018	1,000							
9	-0,108	-0,145	0,072	-0,157	-0,068	-0,068	0,123	-0,046	1,000						
10	-0,026	0,027	-0,014	0,003	-0,010	-0,005	0,054	0,049	0,067	1,000					
11	-0,034	-0,182	0,174	-0,138	-0,015	-0,070	-0,002	-0,026	0,119	0,108	1,000				
12	0,039	-0,014	0,030	-0,022	-0,018	0,057	-0,025	0,038	-0,106	0,183	0,224	1,000			
13	-0,011	0,018	0,000	0,020	-0,015	-0,037	-0,059	0,054	-0,085	0,207	0,083	0,376	1,000		
14	0,005	0,006	0,010	0,015	-0,026	-0,001	-0,056	0,048	-0,142	0,107	0,064	0,362	0,459	1,000	
15	-0,004	-0,014	0,022	0,025	0,015	0,001	-0,007	0,027	-0,018	-0,001	-0,056	0,042	-0,005	0,003	1,000

Fonte: Elaboração própria.

Apêndice D - Modelo de Produtividade Científica Individual com variáveis absolutas

Tabela 41 – Resultados das Regressões – Efeitos Marginais da Produtividade Científica Individual contra as características coletivas e individuais (variáveis absolutas).

	y _{in} /coautores					
	OLS	Tobit	OLS	Tobit	OLS	Tobit
Prod. Cient. do quinquênio anterior			0.697*** (0.0172)	0.716*** (0.019)	0.703*** (0.0196)	0.723*** (0.021)
Prod. dos Pesquisadores	0.189*** (0.0213)	0.197*** (0.022)	0.0685*** (0.0149)	0.071*** (0.016)	0.0669*** (0.0148)	0.069*** (0.015)
Permanentes no quinquênio	-0.0808** (0.0369)	-0.084** (0.039)	0.0253 (0.0256)	0.026 (0.027)	0.0329 (0.0255)	0.034 (0.027)
Pesquisadores Permanentes	0.121*** (0.0266)	0.127*** (0.027)	0.0530*** (0.0192)	0.056*** (0.020)	0.0553*** (0.0194)	0.058*** (0.020)
Pesquisadores Não-Permanentes	-0.00822 (0.0161)	-0.009 (0.017)	-0.0106 (0.0113)	-0.011 (0.012)	-0.0163 (0.0113)	-0.017 (0.012)
Não-Pesquisadores	0.0595 (0.0769)	0.064 (0.079)	0.0207 (0.0572)	0.025 (0.059)	0.0428 (0.0565)	0.048 (0.058)
Dummy de Pesq. Visitantes	0.0135 (0.0170)	0.013 (0.018)	0.000234 (0.0117)	-0.001 (0.012)	0.00568 (0.0118)	0.005 (0.012)
Relação Capital-Trabalho	0.178 (0.214)	0.171 (0.222)	-0.0208 (0.145)	-0.032 (0.151)	0.0320 (0.149)	0.032 (0.154)
Idade do Laboratório	-0.0224 (0.0387)	-0.021 (0.040)	-0.000232 (0.0263)	0.001 (0.027)	-0.00785 (0.0268)	-0.008 (0.028)
Idade do Laboratório ao Quadrado	-0.276*** (0.0613)	-0.286*** (0.064)	-0.0625 (0.0429)	-0.064 (0.045)	-0.0759* (0.0430)	-0.079* (0.045)
Cluster 1	-0.449*** (0.0471)	-0.467*** (0.049)	-0.135*** (0.0338)	-0.140*** (0.035)	-0.134*** (0.0337)	-0.140*** (0.035)
Cluster 3					0.158*** (0.0353)	0.165*** (0.036)
Trabalha em mais de 1 Infraest.					-0.182** (0.0733)	-0.195*** (0.075)
Geração: 1962-1979					-0.0786 (0.0481)	-0.088* (0.050)
Geração: 1980-1989					-0.00466 (0.0349)	-0.012 (0.037)
Geração: 1990-1999					-0.194*** (0.0664)	-0.192*** (0.068)
Pesquisador da área de Ciências Exatas e da Terra					0.0199 (0.0671)	0.031 (0.067)
Pesquisador da área de Engenharias					-0.0379 (0.0647)	-0.063 (0.071)
Pesquisador da área de Ciências Biológicas					0.000477 (0.102)	-0.024 (0.115)
Pesquisador da área de Ciências da Saúde					0.0260 (0.107)	0.026 (0.109)
Pesquisador da área de Ciências Agrárias						
Constante	0.753** (0.354)		0.427* (0.240)		0.277 (0.244)	
N	1.652	1.652	1.652	1.652	1.652	1.652
R2	0.139		0.589		0.598	
Pseudo R2		0.0570		0.3362		0.3438

Fonte: Elaboração própria.

Tabela 42 – Resultados das Regressões – Efeitos Marginais da Produtividade Científica Individual contra as características coletivas por *Cluster* (variáveis absolutas).

	$y_{in}/coautores$											
	OLS			Tobit			OLS			Tobit		
	<i>cluster 1</i>	<i>cluster 2</i>	<i>cluster 3</i>	<i>cluster 1</i>	<i>cluster 2</i>	<i>cluster 3</i>	<i>cluster 1</i>	<i>cluster 2</i>	<i>cluster 3</i>	<i>cluster 1</i>	<i>cluster 2</i>	<i>cluster 3</i>
Prod. Cient. do quinquênio anterior							0.687*** (0.0473)	0.700*** (0.0290)	0.698*** (0.0273)	0.708*** (0.050)	0.710*** (0.030)	0.732*** (0.031)
Prod. dos Pesquisadores	0.213*** (0.0625)	0.196*** (0.0350)	0.173*** (0.0334)	0.220*** (0.065)	0.197*** (0.035)	0.185*** (0.036)	0.0820* (0.0484)	0.064*** (0.0225)	0.067*** (0.0247)	0.082* (0.049)	0.063*** (0.023)	0.073*** (0.026)
Permanentes no quinquênio	0.00281 (0.0961)	-0.169*** (0.0565)	-0.00312 (0.0729)	-0.001 (0.099)	-0.172*** (0.058)	-0.008 (0.078)	0.136** (0.0562)	-0.0269 (0.0403)	0.0595 (0.0486)	0.141** (0.058)	-0.028 (0.041)	0.059 (0.052)
Pesquisadores Permanentes	0.155** (0.0711)	0.0618* (0.0354)	0.243*** (0.0576)	0.168** (0.072)	0.063* (0.036)	0.259*** (0.060)	0.0510 (0.0581)	0.0325 (0.0238)	0.0957** (0.0462)	0.057 (0.058)	0.033 (0.024)	0.103** (0.047)
Pesquisadores Não-Permanentes	-0.0559 (0.0506)	0.0176 (0.0219)	-0.0225 (0.0320)	-0.051 (0.052)	0.017 (0.022)	-0.027 (0.034)	-0.0122 (0.0388)	-0.00214 (0.0156)	-0.0115 (0.0210)	-0.008 (0.040)	-0.003 (0.016)	-0.013 (0.022)
Não-Pesquisadores	-0.136 (0.204)	0.193** (0.0966)	-0.109 (0.170)	-0.155 (0.212)	0.199** (0.096)	-0.097 (0.177)	-0.129 (0.147)	0.125* (0.0714)	-0.0881 (0.132)	-0.138 (0.152)	0.129* (0.071)	-0.079 (0.136)
Dummy de Pesq. Visitantes	-0.0339 (0.0436)	0.072*** (0.0262)	-0.0213 (0.0297)	-0.043 (0.046)	0.073*** (0.027)	-0.023 (0.032)	-0.0589* (0.0322)	0.0118 (0.0181)	-0.00168 (0.0200)	-0.068** (0.034)	0.012 (0.018)	-0.003 (0.021)
Relação Capital-Trabalho	1.779** (0.767)	-0.225 (0.306)	0.433 (0.400)	1.928** (0.817)	-0.232 (0.309)	0.422 (0.424)	0.680 (0.487)	-0.292 (0.208)	0.221 (0.260)	0.795 (0.516)	-0.304 (0.210)	0.209 (0.276)
Idade do Laboratório	-0.331** (0.146)	0.0510 (0.0516)	-0.0717 (0.0762)	-0.361** (0.157)	0.053 (0.052)	-0.069 (0.081)	-0.145 (0.0913)	0.0537 (0.0352)	-0.0491 (0.0489)	-0.169* (0.098)	0.056 (0.035)	-0.048 (0.052)
Idade do Laboratório ao Quadrado	-0.909 (1.095)	0.539 (0.548)	0.432 (0.616)				0.230 (0.661)	0.640* (0.372)	0.00870 (0.415)			
Constante												
N	207	647	483	207	647	483	207	647	483	207	647	483
R2	0.122	0.085	0.085				0.567	0.571	0.557			
Pseudo R2				0.0497	0.0362	0.0338				0.3109	0.3463	0.3023

Fonte: Elaboração própria.

Tabela 43 – Resultados das Regressões – Efeitos Marginais da Produtividade Científica Individual contra as características coletivas e individuais por *Cluster* (variáveis absolutas).

	yin/coautores					
	OLS			Tobit		
	<i>cluster 1</i>	<i>cluster 2</i>	<i>cluster 3</i>	<i>cluster 1</i>	<i>cluster 2</i>	<i>cluster 3</i>
Prod. Cient. do quinquênio anterior	0.723*** (0.0516)	0.716*** (0.0331)	0.693*** (0.0317)	0.745*** (0.054)	0.727*** (0.034)	0.725*** (0.035)
Prod. dos Pesquisadores	0.0566 (0.0493)	0.0579** (0.0226)	0.0680*** (0.0252)	0.056 (0.049)	0.057** (0.023)	0.074*** (0.026)
Permanentes no quinquênio	0.136** (0.0582)	-0.0212 (0.0403)	0.0657 (0.0500)	0.140** (0.059)	-0.022 (0.041)	0.064 (0.053)
Pesquisadores Permanentes	0.0490 (0.0608)	0.0310 (0.0239)	0.0928** (0.0462)	0.055 (0.060)	0.031 (0.024)	0.097** (0.047)
Pesquisadores Não-Permanentes	-0.0133 (0.0404)	0.00124 (0.0161)	-0.0216 (0.0213)	-0.009 (0.041)	0.001 (0.016)	-0.024 (0.023)
Não-Pesquisadores	-0.134 (0.152)	0.126* (0.0706)	-0.0415 (0.130)	-0.144 (0.154)	0.131* (0.070)	-0.031 (0.134)
Dummy de Pesq. Visitantes	-0.0559* (0.0333)	0.0109 (0.0183)	0.00686 (0.0198)	-0.066* (0.035)	0.011 (0.018)	0.007 (0.021)
Relação Capital-Trabalho	0.733 (0.499)	-0.286 (0.204)	0.240 (0.277)	0.844 (0.518)	-0.289 (0.205)	0.247 (0.290)
Idade do Laboratório	-0.150 (0.0929)	0.0531 (0.0343)	-0.0520 (0.0517)	-0.173* (0.097)	0.054 (0.034)	-0.054 (0.054)
Idade do Laboratório ao Quadrado	-0.0474 (0.153)	-0.267** (0.109)	0.00238 (0.126)	-0.051 (0.156)	-0.282** (0.111)	-0.005 (0.125)
Geração: 1962-1979	-0.238 (0.165)	-0.0808 (0.0669)	-0.0195 (0.100)	-0.242 (0.167)	-0.087 (0.067)	-0.031 (0.104)
Geração: 1980-1989	-0.0730 (0.0959)	0.00786 (0.0528)	-0.00393 (0.0639)	-0.074 (0.097)	0.001 (0.054)	-0.018 (0.068)
Geração: 1990-1999	0.263 (0.164)	-0.207* (0.108)	-0.234** (0.0992)	0.299* (0.165)	-0.201* (0.107)	-0.238** (0.104)
Pesquisador de Ciências Exatas e da Terra	0.333** (0.160)	-0.162* (0.0931)	0.209** (0.0946)	0.352** (0.159)	-0.157* (0.092)	0.231** (0.094)
Pesquisador de Engenharias	-0.274 (0.181)	0.0610 (0.0986)	-0.0901 (0.0948)	-0.226 (0.193)	0.059 (0.100)	-0.158 (0.114)
Pesquisador de Ciências Biológicas	-0.392*** (0.139)	0.0412 (0.152)	-0.170 (0.208)	-0.391*** (0.138)	0.043 (0.149)	-0.305 (0.312)
Pesquisador de Ciências da Saúde	-0.669*** (0.218)	0.0374 (0.0996)	0.0809 (0.232)	-0.651*** (0.210)	0.039 (0.099)	0.070 (0.240)
Pesquisador de Ciências Agrárias	0.142 (0.672)	0.654* (0.373)	-0.114 (0.427)			
Constante						
N	207	647	483	207	647	483
R2	0.585	0.581	0.565			
Pseudo R2						

Fonte: Elaboração própria.

Apêndice E – Modelo de Produtividade Científica Individual ponderada pelo IFPA, com variáveis absolutas

Tabela 44 – Resultados das Regressões – Efeitos Marginais da Produtividade Científica Individual contra as características coletivas e individuais ponderados pelo IFPA (variáveis absolutas).

	y _{in} /coautores*IFPA					
	OLS	Tobit	OLS	Tobit	OLS	Tobit
Prod. Cient. do quinquênio anterior			0.719*** (0.0164)	0.739*** (0.018)	0.709*** (0.0180)	0.729*** (0.019)
Prod. dos Pesquisadores Permanentes no quinquênio	0.195*** (0.0199)	0.194*** (0.020)	0.0812*** (0.0137)	0.087*** (0.015)	0.0828*** (0.0139)	0.089*** (0.015)
Pesquisadores Permanentes	-0.0972*** (0.0354)	-0.104*** (0.035)	-0.0154 (0.0239)	-0.024 (0.026)	-0.00977 (0.0239)	-0.018 (0.025)
Pesquisadores Não-Permanentes	0.0991*** (0.0266)	0.104*** (0.026)	0.0361* (0.0195)	0.044** (0.020)	0.0408** (0.0196)	0.049** (0.020)
Não-Pesquisadores	-0.0305** (0.0154)	-0.032** (0.015)	-0.00675 (0.0110)	-0.009 (0.012)	-0.0111 (0.0111)	-0.014 (0.012)
Dummy de Pesq. Visitantes	-0.0547 (0.0720)	-0.073 (0.073)	-0.0372 (0.0587)	-0.055 (0.063)	-0.0312 (0.0585)	-0.046 (0.063)
Relação Capital-Trabalho	0.102*** (0.0142)	0.105*** (0.014)	0.0317*** (0.00994)	0.038*** (0.011)	0.0344*** (0.0101)	0.042*** (0.011)
Idade do Laboratório	-0.0869 (0.265)	-0.101 (0.252)	-0.263 (0.231)	-0.291 (0.234)	-0.261 (0.242)	-0.274 (0.245)
Idade do Laboratório ao Quadrado	0.0126 (0.0490)	0.015 (0.047)	0.0459 (0.0437)	0.051 (0.044)	0.0452 (0.0455)	0.048 (0.046)
Cluster 1	-0.306*** (0.0593)	-0.298*** (0.055)	-0.0962** (0.0384)	-0.111*** (0.041)	-0.0995** (0.0387)	-0.116*** (0.041)
Cluster 3	-0.396*** (0.0454)	-0.397*** (0.044)	-0.113*** (0.0328)	-0.134*** (0.035)	-0.113*** (0.0329)	-0.133*** (0.035)
Trabalha em mais de 1 Infraest.					0.149*** (0.0437)	0.161*** (0.044)
Geração: 1962-1979					-0.0353 (0.0678)	-0.056 (0.070)
Geração: 1980-1989					0.00635 (0.0465)	-0.002 (0.049)
Geração: 1990-1999					0.0608* (0.0328)	0.055 (0.035)
Pesquisador da área de Ciências Exatas e da Terra					-0.0709 (0.0635)	-0.088 (0.067)
Pesquisador da área de Engenharias					0.0411 (0.0675)	0.072 (0.067)
Pesquisador da área de Ciências					0.0107 (0.0669)	-0.033 (0.077)
Pesquisador da área de Ciências da					0.151 (0.157)	0.144 (0.158)
Pesquisador da área de Ciências Agrárias					0.0830 (0.0765)	0.117 (0.077)
Constante	-0.450 (0.395)		0.204 (0.327)		0.130 (0.339)	
N	1.652	1.652	1.652	1.652	1.652	1.652
R2	0.170		0.601		0.606	
Pseudo R2		0.0729		0.3356		0.3406

Fonte: Elaboração própria.

Tabela 45 – Resultados das Regressões – Efeitos Marginais da Produtividade Científica Individual contra as características coletivas por *Cluster*, ponderados pelo IFPA (variáveis absolutas).

	$y_{in}/\text{coautores*IFPA}$											
	OLS			Tobit			OLS			Tobit		
	<i>cluster 1</i>	<i>cluster 2</i>	<i>cluster 3</i>	<i>cluster 1</i>	<i>cluster 2</i>	<i>cluster 3</i>	<i>cluster 1</i>	<i>cluster 2</i>	<i>cluster 3</i>	<i>cluster 1</i>	<i>cluster 2</i>	<i>cluster 3</i>
Prod. Cient. do quinquênio anterior							0.779*** (0.0381)	0.694*** (0.0263)	0.713*** (0.0336)	0.806*** (0.042)	0.712*** (0.028)	0.719*** (0.036)
Prod. dos Pesquisadores	0.230*** (0.0537)	0.195*** (0.0357)	0.177*** (0.0299)	0.218*** (0.051)	0.191*** (0.035)	0.172*** (0.029)	0.0971** (0.0390)	0.0748*** (0.0235)	0.0655*** (0.0207)	0.098** (0.041)	0.074*** (0.024)	0.072*** (0.022)
Permanentes no quinquênio	-0.146* (0.0883)	-0.126** (0.0574)	-0.00395 (0.0595)	-0.170* (0.087)	-0.126** (0.057)	-0.019 (0.057)	0.0394 (0.0463)	-0.0303 (0.0384)	0.000133 (0.0424)	0.018 (0.052)	-0.032 (0.040)	-0.013 (0.044)
Pesquisadores Permanentes	0.112* (0.0655)	0.0772** (0.0356)	0.180*** (0.0582)	0.112* (0.061)	0.083** (0.035)	0.180*** (0.053)	0.0228 (0.0483)	0.0513* (0.0265)	0.0282 (0.0420)	0.027 (0.050)	0.059** (0.027)	0.038 (0.042)
Pesquisadores Não-Permanentes	-0.0764* (0.0439)	0.00133 (0.0227)	-0.0944*** (0.0260)	-0.072 (0.044)	-0.001 (0.023)	-0.091*** (0.026)	0.0124 (0.0307)	-0.0138 (0.0164)	-0.0200 (0.0192)	0.014 (0.033)	-0.016 (0.017)	-0.020 (0.020)
Não-Pesquisadores	-0.160 (0.146)	0.0185 (0.103)	-0.237* (0.142)	-0.172 (0.148)	0.018 (0.101)	-0.271** (0.135)	-0.103 (0.114)	0.0654 (0.0900)	-0.235** (0.115)	-0.129 (0.126)	0.067 (0.092)	-0.274** (0.119)
Dummy de Pesq. Visitantes	0.118*** (0.0372)	0.112*** (0.0229)	0.103*** (0.0232)	0.128*** (0.037)	0.113*** (0.023)	0.099*** (0.023)	-0.00102 (0.0254)	0.0372** (0.0160)	0.0428** (0.0175)	0.010 (0.028)	0.040** (0.017)	0.046** (0.019)
Relação Capital-Trabalho	1.194* (0.674)	-0.562 (0.449)	0.271 (0.333)	1.173* (0.643)	-0.512 (0.439)	0.183 (0.320)	0.342 (0.372)	-0.682* (0.400)	0.00309 (0.224)	0.388 (0.393)	-0.673* (0.403)	-0.063 (0.235)
Idade do Laboratório	-0.219* (0.130)	0.0923 (0.0800)	-0.0554 (0.0615)	-0.219* (0.124)	0.085 (0.078)	-0.038 (0.059)	-0.0729 (0.0698)	0.121* (0.0732)	-0.00872 (0.0418)	-0.085 (0.075)	0.121 (0.074)	0.004 (0.044)
Idade do Laboratório ao Quadrado	-											
Constante	2.552*** (0.972)	0.0409 (0.669)	-1.349** (0.534)				-0.307 (0.580)	0.711 (0.570)	-0.354 (0.359)			
N	207	647	483	207	647	483	207	647	483	207	647	483
R2	0.155	0.103	0.161				0.683	0.546	0.569			
Pseudo R2				0.0677	0.0431	0.0722				0.4228	0.3098	0.3231

Fonte: Elaboração própria.

Tabela 46 – Resultados das Regressões – Efeitos Marginais da Produtividade Científica Individual contra as características coletivas e individuais por *Cluster*, ponderados pelo IFPA (variáveis absolutas).

	yin/coautores*IFPA					
	OLS			Tobit		
	<i>cluster 1</i>	<i>cluster 2</i>	<i>cluster 3</i>	<i>cluster 1</i>	<i>cluster 2</i>	<i>cluster 3</i>
Prod. Cient. do quinquênio anterior	0.792*** (0.0433)	0.698*** (0.0281)	0.694*** (0.0375)	0.827*** (0.048)	0.717*** (0.030)	0.700*** (0.039)
Prod. dos Pesquisadores Permanentes no quinquênio	0.0967** (0.0406)	0.0791*** (0.0236)	0.0670*** (0.0218)	0.097** (0.041)	0.079*** (0.024)	0.075*** (0.023)
Pesquisadores Permanentes	0.0327 (0.0486)	-0.0257 (0.0386)	0.0107 (0.0428)	0.011 (0.053)	-0.028 (0.040)	-0.001 (0.044)
Pesquisadores Não-Permanentes	0.0285 (0.0497)	0.0523** (0.0265)	0.0272 (0.0416)	0.035 (0.050)	0.059** (0.027)	0.036 (0.041)
Não-Pesquisadores	0.0126 (0.0324)	-0.0153 (0.0171)	-0.0223 (0.0192)	0.014 (0.035)	-0.018 (0.018)	-0.024 (0.021)
Dummy de Pesq. Visitantes	-0.116 (0.117)	0.0617 (0.0881)	-0.239** (0.114)	-0.141 (0.124)	0.064 (0.090)	-0.275** (0.120)
Relação Capital-Trabalho	-0.00470 (0.0266)	0.0361** (0.0162)	0.0431** (0.0176)	0.008 (0.029)	0.040** (0.017)	0.047** (0.019)
Idade do Laboratório	0.374 (0.376)	-0.721* (0.407)	-0.0378 (0.225)	0.433 (0.390)	-0.702* (0.408)	-0.082 (0.236)
Idade do Laboratório ao Quadrado	-0.0746 (0.0702)	0.127* (0.0742)	-0.00444 (0.0415)	-0.089 (0.074)	0.125* (0.074)	0.004 (0.044)
Geração: 1962-1979	-0.142 (0.192)	-0.187* (0.0966)	0.331*** (0.115)	-0.173 (0.184)	-0.216** (0.099)	0.314*** (0.120)
Geração: 1980-1989	-0.191 (0.140)	0.0244 (0.0695)	0.0390 (0.0830)	-0.229 (0.143)	0.012 (0.071)	0.050 (0.085)
Geração: 1990-1999	-0.0112 (0.0792)	0.0497 (0.0544)	0.0752 (0.0531)	-0.033 (0.087)	0.041 (0.056)	0.073 (0.057)
Pesquisador de Ciências Exatas e da Terra	0.161 (0.184)	-0.172* (0.0999)	-0.0392 (0.0932)	0.170 (0.169)	-0.183* (0.103)	-0.062 (0.099)
Pesquisador de Engenharias	0.0437 (0.175)	-0.0148 (0.103)	0.150* (0.0905)	0.087 (0.178)	-0.002 (0.102)	0.194** (0.086)
Pesquisador de Ciências Biológicas	-0.0640 (0.306)	0.0673 (0.105)	-0.0725 (0.0862)	-0.123 (0.352)	0.044 (0.112)	-0.135 (0.102)
Pesquisador de Ciências da Saúde	-0.342*** (0.122)	0.0915 (0.262)	0.00320 (0.151)	-0.263** (0.123)	0.100 (0.253)	-0.073 (0.189)
Pesquisador de Ciências Agrárias	-0.469*** (0.109)	0.0211 (0.0889)	0.122 (0.139)	-0.387*** (0.115)	0.039 (0.086)	0.164 (0.139)
Constante	-0.293 (0.590)	0.767 (0.569)	-0.335 (0.367)			
N	207	647	483	207	647	483
R2	0.693	0.553	0.578			
Pseudo R2				0.4345	0.3155	0.3333

Fonte: Elaboração própria.

ANEXO A – QUESTIONÁRIO

LEVANTAMENTO DAS INFRAESTRUTURAS DE PESQUISA

INFRAESTRUTURA DE PESQUISA

O questionário deverá ser aplicado em cada uma das infraestruturas de pesquisa da instituição e deve ser respondido pelo coordenador ou responsável pela infraestrutura.

CARACTERIZAÇÃO

A pergunta abaixo tem o objetivo de classificar o laboratório/infraestrutura em uma taxonomia pré-definida. Assim, pede-se para assinalar com X qual ou quais as categorias que mais representam a infraestrutura. Se necessário, no final da tabela existe um quadro em aberto para sugestões de adição de nova categoria que melhor represente a infraestrutura.

1. Classificação da infraestrutura segundo as categorias abaixo (é possível selecionar mais de uma categoria).

1. TICS, MATEMÁTICA		6. CIÊNCIAS DA VIDA	
	a. Centros de competência em desenvolvimento de		a. Centros de imagem e radiodiagnóstico
	b. Sistemas de computação distribuída		b. Centros de microscopia de luz
	c. Redes de comunicação		c. Bioinformática
	d. Centros de competência em matemática		d. Infraestruturas de bio-nanotecnologia
			e. Centros de genômica, proteômica e transcriptoma
			f. Instalações de fenotipagem
2. ENGENHARIA			g. Atlas biológicos
	a. Pesquisa aeroespacial		h. Coleções de recursos biológicos
	b. Engenharia de transporte		i. Arquivos animais
	c. Laboratórios de testes de materiais		j. Estações de quarentena animal e fazendas
	d. Engenharia civil		k. Laboratórios nível 4 de biossegurança
	e. Centros de pesquisa em manufatura		l. Bancos de sementes
	f. Tecnologias de processo		m. Estufas e câmaras de crescimento
	g. Wave channels		n. Bibliotecas químicas e instalações de screening
3. CIÊNCIA E ENGENHARIA DOS MATERIAIS,			o. Biologia computacional
	a. Laboratórios de difração de raios X e de luz		p. Centros de pesquisa translacional
	b. Fontes de laser		q. Bancos biológicos
	c. Fontes de laser de elétrons livres (free electron		r. Centros de pesquisas clínicas
	d. Fontes de neutron		s. Ciclotrons para aplicações médicas
	e. Muon sources		t. Espectroscopia de massa e outras instalações de
	f. High Power magnets labs		u. Infraestruturas de nutrição e alimentos
	g. Salas limpas		v. Instalações de biologia estrutural
	h. Laboratórios de micro e nanotecnologia		w. Infraestruturas para testes clínicos
	i. Espectroscopia de massa e outros laboratórios de		x. Centros de agronomia e melhoramento de plantas
	j. Laboratórios de análise de superfície		
	k. Laboratórios de microscopia eletrônica		

	l. Laboratórios de alta pressão	7. MEIO AMBIENTE, CIÊNCIAS MARINHAS E DA	
	m. Laboratórios de baixa temperatura		a. Coleções de história natural
	n. Manipulação de actínideos		b. Estações de monitoramento sísmico
4. FÍSICA E ASTRONOMIA			c. Laboratórios de simulação de terremotos
	a. Aceleradores de partículas físicas		d. Outras infraestruturas marinhas
	b. Aceleradores de partículas nucleares		e. Navios de pesquisa
	c. Underground labs		f. Centros de processamento de dados oceânicos e
	d. Detectores gravitacionais		g. Infraestruturas de oceanografia
	e. Telescópios astronômicos		h. Câmeras de testes atmosféricos
5. ENERGIA			i. Infraestruturas de medições atmosféricas
	a. Energia nuclear		j. Aeronaves de pesquisa
	b. Laboratórios de fusão (fusion test facilities)		k. Habitats ecológicos e estações de campo
	c. Plataformas de testes para energia solar		l. Infraestruturas de aquicultura
	d. Instalações para testes de combustão		m. Pesquisa em saúde ambiental
	e. Bioenergia		n. Infraestrutura de gestão ambiental
	f. Geotérmica		o. Outras infraestruturas de ciência ambiental
	g. Eletricidade		
OUTROS TIPOS DE INFRAESTRUTURA, CASO A CLASSIFICAÇÃO NÃO ATENDA AO SEU CASO			
Área		Tipo de infraestrutura	

2. Identificação da infraestrutura (laboratório, biotério, biblioteca etc.)

Nome completo			
Sigla		Início de operação	
Vinculação Instituc.*			
UF		Cidade	
Endereço			
Complemento		CEP	
Telefone		Fax	
E-mail institucional		Website	

* A que unidade (departamento, centro, instituto, escola, faculdade etc), dentro da instituição, a infraestrutura em questão está subordinada ou vinculada.

3. Coordenador(res) *Se houver mais de um coordenador, replicar este campo.*

Nome completo			
CPF		E-mail	
Telefone		Fax	

4. Caracterização

Área física (m2)		Infraestrutura	() Sim	() Não
Prestação de serviços*:	() Sim	() Não	Laboratório acreditado:	() Sim () Não
Linhas de pesquisa:				

* Prestação - por meio de um instrumento formal e mediante alguma forma de remuneração – de um serviço tecnológico, de pesquisa ou de apoio à inovação, incluindo tanto serviços relativos à Tecnologia Industrial Básica quanto serviços criativos voltados ao desenvolvimento de novos produtos ou processos.

5. Serviços tecnológicos prestados

Essa questão será respondida por aqueles que assinalaram “SIM” no quesito “prestação de serviços” do item acima (item 4). Assinalar os principais serviços tecnológicos prestados pela infraestrutura bem como os principais clientes de cada um deles.

Principais serviços prestados:	
Usuários / Clientes	()empresas ()pesquisadores ()governo ()público em geral
Acesso a banco de células, microrganismos etc.	()empresas ()pesquisadores ()governo ()público em geral
Análise de materiais	()empresas ()pesquisadores ()governo ()público em geral
Análise de propriedades físico-químicas	()empresas ()pesquisadores ()governo ()público em geral
Calibração	()empresas ()pesquisadores ()governo ()público em geral
Certificação	()empresas ()pesquisadores ()governo ()público em geral
Consultoria e assessoria técnico-científicas ¹	()empresas ()pesquisadores ()governo ()público em geral
Desenvolvimento e aperfeiçoamento de Processos	()empresas ()pesquisadores ()governo ()público em geral
Desenvolvimento e aperfeiçoamento de Produtos	()empresas ()pesquisadores ()governo ()público em geral
Elaboração e testes de protótipos	()empresas ()pesquisadores ()governo ()público em geral
Ensaio e testes	()empresas ()pesquisadores ()governo ()público em geral
Exames laboratoriais	()empresas ()pesquisadores ()governo ()público em geral
Informação tecnológica	()empresas ()pesquisadores ()governo ()público em geral
Inspeção	()empresas ()pesquisadores ()governo ()público em geral
Manutenção de equipamentos científicos	()empresas ()pesquisadores ()governo ()público em geral
Metrologia	()empresas ()pesquisadores ()governo ()público em geral
Scale up (escalonamento)	()empresas ()pesquisadores ()governo ()público em geral
Serviços ambientais ²	()empresas ()pesquisadores ()governo ()público em geral

1. Inclui, entre outros, os seguintes serviços:

2. Inclui, entre outros, os seguintes serviços:

5.1. Relação das empresas para as quais o laboratório prestou serviços:

Nome da Empresa	CNPJ
-----------------	------

6. Áreas do conhecimento/atuação:

Selecionar uma ou mais áreas do conhecimento a partir da relação/classificação de áreas do conhecimento do CNPq disponível em <http://www.cnpq.br/areasconhecimento/index.htm>. Obs. Utilizar apenas grandes áreas e áreas, sem as subdivisões internas das áreas. É possível selecionar mais de uma área/grande área, no caso de infraestruturas multidisciplinares.

--

7. Breve descrição da infraestrutura de pesquisa (até 1000 caracteres, apenas fatos importantes para o público geral).

--

8. Equipamentos de pesquisa

Descrever, de forma genérica, o conjunto de equipamentos de pesquisa que o laboratório/infraestrutura abriga e estimar o valor atual estimado do conjunto de equipamentos.

Descrição geral do conjunto de equipamentos de pesquisa	Valor global estimado

8.1.Principais equipamentos

Detalhar os principais equipamentos existentes, com custo de aquisição igual ou superior a R\$ 40 mil. Listar apenas aqueles que constituem o diferencial do laboratório/infraestrutura. O foco dessa questão são os equipamentos mais relevantes para a infraestrutura e não aqueles que possuem papel acessório nas pesquisas realizadas ou que possuem valor pouco significativo em relação aos demais equipamentos existentes.

Qd. ¹	Classif. ²	Nome do equipamento	Especificações ³	Ano de	Custo de	Entidade	Estado atual do	Valor estimado ⁷

- Indicar a quantidade de equipamentos, caso sejam iguais e possuam as mesmas características.*
- Classificar o equipamento de acordo com as categorias abaixo. Se possível, indicar nova(s) categoria(s) (suficientemente abrangentes) que considere apropriada(s) para classificar um conjunto de equipamentos de sua área de pesquisa (utilizar o espaço destinado a sugestões no final do questionário).*
 - Cromatógrafos e Espectrômetros*
 - Equipamentos de Informática*
 - Instrumentos Bioanalíticos (analisadores, centrífugas, cintiladores, detectores, sequenciadores de DNA etc.)*
 - Microscópios*
 - Outros (autoclaves, fresadoras, geradores, tratores agrícolas, veículos etc.).*
- Descrever as especificações do equipamento: Marca, modelo e principais características.*
- Indicar o custo unitário de aquisição do equipamento na data em que o mesmo foi adquirido. Se não for possível ter acesso a esta informação, estimar o valor aproximado do equipamento na data em que o mesmo foi adquirido pela instituição. Caso haja mais de um equipamento com as mesmas características, indicar o valor global dos mesmos.*
- Indicar a entidade que financiou a compra do equipamento, dentre as seguintes opções:*
 - A própria instituição ou Entidade Mantenedora*
 - Finep*
 - CNPq*
 - Capes*
 - FAP*
 - Outra instituição pública*
 - Empresa (pública ou privada)*
 - Outra entidade.*
- Indicar o estado atual do equipamento, utilizando as seguintes opções:*
 - Atualizado e em condições de uso*
 - Obsoleto, mas em condições de uso*
 - Sem condições de uso*
 - Inoperante (por falta de manutenção).*
- Estimar o valor econômico do equipamento. Para fazer esta estimativa, calcular o valor aproximado de venda do equipamento em suas condições atuais. Caso haja mais de um equipamento com as mesmas características, indicar o valor global dos mesmos.*

OPERAÇÃO

9. Equipe permanente da infraestrutura

Considera-se como equipe permanente o pessoal (coordenador, principais pesquisadores, técnicos e estudantes de pós-graduação) que trabalha regularmente no laboratório/infraestrutura – independentemente do tipo de vínculo institucional – por, pelo menos, 10 horas semanais. São os profissionais que garantem o funcionamento da infraestrutura e são corresponsáveis pelos equipamentos ou que estejam envolvidos com as principais pesquisas desenvolvidas na mesma.

9.1. Qual o número de profissionais que compõem a equipe permanente da infraestrutura?

		Quantidade
Pesquisadores	Doutores e pós-doutores	
	Mestres	
	Especialistas	
	Graduados	
Estudantes	Doutorandos	
	Mestrandos	
Apoio técnico*	Nível superior	
	Nível médio	
Apoio administrativo		

9.2. Relação dos Pesquisadores (doutores e pós-doutores, mestres, especialistas e graduados) da equipe permanente.

Nome	CPF

10. Caso seja pertinente, informe os principais Programas de Pós-Graduação, com avaliação positiva da CAPES, que utilizam regularmente a infraestrutura/laboratório para o desenvolvimento de suas atividades de ensino e pesquisa:

Área de Avaliação ¹	Programa de Pós-Graduação	Nota na avaliação da Capes ²			
		M	D	F	M/D

1. Listar apenas os programas das seguintes grandes áreas: ciências agrárias; ciências biológicas; ciências da saúde; ciências exatas e da terra; engenharias; e multidisciplinar. Consultar em: <http://www.capes.gov.br/avaliacao/cursos-recomendados-e-reconhecidos>.

2. Legenda:

M - Mestrado Acadêmico

D - Doutorado

F - Mestrado Profissional

M/D - Mestrado Acadêmico/Doutorado

11. Utilização por usuários externos.

Consideram-se como usuários externos todos os pesquisadores e alunos que não fazem parte da equipe permanente, incluindo aqueles vinculados a mesma instituição que abriga a infraestrutura.

11.1. A infraestrutura é aberta à utilização por usuários externos?

() sim () não

11.2. Caso positivo, qual a política de acesso e os procedimentos adotados para a utilização da infraestrutura por pesquisadores externos?

--

12. Número aproximado de usuários externos da infraestrutura

Tipo de usuário	Procedência	Quantidade
Pesquisadores da mesma instituição (da própria unidade ou de outra unidade – depto./faculdade/área)	--	
Pesquisadores de outras instituições	Brasil	
	Exterior	
Alunos de pós-graduação	Brasil	
	Exterior	
Alunos de graduação	Brasil	
	Exterior	
Pesquisadores de Empresas	Brasil	
	Exterior	

13. Utilização da infraestrutura

Assinale quais as principais atividades desenvolvidas pela infraestrutura (é possível selecionar mais de uma alternativa)

Principais atividades desenvolvidas na infraestrutura	Intensidade de uso ¹
Atividades de pesquisa	
Atividades de ensino	
Desenvolvimento de produtos para o setor produtivo	
Prestação de serviços tecnológicos (testes, ensaios etc.)	
Atividades de extensão tecnológica	
Outras atividades (quais?) _____	

1. Informar a intensidade de utilização da infraestrutura com a atividade assinalada, indicando uma das seguintes opções de uso:

- a. Contínuo
- b. Todos os dias úteis da semana
- c. Alguns dias da semana
- d. Alguns dias do mês
- e. Esporádico.

14. Atividades de cooperação

Assinale, caso existam, as principais atividades de cooperação desenvolvidas pela instituição, com o auxílio da infraestrutura. Entende-se por atividade de cooperação aquela formalizada por meio de contratos, convênios ou acordos de cooperação (é possível assinalar mais de uma alternativa).

Tipo de atividade de cooperação desenvolvida	Grau de importância ¹
Cooperação com outras instituições de pesquisa no Brasil.	
Cooperação com outras instituições de pesquisa no Exterior.	
Cooperação com empresas no Brasil	
Cooperação com empresas no Exterior	

	Participação em projetos de cooperação financiados/apoiados	
	Outros: quais? _____	

1. Grau de importância atribuído a atividade assinalada, tomando como parâmetro o conjunto de atividades desenvolvidas pela infraestrutura. Indicar uma das três categorias:
- Alto
 - Médio
 - Baixo

VALOR, RECEITAS E CUSTOS

15. Valor total estimado da infraestrutura

Essa pergunta procura levantar qual o seria o valor estimado da infraestrutura em questão. Esta estimativa inclui o valor das instalações físicas, equipamentos, mobiliário etc. e exclui os custos operacionais e de manutenção.

Valor total estimado da infraestrutura	
	Até R\$ 500 mil
	Acima de R\$ 500 mil até R\$ 1 milhão
	Acima de R\$ 1 milhão até R\$ 3 milhões
	Acima de R\$ 3 milhões até R\$ 5 milhões
	Acima de R\$ 5 milhões até R\$ 10 milhões
	Acima de R\$ 10 milhões até R\$ 20 milhões
	Acima de R\$ 20 milhões até R\$ 30 milhões
	Acima de R\$ 30 milhões até R\$ 50 milhões
	Acima de R\$ 50 milhões até R\$ 100 milhões
	Acima de R\$ 100 milhões até R\$ 200 milhões
	Acima de R\$ 200 milhões

16. Estimativa de custos operacionais

Estimar todos os custos operacionais, incluindo pessoal administrativo e manutenção, com a infraestrutura.

Elemento de despesa	Estimativa de custo
Pessoal (salários, benefícios, bolsas e outros tipos de	
Despesas gerais (água, energia, telefone etc.)	
Material de consumo (insumos para a pesquisa, material de	
Depreciação (predial + equipamentos)	
Serviços de terceiros	

17. Fontes de receitas

Nessa questão, o objetivo é conhecer as fontes de receita da infraestrutura. Inclui o orçamento da própria instituição (ou de sua mantenedora), subvenções, projetos de pesquisa financiados por instituições públicas de fomento, projetos com empresas, venda de serviços tecnológicos etc.

Fonte de Receita	Receitas (Em R\$ mil)	
	Custeio	Capital
Própria Instituição ou Entidade mantenedora		
FINEP		
CNPq		
CAPES		
Fundação estadual de amparo à pesquisa		

Outra instituição pública		
Empresas publicas ou privadas		
Outra (qual?):		
Valor total em R\$ correntes		

CONDIÇÕES ATUAIS DA INFRAESTRUTURA

18. Capacidade técnica da infraestrutura

Essa é uma questão relativamente subjetiva, na qual procura-se saber, do coordenador da infraestrutura, sua opinião sobre a distância e/ou proximidade daquela infraestrutura em relação à fronteira tecnológica da sua área de atuação. Marcar apenas uma opção.

Avaliação da capacidade técnica da infraestrutura	
☐	Avançada e comparável com as melhores infraestruturas mundiais do gênero
☐	Avançada em relação aos padrões brasileiros mas ainda distante das melhores infraestruturas mundiais do gênero
☐	Adequada e com capacidade técnica próxima de outras infraestruturas do gênero no Brasil
☐	Insuficiente em relação a outras infraestruturas do gênero no Brasil
☐	Não sabe

19. Avaliação das condições atuais da infraestrutura

Itens	péssimo	ruim	bom	Muito bom	excelente	Não se aplica
Instalações físicas						
Equipamentos						
Manutenção						
Insumos de pesquisa						
Equipe de pesquisadores						
Equipe de apoio técnico						

20. Há quanto tempo houve a última recuperação/atualização/modernização de parte dos equipamentos ou de toda a infraestrutura (com custo de, pelo menos, 10% do custo total da instalação, segundo a estimativa realizada na questão 15)?

☐	Menos de 1 ano	☐	De 10 a 15 anos
☐	De 1 a 5 anos	☐	Mais de 15 anos
☐	De 5 a 10 anos	☐	