INTERNATIONAL RESEARCH COLLABORATION, RESEARCH TEAM PERFORMANCE, AND SCIENTIFIC & TECHNOLOGICAL CAPABILITIES IN COLOMBIA –A BOTTOM-UP PERSPECTIVE

A Dissertation
Presented to
The Academic Faculty

by

Gonzalo Ordóñez-Matamoros
gonzaloord@hotmail.com

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
Joint Program of the School of Public Policy and the
Andrew Young School of Policy Studies

Georgia Institute of Technology and Georgia State University
December 2008

INTERNATIONAL RESEARCH COLLABORATION, RESEARCH TEAM PERFORMANCE, AND SCIENTIFIC & TECHNOLOGICAL CAPABILITIES IN COLOMBIA –ABOTTOM-UP PERSPECTIVE

Approved by:

Dr. Susan E. Cozzens, Advisor
School of Public Policy
Georgia Institute of Technology

Dr. Alan L. Porter
School of Industrial and Systems Engineering – School of Public Policy
Georgia Institute of Technology

Dr. Gregory B. Lewis
Andrew Young School of Policy Studies
Georgia State University

Dr. Juan D. Rogers
School of Public Policy
Georgia Institute of Technology

Professor J. Adam Holbrook
Centre for Policy Research on Science and Technology
Simon Fraser University

Date Approved: September 17, 2008

To Camilo, Juliana, Margarita, Enna, Alfonso, and my sisters and brothers for their patience, moral support and unconditional love.

ACKNOWLEDGEMENTS[1]

I am grateful to Dr. Susan Cozzens who I admire as a professor, researcher, boss, and mentor. I am indebted for her constant support, which came in many forms during my studies in Atlanta and beyond. Thanks also to the members of my dissertation committee: Dr. Greg Lewis, Dr. Alan Porter, Dr. Juan Rogers, and Professor Adam Holbrook, for the comments and insights provided in the preparation of the dissertation. I also appreciate the encouragement, time and advice given by my colleagues at the Technology Policy and Assessment Center (TPAC), at Georgia Tech.

Thanks to Jorge Lucio, Luis Orozco, Diego Chavarro, Diana Lucio and Rafael Hurtado for their guidance and the information provided for the research. To Dr. Hernán Jaramillo, Dr. Jorge Charúm, and the interviewees for sharing with me their knowledge, understanding and insights of the Colombian S&T system and dynamics.

Last but not least, I would like to specially thank Dr. Fernando Hinestrosa for his generous and invaluable support through all these years; to Dr. Roberto Hinestrosa for his confidence and long term vision; to Dr. Mauricio Perez for introducing me into the research enterprise; to my colleagues at the Universidad Externado de Colombia for their friendship and moral support.

[1] This material is based upon work supported by the National Science Foundation under Grant Number 0647126. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

A&HCI	Arts and Humanities Citation Index
ATE	Average Treatment Effect
ATT	Average Treatment Effect on the Treated
ATU	Average Treatment Effect on the Untreated
CAN	Comunidad Andina de Naciones – Andean Nations Community
C&K	Caliendo and Kopeining
CIA	Conditional Independence Assumption
CIAT	Centro Internacional de Agricultura Tropical
CNS&TS	Colombian Science and Technolgy System
COLCIENCIAS	Colombian Institute for the Development of Science and Technology
CSC	Common Support Condition
FTA	Free Trade Agreement
GDP	Gross Domestic Product
IAC	InterAcademy Council
IDRC	International Development Research Council
IRC	International Research Collaboration
ISI	Institute of Scientific Information
LAC	Latin American and Caribbean Countries
L&F	Long & Freese
LRM	Linear Regression Model
NBRM	Negative Binomial Regression Model
NGO	Non-Governmental Organization

NSF	National Science Foundation
OCyT	Colombian Observatory of Science and Technology
OECD	Organization for Economic Cooperation and Development
PSM	Propensity Score Matching
RC	Research Collaboration
R&D	Research & Development
RICYT	Ibero-American S&T Indicators Network – Red Iberoamericana de Indicadores de Ciencia y Tecnología
RT	Research Team
RTP	Research Team Production or Research Team Productivity
SCI	Science Citation Index
S&E	Science & Engineering or Scientists & Engineers
S&T	Science & Technology
SNC&T	Sistema Nacional de Ciencia y Tecnologia
SSCI	Social Science Citation Index
STI	Science, Technology and Innovation
TWAS	Third World Academy of Science
UN	United Nations
UNCTAD	United Nations Cooperation for Trade and Commerce
UNDP	United Nations Development Program
UNIDO	United Nations Industrial Development Organization
WOS	Web of Science
ZINB	Zero-Inflated Negative Binomial Model
ZIP	Zero-Inflated Poisson Model

SUMMARY

Recent trends show that Colombian science and technology (S&T) performance is improving rapidly. This is presumably the result of two 'mega trends' characterizing the Colombian S&T system: 1) the rapid professionalization of the R&D enterprise, as reflected by the formation of research teams with the support of the Colombian government and the elite research institutions; 2) the internationalization of its scientific community, especially since the 1990s after the opening of the economy to foreign trade.

This dissertation examines the factors affecting Colombian S&T performance, and particularly the ways international research collaboration affects local scientific and technological capabilities. S&T capabilities are measured by the ability of research teams to produce bibliographic outputs, and to contribute to local knowledge.

Research hypotheses are tested using Zero Inflated Negative Binomial Regression models and logistic regressions to account for the effects of international research collaboration on team output while controlling for team characteristics, partner characteristics, scientific discipline, sector, the characteristics of the teams' home institution, and team location. The study uses control groups and the Propensity Score Matching approach to assess the overall impact of international research collaboration on research team performance while controlling for the effects of endogeneity and selection bias.

Results show that international research collaboration is positively associated with both team output and teams' ability to contribute to local knowledge. The study shows that such effects depend on the type of collaboration chosen and the type of partner involved. Particularly, it shows that while co-authoring with colleagues located overseas or receiving foreign funding increases team output, hosting foreign researchers does not seem to affect a team's productivity once all other variables are held constant. It also finds that collaborating with partners from the South yields greater productivity counts than collaborating with partners from the North, and that funding from southern countries is associated with greater productivity rates than any other combination of collaboration activity and origin of partners.

The study also finds that hosting foreign researchers does not appear to be associated with the probability of teams to involve Colombia in their research process either, and that receiving foreign funding or co-authoring with colleagues located overseas increases a team's probability to contribute to local knowledge. Similarly, the study finds that collaboration with partners from northern countries is strongly associated with a team's ability to contribute to local knowledge, while collaboration with partners from southern countries is not. The study finds that although the number of participating researchers holding doctorates positively affects team output, it negatively affects a team's ability to contribute to local knowledge -- but as team size increases beyond 9 members with a PhD, its effects become positive at an increasing rate. Finally, the study finds curvilinear effects of team size, team age and number of active R&D projects a team manages. Theoretical and policy implications of these and other counterintuitive findings are discussed.

CHAPTER 1

INTRODUCTION

1.1 Internationalization and Institutionalization of Science, Technology and Innovation

International research collaboration is a growing social phenomenon (Wagner and Leydesdorff 2006; NSF-NSB 2008). It results in part as a strategy to deal with increasingly complex problems and the rising costs of research (Luukkonen, Persson; et al. 1992; Gibbons, Limoges et al. 1994; Adams, Black et al. 2005). It also responds to government policies oriented to favor globalization (Georghiou 1998; Wagner, Brahmakulam et al. 2001). Finally, the continuous fall of communication costs and the increased mobility of scientists and students across borders are also contributing to this phenomenon.

According to the US National Science Foundation (NSF), the number of international articles with authors from at least two countries more than doubled in share between 1988 and 2003 from 8% to 20%. The number of countries collaborating on an article also expanded. In 2003, more than 60 countries had co-authored with other countries, compared with 32 in 1996 (NSF-NSB 2006). Over the period, 1995-2005, intercontinental co-authorship increased as a percentage of total article output for the US (from 17% to 27%), for the EU (from 18% to 26%), and for Asia (from 16% to 19%)(NSF-NSB 2008), resulting in an increasing level of international interdependence of the research enterprise (Narin, Stevens et al. 1991; Glänzel and Schubert 2004; Glanzel and Schubert 2005; NSF-NSB 2008).

A second and growing trend in addition to the internationalization of the S&T community is the professionalization and institutionalization of the scientific and technological enterprise (Gibbons, Limoges et al. 1994; Etzkowitz and Kemelgor 1998; Laredo 2003). Indeed, the model shift of knowledge production described by Gibbons and colleagues more than a decade ago (Gibbons, Limoges et al. 1994), portraying a shift towards multi- and inter-disciplinary research and the decline of single individual and single discipline research, seems to be now largely confirmed by the emergence of research teams or groups (Kretschmer 1985; Cohen 1991; Seglen and Aksnes 2000; Rey-Rocha, Martin-Sempere et al. 2002; Laredo 2003; Newman 2004; Carayol and Matt 2004a; Carayol and Matt 2004b; Adams, Black et al. 2005; Lima, Liberman et al. 2005; Calero, Buter et al. 2006; Carayol and Matt 2006).

From the policy perspective, these research teams are not only indicators of local S&T capabilities but multipliers of such capacities. They are increasingly regarded as vehicles of S&T progress and the building blocks of science, technology and innovation systems (Crow and Bozeman 1998; Etzkowitz and Kemelgor 1998; Laredo and Mustar 2001; Amsterdamska 2008; Mirowski and Sent 2008).

These two trends (internationalization and institutionalization) are not only taking place in developed countries but are arguably happening at a particularly rapid pace in developing countries. Research on these phenomena and on their consequences in developing countries is rather scarce, however. This dissertation contributes to current knowledge and understanding of the extent, characteristics, and ways international research collaboration affects S&T capabilities, as reflected by the performance of research teams in the context of a developing country: Colombia.

1.2 Colombian S&T Performance

As most developing countries, Colombia has S&T strengths in research areas such as tropical medicine and agriculture but lacks important aspects of S&T capacity in personnel, infrastructure, investment, and institutional environment. As reported by the Interamerican/Ibero-American Network on S&T Indicators (RICYT for its name in

Spanish), and based on comparative statistics gathered for most countries in the region, Colombia, with the third largest population in Latin America and the fourth largest GDP in the region, a) spends a very low percentage of its GDP on S&T (0.5%); b) allocates a small portion of its human resources to the performance of S&T activities (620 researchers per million inhabitants of working age, less than half of the region's average); c) performs poorly in S&T as reflected by its research outputs (0.08% of world articles and an average of 7.1 articles published in high impact journals per 100 researchers, which is half the region's average); and d) has low innovative capacity (220 patents per million inhabitants compared to the average of 1,620 patents per million people in the 10 largest economies in Latin America) (RICYT 2004). Table 1 summarizes these indicators.

Table 1. Colombian Basic S&T Indicators

Latin America: Selected Input and Output Indicators. 2005

Country	Population millions	Expenditure on S&T as % of GDP (a)	Researchers per thousand labor force (b)	% Researchers with PhD (c)	Invention Coefficient (d)	Publications in SCI Search as % of World (e)	Publications in SCI Search per 100 researchers (f).
Argentina	37.8	0.53	3.16	23.7	2.79	0.49	11.62
Brazil	184.2	1.12	1.55	61.8	5.99	1.6	12.36
Chile	16.3	0.68	2.78	NA	3.52	0.28	16.29
Colombia	45.29	0.51	0.62	17.2	0.22	0.08	7.14
Costa Rica	4.3	1.10	0.76	25	0.88	0.03	23.2
Ecuador	13.23	0.18	0.16	10.4	0.38	0.02	22.8
Mexico	103.83	0.46	1.03	NA	0.56	0.64	17.17
Peru	27.97	0.16	0.41	NA	0.14	0.03	6.67
Uruguay	3.31	0.28	3.1	11.9	0.82	0.04	10.37
Venezuela	26.6	0.23	0.59	51.5	0.89	0.11	15.63
Average 10		0.53	1.42	28.79	1.62	0.33	14.33

(a) Costa Rica: 2004; Ecuador: 2003; Uruguay: 2002; Chile and Peru: R&D, 2004
(b) Head Count. Brazil, Chile, Colombia, Peru and Venezuela: 2004: Ecuador: 2003; Uruguay: 2002; Mexico: Full Time Equivalent -FTE
(c) Brazil, Colombia: 2004; Ecuador and Venezuela: 2003; Uruguay: 2002
(d) Patents applied for by residents per thousand inhabitants. Brazil, Ecuador, Peru, and Venezuela: 2004
(e) Counties may be counted twice in international articles
(f) Based on head count except for Mexico (FTE). Brazil, Chile, Colombia, Peru, and Venezuela: 2004; Ecuador: 2003; Uruguay: 2002
Source: RICYT, calculations by the author

Arguably, this rather poor performance is explained in part by the country's isolation from the global market experienced during the import substitution period of the 1970s and 1980s (Garay 1998), which seems to be affecting Colombian competitiveness[2].

Similarly, Colombian capacity to contribute to local knowledge and understanding is relatively poor. Based on the analysis of the documents published between 1980 and 2005 in journals indexed by the ISI's Web of Knowledge3, local scientists scarcely write more about Colombian issues or use Colombia as their unit of analyses than scientists located overseas. In fact, as shown in Figure 1, Colombian S&T is barely self-sufficient (countries above the 0 are self-sufficient, and those below 0 are dependent on international STI capacity).

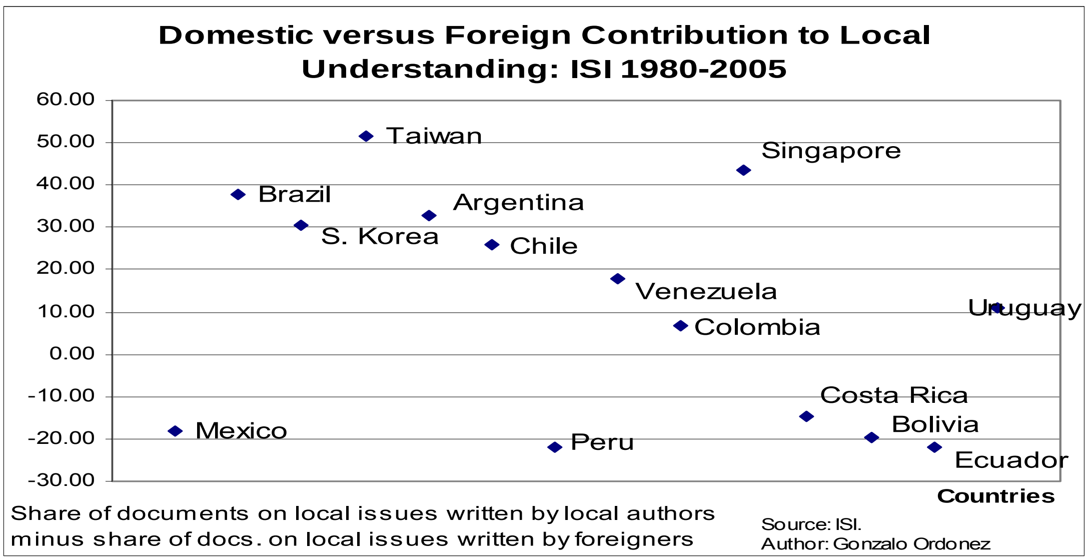

Figure 1. Colombian Contribution to Local Knowledge: 1980-2005

---

[2] According to a survey census of the Colombian manufacturing firms in 2005, only 8.3% of the more than 6,000 establishments surveyed can be considered 'radical innovators'; 17.2% are classified as 'incremental innovators'; 7.9% as 'organizational innovators'; 43.1% as 'technologically adequate'; and the remaining 23.5% as 'non-innovative firms' since they do not show having invested on innovation or development activities, or do not report progress on the level of attainment of their innovation objectives DANE (2006). Innovacion y Desarrollo Tecnologico en la Industria Manufacturera. Colombia 2003-2004. Bogota, D.C., Departamento Administrativo Nacional de Estadistica -DANE, Departamento Nacional de Planeacion - DNP, Insituto Colombiano para el Desarrollo de la Ciencia y la Tecnologia -COLCIENCIAS..

[3] See http://www.isiwebofknowledge.com/

However, since the 1990s Colombian scientific and technological capacity has experienced a rapid improvement. Based on the analysis of the data from the Web of Knowledge, the number of Colombian scientific publications appearing in high quality journals has doubled in the last 10 years, revealing the highest growth rate in the region.

Many plausible explanations of this recent performance have been offered in public debates. These include a) the leading role played by Colciencias, the Colombian national science foundation, in encouraging higher quality of research by ranking research teams and using this rank to support funding decisions; b) the process of academic accreditation led by ICFES, the Colombian Institute of Higher Education, oriented at encouraging the transition of higher education institutions to research-based institutions; c) the loans contracted with IDB, the Inter-American Development Bank, to fund R&D and innovation activities as well as masters and doctoral education; d) the increased market competition resulting from the opening of the economy to foreign products and services; and e) the increased interaction between the Colombian S&T community and their foreign partners. None of these hypotheses have been empirically investigated, however. This dissertation chooses to test the hypothesis of the internationalization process and acknowledges the leading role currently played by research teams in Colombia.

In this sense, as reported in a preliminary paper written by the author using data from the Web of Knowledge on more than 5,400 journal articles published by Colombian scientists and engineers between 1980 to 2005, this recent good performance seems to be explained by the country's increased international collaboration (Ordonez 2005). As shown in Figure 2, while the number of articles published by Colombians alone is rather small, that published in collaboration with foreign partners is large and rising rapidly. The causes, drivers and implications of this pattern are still to be explained, however, and that is one of the goals of this dissertation.

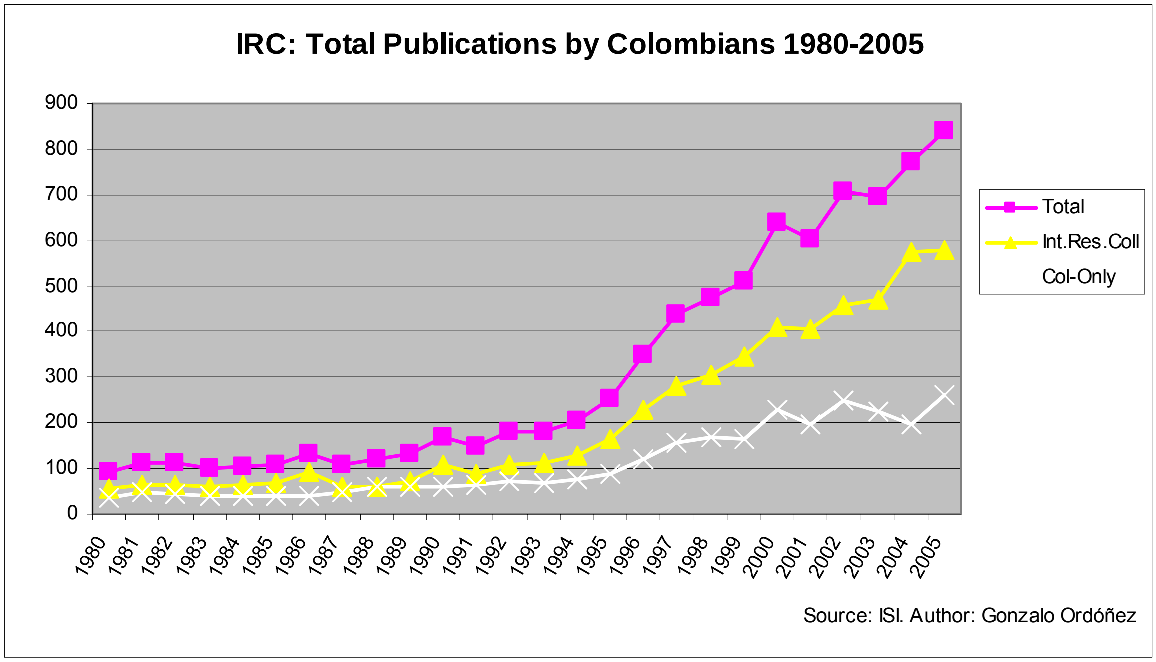

Figure 2. Publications and Research Collaboration: 1980-2005

In addition, a second and important trend taking place in Colombia is the rapid process of institutionalization of the scientific enterprise. The analysis of the data provided by the Colombian Observatory of Science and Technology (OCyT)[4] shows that the number of research teams responding to the calls made by the Colombian institute for S&T development (Colciencias) to update its directory has dramatically increased: It jumped from fewer than 600 to more than 3,000 in the last decade. In fact, during the last decade, the number of individuals reporting collaborative activities and institutional affiliation with a research team nearly quadrupled: it rose from less than 5,000 in 1995, to more than 12,000 in 2000, to nearly 20,000 in 2005. Today, these teams host most of the Colombian scientific community estimated by the OCyT to be of more than 24,000 individuals, of which more than 10,000 people report research outputs (OCyT 2007).

[4] See www.ocyt.org.co

Finally, the analysis of the autonomous capacity of the country to contribute to local knowledge shows a small but important increase experienced during the last decade. While the difference between the share of documents on Colombian issues written by local scientists and engineers and the share of documents on those issues written by foreign scientists and engineers was nearly -30 in the 1970s, it felt to -10 in the 1980s, it became positive in the 1990s and today is somewhere around +20. This capacity remains very low compared to that shown by Asian countries and other comparable Latin-American countries, however.

However, whereas there is a relatively well established research team policy in Colombia (Jaramillo 2007), the country still lacks a coherent internationalization policy involving science, technology and innovation activities. In fact, little is known on the determinants, characteristics, processes and impacts of international research collaboration in Colombia.

Thus this dissertation contributes to current understanding of the extent international research collaboration affects S&T capabilities in Colombia, as reflected by the performance of its research teams. In this framework, S&T capabilities are measured by the production of scientific results by local teams and by their ability to contribute to the study of issues of the home country's interests. Mediating factors such as team characteristics, partner characteristics, scientific discipline, sector, location, characteristics of the teams' home institution, team size, team age, and characteristics of the team leader are taken into account to better understand the ways international research collaboration affects research team performance. International research collaboration is measured through the co-authorship of journal articles, the participation of foreign researchers in local research teams, and the reliance on foreign funding to team R&D projects.

The analyses tests several research hypotheses using zero-inflated negative binomial regression models to predict counts of scientific production, and using logistic regression to evaluate the factors explaining the probability of teams to work on issues of local relevance. In each case, the impacts of different types of collaboration and of different types of partners (North and South) are investigated. The propensity score matching approach is used to assess the impact of international research collaboration while controlling for selection bias and prevent endogeneity. The analyses are based on cross sectional data of 1889 Colombian research teams active between 2003 and 2005 working in all scientific fields, and on a sample of 672 teams.

CHAPTER 2

INTERNATIONAL COLLABORATION AND S&T CAPABILITIES

This chapter presents the literature found on the definitions, processes, and impacts of research collaboration, discusses the specific contributions this dissertation makes to current work done on the topic by sociologists, economists, S&T policy evaluation scholars, and the international relations students. The chapter ends with the discussion of the theoretical model and presents the hypotheses that guide the study.

2.1 Research Collaboration

The literature on the characteristics and on the determinants of research collaboration is rather abundant. Katz and Martin define research collaboration as the working together of researchers to achieve the common goal of producing new scientific knowledge (Katz and Martin 1997). A variety of 'collaborative activities' has been identified as falling under this broad concept. As Bordons and Gomez (2000) claim, these include the expression of opinions, the exchange of ideas and data, working together during the course of a project, working separately on different parts of a project with the purpose of integrating the results at the end, sharing equipment, and exchanging personnel. (Bordons and Gomez 2000).

Similarly, several concepts have been proposed in the literature referring to research collaboration, including **a)** 'Invisible Colleges' (Price and Beaver 1966; Crane 1972; Cronin 1982; Gmur 2003), **b)** 'Research Networks' (Thorpe and Pardey 1990; Callon, Courtial et al. 1991; Callon 1992; Hicks, Isard et al. 1996; Hicks and Katz 1996; Malo and Geuna 2000; Newman 2001; Newman 2001; Landry, Amara et al. 2002; Heimeriks, Horlesberger et al. 2003; Helble and Chong 2004; Rigby and Edler 2005), **c)** 'Research Partnerships' or 'Strategic Alliances' (Carayannis, Alexander et al. 2000;

Hagedoorn, Link et al. 2000; Hagedoorn 2002; Link, Paton et al. 2002; Carayannis and Laget 2004; Kastelli, Caloghirou et al. 2004), **d)** 'Sabato Triangle' or 'Triple Helix' (Sabato 1975; Sabato and Mackenzi 1982; Leydesdorff and Etzkowitz 1998; Etzkowitz and Leydesdorff 2000; Heimeriks, Horlesberger et al. 2003; Leydesdorff and Meyer 2003), **e)** 'Innovation Systems' (Lundvall 1992; Nelson 1993; Acs, de la Mothe et al. 1996; OECD 1997; Holbrook and Wolfe 2000; Holbrook and Salazar 2004), **f)** 'Innovation Clusters' (Saxenian 1994; OECD 1999; Porter 2001; Holbrook and Wolfe 2002; Andersson, Serger et al. 2004; Dahl and Pedersen 2004) **g)** 'Knowledge Value Alliances' (Rogers 2001; Rogers and Bozeman 2001), **h)** 'Knowledge Value Collectives' (Bozeman and Rogers 2002), or 'simply' **i)** 'Research Collaborations' (Beaver and Rosen 1979; Beaver and Rosen 1979; Katz and Martin 1997; Bordons and Gomez 2000; Hagedoorn, Link et al. 2000; Beaver 2001).

However, as Katz and Martin (1997) acknowledge, both the concept of 'working together' and the assumption of a 'common goal' as a distinctive characteristic of a collaborative activity are rather conceptually and empirically problematic since, **a)** it is not clear how closely researchers have to work together in order to constitute a collaboration, and **b)** either no two researchers ever have precisely the same goals, or, conversely, every single researcher in the world is in fact a member of a big collaboration called 'scientific community' for they all work to advance scientific knowledge and are all somewhat interrelated: they all exchange ideas on what experiments to do next, what hypothesis to test, what new instrumentation to build, how to relate their latest experimental results to theoretical models, and so on" (Katz and Martin 1997).

As Bordons and Gomez acknowledge, if we take a narrow definition and agree that collaboration is defined as two or more scientists working together on a joint research project, sharing intellectual, economic and/or physical resources, a wide range of situations still can be included, and a wider array of contributions will in fact be excluded under such definition.

It seems therefore that, as the authors acknowledge, a research collaboration has a very “fuzzy” or ill-defined border, and exactly where that border is drawn is a matter of social convention and is open to negotiation. Furthermore, perceptions regarding the precise location of the ‘boundary’ of the collaboration may vary considerably across institutions, fields, sectors, countries, actors, and purposes over time. The fact is that, as any other social process, research collaboration is mainly governed by the complexity of human interactions, which we still don’t understand completely.

Nevertheless, several types of collaboration are identified in the literature. As Bordons and Gomez (2000) point out, they can be theoretical or technical, the former being based on the exchange of ideas, the provision of advice, or criticism, and the latter being based the share of resources, methods, etc. (Bordons and Gomez 2000). Another typology of collaboration is offered by Hagedoorn, Link et al (2000), who claim that research partnerships can be either formal or informal and can involve any type of partners (i.e. scientists, technicians, students, employees, etc.), belonging to universities, enterprises or government agencies committed to research projects. While formal research partnerships include research corporations (equity joint ventures focusing on research, and research joint ventures) and contractual arrangements such as strategic technical alliances, etc., informal agreements includes short-term research project-specific endeavors (Hagedoorn, Link et al. 2000), and less visible but not less important social contacts.

Why do scientists collaborate? According to Beaver (2001) researchers collaborate to gain access to equipment or other types of resources; to access to new funds; to obtain prestige or visibility; for professional advancement; to make progress more rapidly; to tackle “bigger” problems (more important, more comprehensive, more difficult, global); to enhance research productivity; to claim primacy, ownership and rewards; to get to know more people and to create a network; to learn new skills or techniques; to share the excitement of an area with other people; to find flaws more efficiently, reduce errors and mistakes; to keep one more focused on research and avoid doing other activities; to reduce isolation, and to recharge one's energy and excitement; to educate (a student, graduate student, or oneself); to advance knowledge and learning; and for fun, amusement, and pleasure (Beaver 2001).

In a survey administered on 195 first-listed authors of institutionally co-authored journal articles registered by the 1994 CD-ROM version of the Science Citation Index with at least one address at a Swedish University, Melin (2000) found that 41% of the interviewed collaborated mainly because of his/her co-author's special competence; 20% because of his/her co-author's special data or equipment; 9% were more interested in collaborating mostly for developing and testing a new method; 16% because of social reasons (old friends, past collaboration, etc.); and 14% mostly motivated by supervisor-student relations. The author found that in many cases collaboration started up from attending conferences and attending social and academic events (Melin 2000).

Finally, the choice of collaborating also depends on the characteristics of the discipline one works in. In fact, some R&D projects belonging to disciplines such as physics are more likely to be collaborative than projects belonging to, for example, the social sciences and the humanities such as sociology or philosophy. Indeed, As Frame and Carpenter claim, the fact that most disciplines differ in their epistemological and methodological characteristics makes research collaboration a complex enterprise (Frame and Carpenter 1979). Whereas such differences can translate into practices or ethos that negatively affect the progress of inter-disciplinary collaboration, in some cases they can affect it positively.

2.2 What is International Research Collaboration?

Arguably, the similarities between research collaboration and international research collaboration are greater than the differences between the two. However, distinctive aspects of international research collaboration, besides the 'obvious' condition that partners belong to different nations, include a different set of drivers, enablers, modalities, and consequences.

As for the drivers of International Research Collaboration, and according to Wagner and Leydesdorf (2004), these include: **a)** location of specific resources. Marine research for example would probably require accessing different ocean resources from different countries; **b)** unique expertise. The treatment of some disease may well require local expertise in those areas where it has developed and being investigated from the past; **c)** location of large-scale equipment. A space research initiated in Russia would probably need to work at NASA to do some of their experiments; **d)** global problems requiring global solutions. Global warming would probably require research performed in different places of the planet to monitor and understand the causes (Wagner and Leydesdorff 2004).

As for the enablers of international research collaboration is concerned, the literature identifies the following: **a)** the return to home country of former 'brain drained'. It is well known (thought barely tested empirically) that one of the factors driving international research collaboration are the social networks created by foreign students and professors who return to their home countries and maintain their contacts with their mentors, colleagues or students in the countries where they spend part of their academic lives (Melin 2004); **b)** the Diaspora. Many of those who do not return to their countries of origin keep the contacts made in the past or develop new ones with their co-nationals they meet in international workshops or other academic and social events (Basu and Kumar 2000; Chaparro, Jaramillo et al. 2004); and **c)** the Cultural-, geographic-, historical-, linguistic-, proximity. One is more likely to collaborate with whom one shares more basic characteristics than with those one shares less common characteristics (Frame and Carpenter 1979; Narin, Stevens et al. 1991; Katz 1994; Farrell 2001; Lee 2004; Levine and Moreland 2004; Wagner 2005); In addition, relatively low costs of transportation and communication have contributed importantly to the collaborative enterprise across borders.

Some of the barriers to international research collaboration identified in the literature include **a)** low absorptive capacity. According to Cohen and Levinthal, it is the lack of absorptive capacity of the knowledge and technology produced in developed countries what keeps developing countries from benefiting from the advances of the modern world (Cohen and Levinthal 1990). In fact, very often, researchers from developing countries are not able to take advantage of the knowledge and techniques offered by partners working in developed countries mostly because they lack the basic resources and knowledge necessary to exploit such opportunities (Bayona, Garcia-Marco et al. 2001; Penner-Hahn and Shaver 2005); **b)** strong intellectual property protection (Forero-Pineda and Jaramillo-Salazar 2002); and **c)** political reasons oriented at controlling migration, ensuring national security, etc.

Finally, the modalities of international research collaboration include working with foreign partners affiliated with local teams, working in projects with foreign funding, and co-authoring with partners located overseas. As will be explained later, arguably each type of collaboration yields different effects on local research. This is one of the issues investigated in this study.

In contrast to the literature on the characteristics and on the determinants of research collaboration and of international research collaboration, the literature on the impacts of international research collaboration on research performance is rather scarce. Fortunately, that related to the effects of research collaboration without distinction of origin of the partners is abundant and it's helpful for better understanding the ways international collaboration affects research performance. Section 2.3 discusses the literature on the effects of research collaboration on research performance. Section 2.4 discusses the specific contribution this dissertation makes to current literature in the topic and introduces the research hypotheses relating to both research team productivity and research team orientation.

2.3 Research Collaboration and Research Performance

In the literature, research collaboration is mostly portrayed as an important enabler of science and technology development. It is considered to be 'better' than individualistic research in several respects. Many argue that research collaboration has greater epistemic authority (Wray 2002; Beaver 2004); facilitates diffusion of information and ideas; increases access to new knowledge and research tools; and offers visibility and feedback (Crane 1972; Beaver and Rosen 1979; Rigby and Edler 2005). These are crucial elements for the use and production of new knowledge and technology.

More importantly, most of the literature on the topic claims that research collaboration is an important source of creativity (Farrell 2001; Burt 2004; Levine and Moreland 2004; Uzzi and Spiro 2005), which in the right set of conditions may increase **a)** scientific productivity (Beaver and Rosen 1979; Landry, Traore et al. 1996; Adams, Black et al. 2005; Lee and Bozeman 2005; Turner and Mairesse 2005), **b)** research quality (Diamond 1985; Katz and Hicks 1997; Basu and Aggarwal 2001; Frenken, Hölzl et al. 2005; Rigby and Edler 2005), **c)** innovative capacity (Allen 1977; Georghiou 1998; Le Bas, Picard et al. 1998; Tsai and Ghoshal 1998; George, Zahra et al. 2002; Landry, Amara et al. 2002; Belderbos, Carree et al. 2004; Granovetter 2005), **d)** science and technology human capital (Coleman 1988; Rogers 2001; Rogers and Bozeman 2001; Seibert, Kraimer et al. 2001; Bozeman and Rogers 2002; Bozeman and Corley 2004), and **e)** help the consolidation of research agendas and the expansion of research areas.

Others, however, warn about the negative impacts of research collaboration on productivity (Fox and Faver 1984; Landry and Amara 1998; Carayol and Matt 2004b; Cummings and Kiesler 2005); output quality (Herbertz 1995; Kleinman 1998); innovative capacity (Gelijns and Thier 2002); human capital (Behrens and Gray 2001;

Stephan 2001; Slaughter, Campbell et al. 2002); and relevance of the research (Kleinman 1998; Florida 1999; Sagasti 2004; Shrum 2005). Risks and costs identified include the privatization and capture of traditional 'public' knowledge, the 'mercantilization' of knowledge and human capital as resulting from public-private research partnerships, opportunity costs, and crowding out effects.

The following is the literature found on the topic.

2.3.1 Research Collaboration and Creativity

Governments and institutions encourage or require the collaborative production of knowledge when scientists apply for funding because of the assumed positive effects this has on creativity. The mechanism through which collaboration increases creativity is little understood, however. While the literature on the virtues of external peer review on research quality is rather well developed (Cozzens, Popper et al. 1994), that related to the phenomena occurring within the collaborative process between partners is relatively new.

The issue is the object of study by sociologists, psychologists, economists, organizational theorists, and recently by policy scholars. Social capital and lately social network theorists have taken the lead in providing insights on role played by research collaboration on creativity (Granovetter 1973; Allen 1977; Coleman 1988; Fountain 1998; Nahapiet and Ghoshal 1998; Tsai and Ghoshal 1998; Farrell 2001; Laudel 2001; Seibert, Kraimer et al. 2001; Landry, Amara et al. 2002; Burt 2004; Granovetter 2005; Rigby and Edler 2005; Uzzi and Spiro 2005).

According to Granovetter (1973), individuals with a large number of "weak ties," that is, relationships with people from outside of their closest circle, are more likely to access information from distant parts of the social system and less likely to be confined to the provincial news and views of their close friends, placing them into an advantageous position in the market (Granovetter 1973; Granovetter 1983).

Allen (1977) claims that individuals with more contacts outside the organization ("gatekeepers") are advantageously situated for facilitating information flow and serve as the primary link to external sources of information and technology: a critical role for importing novel information and linking the organization with its environment (Allen 1977). Burt (2004), inspired by Mills (1848), claims that people connected with a greater diversity of groups are more familiar with alternative ways of thinking, which gives them more options to select from and synthesize, increasing their probability of having good ideas (Mills 1848; Burt 2004).

In fact, the impact of research collaboration on creativity is closely related to its impact on scientific productivity. How are they connected? The following discussion is based on what the literature says on the issue.

2.3.2 Collaboration and Research Productivity

From the policy point of view, one of the most important expected results of research collaboration is increased productivity. The idea that two or more heads produce more than one has implicit the assumption of efficiency resulting from the combination of skills needed to increase productivity. As Beaver (2001) reports citing one of his interviewees "[one] can put one student into the field for the summer, 3 months. After 5 years, [one will] have enough data to produce a research publication. A large research group can put 5 students in the field for the summer, 3 months. But in 3 months, the research group already has the data for a publication" (Beaver 2001). As the author adds, like the advantages (…) of parallel processing, one can parcel out parts of a problem, and finish more rapidly than one's competition.

However, empirical literature on the impact of research collaboration on research productivity is rather mixed. While some authors find positive effects on productivity as a result of division of labor (Landry, Traore et al. 1996; Lee 2004; Adams, Black et al. 2005; Lee and Bozeman 2005; Turner and Mairesse 2005), others find negative or no effects as a result of high transaction costs (McDowell and Smith 1992; Landry and Amara 1998; Slaughter, Campbell et al. 2002; Cummings and Kiesler 2005; Bonaccorsi, Daraio et al. 2006; Carayol and Matt 2006).

2.3.2.1 Positive Impacts

Landry, Traore et al (1996), performed an econometric analysis using survey data from Canadian academic researchers of all scientific disciplines and found that "collaboration, whether undertaken with universities, industries or institutions, may indeed increase researchers' productivity." According to the authors, the effect of collaboration on productivity varies according to scientists' field of research, however. Adams, Black et al (2005), who studied data derived from 2.4 million scientific papers written in 110 top U.S. research universities between 1981 and 1999, found that scientific output (as measured by paper publication) increases with team size. The authors conclude that "[s]ince increasing team size implies an increase in the division of labor, these results suggest that scientific productivity increases with the scientific division of labor" (Adams, Black et al. 2005).

Turner & Mairesse (2005) studied non-individual determinants of productivity by analyzing publications of 497 French physicists working at the Centre National de la Recherche Scientifique (NRS) over the period 1986-1997. They found that "the size of the laboratory has a small effect on individual productivity even though 'talented' [quotations in original] researchers seem more likely to be affiliated with larger labs." They measured productivity as the mean number of articles per researcher per year, the average impact factor and the mean number of citations to the articles (Turner and Mairesse 2005).

Lee (2004) studied the differences in performance of foreign-born and native-born scientists in the USA with data from 443 curricula vitae and a survey of scientists and engineers. He found that research collaboration (measured by the number of self-reported collaborators the respondent had, collaboration motives, research time to collaborate, cosmopolitan scale (quasi-geographical dispersion of collaboration), and the co-authorship pool) has a positive impact on productivity (measured by both normal of simple number of publications, and fractional counts, that is, dividing by the number of co-authors) of scientists.

According to Lee and Bozeman (2005), based on the curricula vitae and survey responses of 443 academic scientists affiliated with university research centers in the USA, publication count of peer-reviewed journal papers is strongly and significantly associated with the number of collaborators (Lee and Bozeman 2005).

2.3.2.2 Negative Impacts

Critics argue that high transaction costs in collaborative activities reduce research productivity. Katz and Martin (1997) claim that research collaboration also increases costs on travel, administration, and time spent on keeping all collaborators informed of the progress, deciding what to do next, developing new working relationships, resolving different opinions, and reconciling differences in management cultures, financial systems, rules on intellectual property rights, rewards systems, and promotion criteria.

Beaver (2001) identifies two main problems associated with research collaboration:

1. Principal Investigators lose touch with direct research: it may reduce creativity inspired by directly acquired tacit knowledge of how things work in practice; it may reduce the possibility of being a bench scientist; it may divert creative talents to administration and competition for limited resources.
2. Privatization of research is harmful to the research ethos: creation of entrepreneurial fiefdoms may promote negative strategies, especially secrecy or additional limits on the free sharing of ideas and materials in science; cooperation

with other laboratories (competitors) may be for purposes of cooptation or espionage, practices potentially harmful to science; even for the more positive purpose of alliance, competitive advantage may deter "smaller" laboratories or individuals.

Recently, empirical work has provided support to these claims:

Cummings & Kiesler (2005) investigated scientific collaboration across disciplinary and university boundaries to understand the need for coordination in these collaborations and how different levels of coordination predicted success. Their sample of 62 research collaborations supported by the US National Science Foundation in 1998 and 1999 showed that "[p]rojects with [principal investigators] from more universities were significantly less well coordinated and reported fewer positive outcomes than projects with principal investigators from fewer universities" (Cummings and Kiesler 2005).

Carayol & Matt (2006) analyzed the scientific research production of more than a thousand faculty members of Louis Pasteur University in France and found that the size of the lab affects negatively on productivity, as measured by fractional counts. According to the authors, researchers publish more when they are in smaller labs.

Negative effects associated with type of partner have also been reported. Slaughter, Campbell, et al. (2002) studied interview data from 37 science and engineering faculty members involved in university-industry relations in the USA and found that faculty face difficulties and tensions centered on intellectual property and restrictions on publication of research results when they work on industrial or corporate projects (Slaughter, Campbell et al. 2002)

Bonaccorsi, Daraio, et al. (2006) studied the Italian system of universities and found that collaboration with industry may improve productivity, but beyond a certain level the compliance with industry expectations may be too demanding and deteriorate the publication profile (Bonaccorsi, Daraio et al. 2006).

Landry, Traore et al (1996) found that scientists involved in collaboration aimed mostly at producing patented and unpatented products, scientific instruments, software and artistic production were less productive than their peers (Landry, Traore et al. 1996).

Similarly, there are empirical studies that report no meaningful effects.

2.3.2.3 No Relationship

Landry and Amara (1998) investigated the factors explaining why university researchers choose a given institutional structure when they engage in collaborative research projects using survey data from 1566 Canadian university researchers from the disciplines of engineering, natural sciences and health sciences. They found a trade-off between the capture of benefits measured in terms of additional publications and research funds and the coordinating costs of collaborative research (Landry and Amara 1998)

McDowell and Smith (1992) investigated the implications of academic promotions of the effect of gender-sorting on propensity to co-author of a cohort of 178 PhDs in economics from the top twenty institutions between 1968 and 1975. By analyzing their publications as registered by the American Economic Association's Index of Economic Articles, they found no significant effect of co-authorship on productivity (McDowell and Smith 1992)

Lee and Bozeman (2005) found that although (normal or simple) publication count of peer-reviewed journal papers is strongly and significantly associated with the number of collaborators, fractional count is not (Lee and Bozeman 2005).

Cummings & Kiesler (2005) found that "[p]rojects with principal investigators in more disciplines reported as many positive outcomes as did projects involving fewer disciplines."

Duque, Ynalves et al (2005), who examined the ways in which the research process differs in developed and developing areas, found that "collaboration is not associated with any general increment in productivity," the latter being measured by self-reported publication counts, and the former being measured by self-reported number of individuals the respondent worked with and the proportion of projects collaborated on by the respondents (Duque, Ynalvez et al. 2005).

Probably Beaver is right by claiming that "[a]t worst [research collaboration] doesn't influence, at best it enhances" (Beaver 2001). In fact, based on the literature, it seems that the effects of research collaboration on research performance depend on a set of mediating factors. These factors can be arranged into five groups as follows:

1. Factors related to the researchers' characteristics participating in the collaborative enterprise including **a)** age (Cole 1979; Diamond 1985; Levin and Stephan 1991; Stephan and Levin 1997; Dietz 2004; Smeby and Try 2005), **b)** sex (Fox and Faver 1985; Long 1992; Long, Allison et al. 1993; Prpic 2002), **c)** level of education (Becker 1964; Barro and Lee 2001; Bozeman, Dietz et al. 2001; David and Goddard L 2001), **d)** professional experience (Dietz 2004; Melin 2004), **e)** 'foreignness' (Lee 2004), and **f)** cosmopolitanism (Lee and Bozeman 2005);
2. Factors associated with the motivations for collaboration (Melin 2000), and the type of collaboration activities and strategies (Moed 2000);
3. Factors associated with the scientific discipline (Frame and Carpenter 1979; Becher 1981; Bauer 1990; Becher 1994; Landry, Traore et al. 1996; Bordons and Zulueta 1997; Qin, Lancaster et al. 1997; Okubo, Dore et al. 1998; Whitley 2000; Rinia, Van Leeuwen et al. 2002; Frederiksen 2004; Schummer 2004; Cummings and Kiesler 2005; Wagner 2005);

4. Factors regarding the type of partners involved, including **a)** sector of institution of affiliation (Landry, Traore et al. 1996; Etzkowitz and Leydesdorff 2000; Godin and Gingras 2000; Hagedoorn 2002; Cummings and Kiesler 2005; Frenken, Hölzl et al. 2005), **b)** localization or agglomerate (Saxenian 1994; Acs, de la Mothe et al. 1996; Landry and Amara 1998; Malo and Geuna 2000; Scott 2001; Liang and Zhu 2002; Stolpe 2002; Casper and Karamanos 2003; McKelvey, Alm et al. 2003; Zitt, Ramanana-Rahary et al. 2003; Bonaccorsi and Daraio 2005), and **c)** geographic and cultural proximity (Frame and Carpenter 1979; Narin, Stevens et al. 1991; Luukkonen, Persson; et al. 1992; Katz 1994; Leclerc and Gagne 1994; Landry, Traore et al. 1996; Cardinal and Hatfield 2000; Turner and Mairesse 2000; Liang and Zhu 2002; Mora-Valentin, Montoro-Sanchez et al. 2004; Wagner 2005; Waguespack and Birnir 2005); and
5. Public policies (Georghiou 1998; Georghiou 2001; Wagner, Brahmakulam et al. 2001; Smeby and Trondal 2005).

Some of this material is analyzed in the discussion of the factors affecting research performance.

Finally, to the author’s knowledge, no empirical work has been done on the effects of research collaboration on research orientation. In fact, that is one of the areas in which this dissertation makes its greatest contribution.

The following section, hence, discusses the contribution this dissertation offers to the understanding of the effects attributable to international research collaboration on research productivity and research orientation. This is done mostly by studying the case of a developing country while using research teams as unit of analysis in recognition of its importance as indicators and multipliers of local S&T capacity and, therefore, as key S&T policy targets.

2.4 Contribution this Dissertation Makes to Current Literature

This dissertation attempts to contribute to at least four research streams: research evaluation; sociology of science and technology; science, technology, and innovation policy in developing countries; and international relations and foreign policy. In fact, while the literature on the determinants and processes of research collaboration and of international research collaboration is relatively abundant and 'mature'[5], that on their impacts is rather rare and is still in its infancy[6]. New statistical tools and better information are contributing to its rapid evolution, however.

In this framework, and in contrast to the relatively extant literature found on the effects of research collaboration on research productivity, that on the effects of *international research collaboration* on the same variable is even scarcer. Not to mention the relative silence of the literature on the effects of international research collaboration on research *orientation*; and on the impacts on productivity and orientation in the context of a *developing country*.

2.4.1 Conceptual Framework

Based on the research collaboration literature, on recent literature on the effects of international research collaboration, and on the interviews done in the framework of this dissertation, several arguments can be proposed to explain the impact of international research collaboration on research performance in developing countries. These include

---

[5] See the work done in the framework of the Society for Social Studies of Science.
[6] See recent literature on the journals Scientometrics, Research Evaluation, and Research Policy.

arguments associated with the type of collaboration and the type of partner and their impact on both research productivity and research orientation.

2.4.1.1 International Research Collaboration, Creativity, and Productivity in Developing Countries

The literature on the effects of international research collaboration on research performance is rather recent, and similarly to the claims found in the literature on research collaboration, it arrives at contradictory results.

Turner & Mairesse (2005) analyzed publications of 497 French physicists working at the Centre National de la Recherche Scientifique (NRS) over the period 1986-1997 and found that the international openness of the laboratory positively influenced individual performance. They found that the accessibility of the technologies for experiments has a positive impact on productivity. Productivity is measured by the mean number of articles per researcher and per year, the average impact factor and the mean number of citations to the articles (Turner and Mairesse 2005). In contrast, Carayol & Matt (2004a and 2004b) found that the labs with more international collaborations did not have higher average publication performance (Carayol and Matt 2004a; Carayol and Matt 2004b).

Positive effects of international research collaboration on research productivity can be based on four arguments: a) the “more-is-better” argument, b) the “complementarity-based-on-diversity” argument, c) the “complementarity-based-on-similarity” argument, and d) the “linear-model” argument.

The “more-is-better” argument is the simplest and more commonly found in the literature. In the framework of this dissertation, this argument can be adapted to hypothesize that as long as a foreign researcher, a project is funded by a foreign institution, or a co-author located overseas is involved in the research process, more bibliographic outputs can be produced.

The "complementarity-based-on-material-diversity" argument is based on the literature in sociology of science and differs to the previous argument in the sense that it includes a qualitative criterion associated with the characteristics of the partner. In this framework, the greater the differences between the partners, the better, as in a collaborative enterprise everyone would offer something the other lacks and would get something would not be possible or easier to get otherwise. By collaborating with partners of different characteristics, one can get a better understanding of one's own problems by studying one's partners' problems and/or working on their solutions. By doing so, we complement our knowledge with that of our peers. In a sense, this is a variation to the "strength-of-weak-ties" argument proposed by Granovetter and Burt who claim that one has more to learn from those that see or have things one does not see or have, than from those of similar characteristics (Granovetter 1973; Burt 2004; Granovetter 2005).

Levine and Moreland (2004), for whom human cognition is an interpersonal as well as an intrapersonal process, claim that research collaboration increases creativity, particularly when it involves some degree of diversity, which may stimulate divergent thinking (Levine and Moreland 2004). Beaver (2001) claims that "multiplicity of viewpoints energizes and excites participants, makes actual work more intense and stimulates creativity." Research collaboration among members of different epistemic communities is one of the most important causes of the rapid progress in S&T in most developed countries, where "complex problems are better faced by teams appealing to multiple approaches in a process where each of the participants learns something new and sometimes unexpected from their colleagues" (Beaver 2001). As Fleming (2001) argues, the main function of R&D is indeed to generate new knowledge by recombining existing knowledge, and "when expertise is shared, it makes the sum stronger than the parts" (Fleming 2001).

The "complementarity-based-on-epistemological-similarity" argument is also based on the literature in sociology of science and also takes into account the characteristics of the partners. Based on this argument, a collaborative research is more productive when it involves partners that are compatible in many senses. This argument claims that for practical reasons, and to be successful in the research enterprise, one needs to work with partners with whom one shares similar paradigms, methods, views and values. It also draws from the literature that claims that personal empathy in terms of gender, age, social status, origin, language, ideology, experience, professional practice, professional ethos, religion, etc., is decisive.

As Levine and Moreland (2004) claim, similarity among partners may facilitate communication and interaction and by that means creativity: "[c]reativity in science, as in most other domains, involves more than simply generating a set of novel ideas (divergent thinking). It also involves narrowing this set to one alternative (convergent thinking) and then implementing this alternative by empirically testing and communicating it to the scientific community" (Levine and Moreland 2004). To Farrell, shared cognition, which constitutes the basis for research collaboration, implies a "shared set of assumptions about their discipline, including what constitutes good work, how to work, what subjects are worth working on, and how to think about them" (Farrell 2001).

The “linear-model” argument also claims positive effects of international research collaboration as it sees the collaborative process as an input-output process, where every collaborative input (foreign researcher or foreign funding) results in an S&T product. It differs from the “more-is-better” argument as it sees a more deterministic relationship between efforts and results.

Finally, several arguments can also be proposed to explain the negative effects of international research collaboration on research productivity based on the collaboration literature and on the opinions of the scientists interviewed. Hence, negative or no effects of international research collaboration can be attributed to the costs associated with the management of the collaborative enterprise. For the purpose of this dissertation, this is referred to as the “transaction-costs” argument. This argument contradicts the “more-is-better” argument as it claims that each additional researcher or funding source involved in the collaborative enterprise comes with a cost associated with it, which may affect research productivity.

Other arguments associated with the negative effects of the collaborative activity include the fact that sometimes partners collaborate without the intention to make public their findings (i.e the “inconvenience argument”), or that the lack of match between partners makes collaboration difficult and therefore unproductive.

To the author’s knowledge, current literature does not offer empirical support to most of these arguments. The use of a developing country as a case study to better understand the effects of international research collaboration on S&T capabilities seems to be better for this purpose than studying the effects of collaboration between developed countries, mostly because the differences between a developed and a developing country partners tend to be larger, which makes the assessment of impact or gains easier from the methodological point of view. This will allow testing the assumption that asymmetries lead to important gains for those in the seemingly disadvantaged position. This is the basis of the "diversity argument" discussed earlier.

Similarly, the study of the research collaboration pattern and effects in the context of a developing country can also contribute to the testing of the "similarity argument" as South-South collaboration mostly happens among neighbor countries sharing similar resources, views and problems (not to mention history, language, religion and culture characterizing most Latin-American countries).

In fact, besides the effects attributed to research collaboration as discussed earlier, international research collaboration can affect developing countries in a variety of ways. It can give local scientists and engineers access to new knowledge and research resources they would not have otherwise within their national boundaries (Wagner, Brahmakulam et al. 2001). It may raise the quality of the research performed in those countries, increasing the possibility for local scientists and engineers to benefit from the expertise brought about by international partners. These benefits can hardly be obtained in isolation from the global science and technology system.

However, international research collaboration can also increase their loss of autonomy and 'distract' local capabilities and critical mass needed to face local concerns, forcing them to address 'irrelevant' issues (Sagasti 2004). This is the topic discussed in the next section.

2 4.1.2 International Research Collaboration and Research Orientation in Developing Countries

The literature on the impact of international research collaboration on the orientation of the research performed is astonishingly silent. The reasons why there are few studies on the issue may be that, on the one hand, it is usually hard to define and account for the concept of "orientation" or "relevance" implicit in this variable. In fact, given the intrinsic characteristics of the scientific activity and its outcomes (it is a public good, it does not extinguish once it is used, etc.) it is hard to judge whether a specific contribution to knowledge is relevant or not. Questions such as "for whom?", or "when?" are often well grounded as there is no way to know whether what today is "irrelevant" is not going to be "crucial" for tomorrow's scientific development (Kuhn 1966).

Nevertheless, from the public policy perspective, the issue of "relevance" or "pertinence" is a matter of concern that has been around for a long time (e.g. Knowledge for what? (Polany 1962; Smith 1990)). Indeed, as any other human activity that typically demands large support from governments, the performance of science and technology activities are perceived to have the moral obligation to make effective contributions to the betterment of the societies that sponsor their activities (Cozzens 1999; Cozzens 1999; Cozzens, Bobb et al. 2005).

Regarding international research collaboration, the hope from the policy perspective is that local teams take advantage of the cognitive and material resources provided by their foreign partners to increase their contribution to the stock of local knowledge, hence increasing local S&T capacity to solve local problems.

This is considered to be particularly true in the case of developing countries, where local endowments of S&T capabilities are relatively scarce. This concern is consistent with the literature that sees knowledge as an opportunity for development, and "development as freedom" (Sen 2000; Cozzens, Gatchair et al. 2008). In this framework, the hope is therefore that by doing R&D activities in these developing countries, working on their own problems or using their countries as laboratories thanks to a collaborative activity with foreign partners will benefit their society and economy in the long run. The opposite may entail large opportunity costs.

In fact, if working on R&D activities in the framework of a collaborative activity is considered good for the developing country, working in their own country or using their country as the focus of their collaborative research should be considered as even better.

Hence, four arguments can be proposed to explain the effects of international research collaboration on research orientation. Arguments claiming positive effects include the "complementarity-based-on-epistemological-similarity" argument discussed earlier and the "commitment argument." In contrast, arguments claiming negative or no effects of international research collaboration on research orientation include the "opportunity" argument and the "outsourcing" argument.

Positive effects of international research collaboration may be based on the fact that we can get a better understanding of our problems by working on issues that are common to partners of similar characteristics in all relevant aspects (i.e. the "complementarity-based-on-epistemological-similarity" argument). Similarly, international research collaboration can also have a positive effect on research orientation in the sense that sometimes there might be bounds of some sort (contractual, personal, etc.) that leads to a commitment to work on local issues (i.e. the “commitment” argument).

In contrast, negative effects of collaborating with international partners on research orientation may be based on the existence of a relationship characterized by subordination. Foreigners may be interested in working with researchers and engineers from developing countries because of their calculations of the quality/price ratio (i.e. the “outsourcing” argument). In addition, researchers may be required to work on foreign issues because they do not have any other choice, or because they perceive in the collaborative activity an opportunity to work on issues of their own interest or expertise, which may not in turn be related to local issues (i.e. the “opportunity” argument).

Another reason why there are so few studies on the issue may not only be because it is risky to draw conclusions from, but also because it is materially hard to operationalize. However, the fact that it is hard to measure, and potentially misleading, should not be considered as a reason for not attempting to study it, as there is a real demand of information on that issue. Caution in its interpretation is needed, however. Fortunately, new and better information and software tools are increasingly making this task easier, allowing policy researchers to make useful contributions to the on-going and never-ending discussions on the topic in the S&T Policy arena.

To sum up the discussion presented here regarding the ways international research collaboration potentially affects research performance in developing countries, one can hypothesize that while collaborating with partners from the North positively affects team productivity because of its contribution in terms of material complementarity (i.e. the “diversity argument”), collaborating with partners from the South positively affects team orientation because of its contribution in terms of cognitive complementarity (i.e. the "similarity argument").

Taking this debate to the international relations literature, and from the perspective of a developing country, it could be the case that, if the "similarity argument" discussed earlier is right, international collaboration with partners from 'similar' countries (i.e. South-South collaborations) would have better effects than collaboration with partners with different characteristics (i.e. North-South collaborations). Such a finding would lead to a policy emphasizing South-South collaborations in developing countries.

Similarly, the choice of the collaboration strategy (hosting foreign researchers, working with foreign funding or co-authoring with colleagues located overseas) can also be supported by testing the arrangements associated with the gains and costs of each alternative. More on this will be discussed later.

Finally, another contribution this dissertation attempts to make refers to the use of research teams as the unit of analysis and policy targets to better account for the effects of international research collaboration on locals S&T capabilities. The next section discusses this choice.

2.4.1.3 Research Teams as Policy Targets and Unit of Analyses

As discussed earlier, the social organization of scientists into teams is today characteristic of most national science and technology systems (Gibbons, Limoges et al. 1994; Etzkowitz and Kemelgor 1998; Laredo 2003). Although researchers are generally members of an institute or department defined by discipline or thematic field, they work mostly in laboratories and within teams, and very often these teams result from working on projects that cut across administrative boundaries (Laredo 2003).

Their role as multipliers of S&T capabilities is what makes research teams an appropriate unit of analysis and focus of research and innovation policy (Laredo and Mustar 2001). However, this is rarely explicitly acknowledged in the public policy literature. Indeed, whereas the process of institutionalization of S&T as an indicator of local capacity has been implicitly recognized (Beaver and Rosen 1979; Gibbons, Limoges et al. 1994; Crow and Bozeman 1998; Etzkowitz and Kemelgor 1998; Etzkowitz and Leydesdorff 2000), their consideration as 'multiplier devices' or vehicles through which S&T capacities are created (Andrews 1979; Beaver and Rosen 1979) is still underdeveloped.

Research teams are particularly important as social structures within the system in the sense that they facilitate scientific and technological progress. They **a)** provide the framework necessary for interactive learning and creativity through the exchange of tacit knowledge and the sharing of resources and feedback among actors within a research system; **b)** facilitate the expansion of research areas of high levels of complexity as they allow research and education institutions to develop themes that would not be warranted in disciplinary units such as university departments and research centers designed to cover the full range of a discipline or a sub-field (Etzkowitz and Kemelgor 1998); **c)** have greater flexibility than departments or institutions in incorporating external influences into the research process; **d)** facilitate the performance of R&D projects, the internal coordination of tasks, the management of pooled resources, and external control (it is better to have one contract than a multiplicity of them); and **e)** may contribute to local innovation and economic development, particularly when they are incorporated into research clusters (Andersson, Serger et al. 2004) and regional systems of innovation (Saxenian 1994; Acs, de la Mothe et al. 1996; Holbrook and Salazar 2004).

As Etzkowitz and Kemelgor (1998) posit, “achieving a critical mass of research in distinctive fields is not only an essential part of the struggle to raise the status of an academic institution; it is also increasingly recognized as a path to local economic growth, initially through the research itself and then through the economic development that may be generated from that research” (Etzkowitz and Kemelgor 1998).

In general, policymakers are more interested in supporting research teams based on the conviction that the outcome of their combined effort outweighs the outcome obtained by summing up individual efforts. That is, the ‘whole-is-greater-than-the-sum-of-its-parts’ type of argument, an appealing one but hardly demonstrated empirically. More on this is will be discussed in the next chapter.

There are many reasons why some researchers like to work within such social structures. According to Landry and Amara (1998), these include: a) additional funding; b) additional equipment and facilities; c) additional information and data; d) additional resources; e) increased number of publications; f) increased number of innovations; g) improvement in the quality of teaching and training; h) more opportunities for students; and I) more networks of collaborators (Landry and Amara 1998).

In a developing country, in addition to the benefits of research collaboration mentioned, researchers benefit from team membership as it usually implies some level of formality and institutional support, which translates into labor stability and social recognition. More importantly, as Adam Holbrook posited “in Sri Lanka, where they tend not to operate in teams… [t]hey have a major outflow of human capital, possibly because there is no 'team-building,' no social structure created by the research effort there."[7]

However, besides the positive aspects of research teams mentioned, such structural arrangements also entail administration, coordination, and negotiation costs or, what economists call "transaction costs" (Williamson 1985; North 1990), that may affect productivity. The hope is that the gains in creativity and research quality outweigh these costs, however. In this framework, international research collaboration may not only increase creativity but it may also increase those costs. It seems therefore that teams' intrinsic characteristics are very important. In the author's opinion, these issues have not been satisfactorily raised in the current literature yet.

2.4.1.3.1 Determinants of Research Team Performance

The literature on the determinants of individuals' research productivity is abundant and it ranges from studies interested on the role of individual's characteristics to the role played by public policies. Babu and Singh identified more than 200 variables affecting individual's productivity (Babu and Singh 1998).

In contrast, the literature concerning the determinants of team productivity and of research orientation is indeed rare. Furthermore, the literature on the role international research collaboration plays on the various ways S&T activities are performed by teams is even rarer.

---

[7] Personal communication to the student, Vancouver, June 17th, 2006.

Arguably, the same way research collaboration affects research performance depending on a set of mediating factors including researchers' characteristics, international research collaboration affects local performance depending on teams' characteristics, including **a)** team size (Stankiewicz 1979; Cohen 1980; Cohen 1981; Qurashi 1984; Kretschmer 1985; Noltingk 1985; Cohen 1991; Qurashi 1991; Qurashi 1993; Kyvik 1995; Bordons, Gomez et al. 1996; Bordons and Zulueta 1997; Bordons, Zulueta et al. 1998; Landry and Amara 1998; Seglen and Aksnes 2000; Martin-Sempere, Rey-Rocha et al. 2002; Rey-Rocha, Martin-Sempere et al. 2002; Guan and Wang 2004; Adams, Black et al. 2005; Bonaccorsi and Daraio 2005; Lima, Liberman et al. 2005; Wang and Guan 2005; Bonaccorsi, Daraio et al. 2006); **b)** team age (Stankiewicz 1979; Cohen 1991; Landry and Amara 1998; Harrison, Price et al. 2002; Rey-Rocha, Martin-Sempere et al. 2002; Smeby and Try 2005); **c)** cohesiveness or empathy among team members (Harrison, Price et al. 2002; Martin-Sempere, Rey-Rocha et al. 2002; Hoegl and Proserpio 2004; Lima, Liberman et al. 2005); **d)** diversity or complementarity of skills (Ettorre 2000; Harrison, Price et al. 2002; Porac, Wade et al. 2004; Danilovic and Mats 2005; Waguespack and Birnir 2005); **e)** leadership; **f)** team reputation, and **g)** institutional recognition and support, among others.

Consistent with the collaboration literature, empirical work on the effects of team size and team age on team performance arrives at conflicting conclusions. Some authors claim that team size positively affects team productivity resulting from economies of scale, scope, division of labor, and complementarity (Adams, Black et al. 2005). Others claim negative impacts due to transaction costs (Bordons, Zulueta et al. 1998). Others claim no statistically significant effect (Cohen 1991; Seglen and Aksnes 2000). And yet others claim curvilinear effects showing positive effects up to a maximum number of team members after which productivity starts to decline (Qurashi 1991; Qurashi 1993).

To the author's knowledge, the effects of team size on research orientation have not been explored so far. In fact, it is not easy to hypothesize regarding such effects.

Regarding team age, the empirical literature shows that it either has positive effects (Harrison, Price et al. 2002; Rey-Rocha, Martin-Sempere et al. 2002) or no effects (Cohen 1991). The literature arguing positive effects claim that older teams are more stable in their process of knowledge production, 'marketing,' and publication, which may contribute to their cohesion, reputation and specific advantages hardly found in new teams.

Similarly, to the author's knowledge, the effects of team age on research orientation have not been explored yet. However, one could argue that throughout the years of experience, teams specialize on issues for which they can be considered as "unique" by their local clients, and therefore are demanded to work on issues of local concern.

The literature regarding the effects of the other variables identified as important for explaining team productivity and team orientation is even more silent. However, some plausible speculations can be made by extrapolating some of the effects identified in the literature on research collaboration (see literature reviewed above), while other hypotheses can be made based on mere speculation.

As PhD holders are often people able to make important contributions to team productivity given their knowledge, production skills, and experience gathered through their research careers, one would expect that the more PhD holders the team has, the higher its productivity. However, the effects of having doctorates may be negatively associated with team contribution to local knowledge, as most PhD holders working in developing countries obtained their title in a foreign country and therefore may be overspecialized and underemployed, which leads them to prefer to work with their foreign mentors or colleagues overseas. Arguably, the overspecialization or underemployment of PhD holders result from the lack of research resources available locally, which in turn may force PhD holders to work on projects of foreign interests.

Having many projects active is one of the characteristics of the more dynamic and more productive teams. And the more dynamic the team is, the greater the local demand it receives for working on R&D projects of local interests. Hence, the number of R&D projects active is positively associated with team productivity and team ability to contribute to local knowledge.

No plausible speculations on team productivity can be made related to the disciplines the team works in, except that, given the specialty of the teams working in the engineering one would expect that they produce relatively less bibliographic products than teams working in the other disciplines.

Regarding team orientation, one can arguably claim that teams working in the natural sciences are less like to work on local issues than teams working in the agricultural or social sciences or the humanities. The assumption is that, contrary to the former type of teams, teams working in the latter fields tend to be asked more frequently to solve problems typically related to the local everyday life. These are hypotheses to be tested, however.

Teams affiliated with firms tend to produce and publish fewer R&D products than teams affiliated with universities mostly because they are less interested in disclosing their findings, which may have strategic or commercial value. However, based on one of the critiques commonly raised regarding the role of the university in developed countries, one can hypothesize that teams affiliated with such academic institutions are more likely to work in universal concerns than in local issues as compared to similar teams affiliated with the other sectors.

Teams affiliated with institutions with large R&D budgets tend to have better ways of facilitating and persuading the research teams to be more productive. Similarly, internal competition for R&D funds in institutions with large R&D budgets may lead the teams to work on local issues as they may prefer local funding as opposed to foreign funding, which keeps them from having to look for funding elsewhere forcing them to work on issues they are not necessarily interested in. In this sense, competitiveness translates into autonomy or independence.

Extant literature claims that localization also matters for explaining research productivity. The reason is that big cities or agglomerates offer researchers large opportunities and resources to be more productive. In fact, the idea of 'scientific districts,' 'clusters,' 'science parks,' 'technopoles,' or 'poles of excellence' has been the focus of S&T policy for at least two decades, since the publication of an influential study done by Saxenian regarding the Silicon Valley and Route 128 (Saxenian 1994). The idea portrays the co-presence and interaction of diverse actors including higher education and research institutions, firms, government agencies, financial services, technology transfer facilitators, and other intermediary organizations (Acs, de la Mothe et al. 1996). Behind this idea is the assumption that proximity favors spillovers that can be translated into increased scientific productivity. Personal interaction, on-site demonstrations, and transfer of tacit knowledge, are enablers of creativity, productivity and innovation (Saxenian 1994; Acs, de la Mothe et al. 1996; Landry and Amara 1998; Malo and Geuna 2000; Scott 2001; Liang and Zhu 2002; Stolpe 2002; Casper and Karamanos 2003; McKelvey, Alm et al. 2003; Zitt, Ramanana-Rahary et al. 2003; Bonaccorsi and Daraio 2005). We can extrapolate these findings to explain team performance.

The ways team location may be related to team orientation can also be explained based on the same grounds as teams located in big cities benefit more from local capabilities, better quality of information, attract more students, have greater access to local contracts and, as a consequence, are better placed for working on local problems.

Finally, there are other factors much harder to observe that can also help to explain both team production and their ability to contribute to local knowledge. In fact, a combination of diversity (Ettorre 2000; Porac, Wade et al. 2004; Waguespack and Birnir 2005), division of labor (Adams, Black et al. 2005), cohesion among partners (Martin-Sempere, Rey-Rocha et al. 2002; Hoegl and Proserpio 2004; Lima, Liberman et al. 2005), leadership, institutional support, etc. seem to affect the creation of value within a collaborative context.

In sum, the study of the effects of the mediating variables identified here to explain the performance of research teams has not yet been explored empirically in the framework of international collaboration. This is one of the contributions this dissertation attempts to make.

2.4.2 Theoretical Model and Hypotheses

2.4.2.1 Theoretical Model

This work attempts to add to current understanding of the contribution of research teams to the creation of local S&T capabilities. Furthermore, it focuses on the role international research collaboration plays in that process. Arguably, as suggested above, whereas on the one hand the structure of teams implicitly brings the cohesion necessary between peers required for a positive performance, on the other hand international collaboration brings the complementarity needed for facilitating creativity, productivity, quality, innovative capacity, and relevance. International research collaboration may also open access to knowledge, provide resources allowing the team to engage students, and helps shape and strengthen the team's collective research agenda and orientation.

However, international research collaboration may also entail negative effects on team performance. It can decrease team productivity and detour team research orientation. Team characteristics, type of partner, and type of collaborative activity may affect the ways international research collaboration affects team performance.

Regarding the effects of different types of collaboration, one can hypothesize that when the collaboration implies hosting foreign researchers, it may contribute to local knowledge as these researchers probably work in local research teams in part because they are interested in local issues to which they are exposed to. This the "commitment" argument discussed earlier.

However, hosting foreign partners may affect team productivity as it may increase management and coordination costs. This is the “transaction-costs” argument explained above.

Based on the “linear-model” argument, one can hypothesize that working with foreign funding leads to greater team productivity. However, if we agree on the “opportunity” argument discussed earlier, team contribution to local knowledge may be negatively affected by the fact that foreign funded projects often are designed in funding countries, which may be interested on working on foreign issues more than on local issues. Funding countries may seek to rely on R&D capacity located overseas to meet their own research goals, which may lead to an outsourcing of local capacity and therefore to a sort of brain drain without mobility. Teams located in developing countries may see this as an opportunity but they may be force to work on foreign issues.

Similarly, co-authoring with colleagues located overseas may positively affect team productivity as it may add to local scientific capacity (i.e. the “more-is-better” argument) but it also risks of diverting local capacity to researching in foreign issues, therefore negatively impacting research team contribution to local knowledge (i.e. the “outsourcing” argument).

Regarding the effects of collaborating with different types of partner, and extrapolating the research collaboration literature discussed earlier, one can hypothesize that, based on the “complementarity-based-on-material-diversity” argument” teams that collaborate with partners from the North are more productive than teams collaborating with partners from the South as the former partners tend to have more to offer in terms of materials and research experience than partners from the same country or partners from southern countries.

However, when considering the effects on team orientation, the similarity among partners may have better impacts. In this sense, partners from the South may be more interested on local issues than local partners and than partners from the North as they may learn more how to solve their own problems from their partner's experience. In fact, southern countries share many common characteristics such as history, climate, natural resources, language, traditions, etc., as opposed to what they have in common with their northern partners. This is the "similarity argument" discussed earlier. In this sense, cognitive and epistemological complementarity resulting from collaborating with partners from southern countries contributes to team capacity to contribute to local knowledge.

Inversely, given the marginal role developing countries play in global research streams led by northern countries, their relatively weak negotiating capacity leads them to engage in projects of their partners' interests more than on their own interests. This is the "outsourcing" argument discussed already.

Figure 3 summarizes the theoretical model proposed.

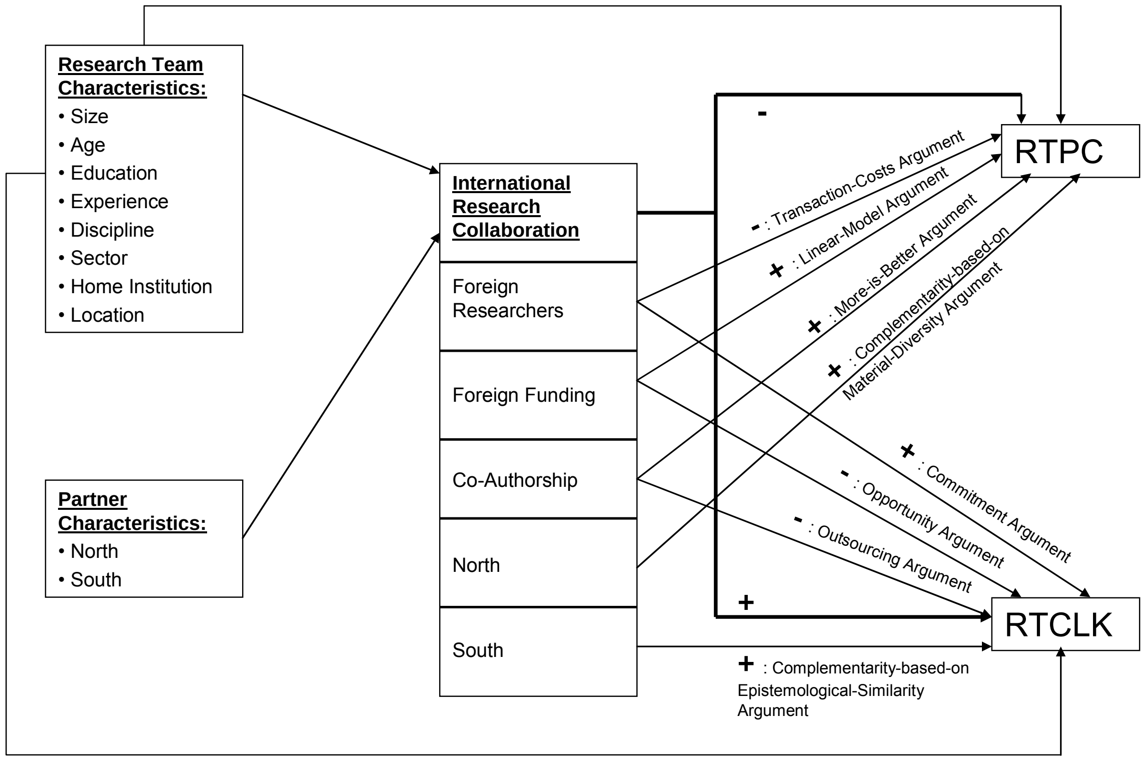

Figure 3: Theoretical Model

2.4.2.2 Summary of Hypotheses

In previous sections we saw that a) there is no consensus on the characteristics of the effects of research collaboration on research performance; b) there is little guide provided by the literature on the ways and extent to which international research collaboration affects research productivity; and c) there is even less information as to how different types of collaboration affects research productivity and research orientation. For these reasons, as explained earlier, some of the hypotheses proposed are exploratory and intuitively based.

The hypotheses proposed include, on the one hand, the relationships between the three types of international research collaboration discussed and the productivity and orientation of the research teams, and on the other, the relationship between different types of partners and team performance.

2.4.2.2.1 International Research Collaboration and Research Team Output

The following are the hypotheses regarding the overall effects of international research collaboration; of the different types of collaboration considered (co-authoring with partners located overseas, working on projects with foreign funding and hosting foreign researchers); and of different types of partners: North and South):

H1. International research collaboration (IRC) positively affects team productivity in Colombia. This hypothesis is grounded on the literature that claims that research collaboration facilitates access to materials, financial resources, new knowledge, and relevant information, and that by so doing it increases creativity and productivity.

H2. Hosting foreign researchers reduces team output. This hypothesis is based on the "transaction costs" argument discussed earlier.

H3. Receiving foreign funds to support R&D activities increases team output. The reason supporting this hypothesis is apparent as foreign funding usually implies the elaboration of research products. For our purposes this is the "linear-model" argument.

H4. Co-authoring with foreign partners located overseas increases overall team output. This hypothesis is based on the "more-is-better" argument.

The origin of the partner may also have mixed effects on the performance of the teams. Whereas research teams may have more to learn and may gain more access to scarce resources by collaborating with researchers from the North, they may gain more understanding of local issues when collaborating with researchers from the South, who have similar research questions, approaches, resources and historical background.

H5. Teams that collaborate with partners from the North have more bibliographic products. This hypothesis is based on the "complementarity-based-on-material-diversity" argument.

Finally, there might be combined effects regarding partner origin and type of collaboration.

H6. Working with projects funded by foreign institutions increases team output more for teams that collaborate with northern countries than for those that collaborate with partners from the south. This implies the combination of two positive effects: the effects derived from working on projects with foreign funding (the "linear-model" argument), and the effects derived from the "complementarity-based-on-material-diversity" argument.

2.4.2.2.2 Hypotheses on Research Team Ability to Contribute to Local Knowledge

The hypotheses associated with the effects of international research collaboration both considering the different types of collaboration and partners include:

H7. Teams that collaborate internationally are less likely to use 'Colombia' in their research activities.

H8. Hosting foreign researchers increase the probability of teams to involve 'Colombia' in their research. This is based on the "commitment argument."

H9. Receiving foreign funding reduces the probability of teams to work on research activities involving 'Colombia'. This hypothesis is based on the "opportunity" argument.

H10. Co-authoring with foreign partners located overseas reduces the probability of teams to work on research activities involving 'Colombia'. This hypothesis is based on the "outsourcing" argument.

H11. Teams that collaborate with partners from the South are more likely to use 'Colombia' in their research projects and products. This hypothesis is based on the "similarity argument."

H12. Working on projects funded by foreign institutions decreases the probability of using Colombia in their research more for teams that collaborate with northern countries than for those that collaborate with partners from the south. This implies the combination of the two effects discussed earlier: the negative effects derived from working on projects with foreign funding, and the negative effects derived from the "outsourcing argument."

Table 2 summarizes the hypotheses discussed in this chapter. The next chapter introduces the operation definitions, the data and the models used.

Table 2: Summary of Research Hypotheses

Variable	Research Team Output	Research Team Contribution to Local Knowledge
Internat. Res. Collab.	> than No Internat. Res. Collab.	Less likely than No- Internat. Res. Collab.
Foreign Researchers	< than No Foreign Researcher	More likely than No Foreign Resear.
Foreign Funding	> than No Foreign Funding	Less likely than No Foreign Funding
Internat. Co-Author	> than No Inter. Co-Author	Less likely than No Internl. Co-Author
Partner from North	> than No Partner from North	Less likely than Partner from South
Fore. Fund from North	> than Foreign Funding from South	Less likely than Foreign Funding from South

CHAPTER 3

METHODOLOGY

This dissertation attempts to answer the following basic research question: What is the role of international collaboration on the development of Colombian S&T capabilities as reflected in the performance of research teams?

To answer to this question this chapter presents the operational definitions, the data, and the models used.

3.1 Operational Definitions

Three key definitions were operationalized: Science and Technology Capabilities, International Research Collaboration, and Research Team.

3.1.1 Science and Technology Capabilities

S&T Capabilities are defined in this dissertation by the revealed ability of research teams to produce and disseminate knowledge and to contribute to the study of issues that may be of local interest. In this framework, Research Team Output is measured by the team's revealed productivity, that is, their production of journal articles, books, book chapters, proceedings, working papers, and other bibliographical products done by the teams during the period observed (See the list of bibliographical products in Appendix A)[8]. Research Team Ability to Contribute to Local Knowledge (also referred in this

[8] All these products are given equal weight in the econometric analysis presented. Although this may be seen as problematic, the reason is that we are interested more on the scientific capacity of the teams to

dissertation to as research “orientation” or research “relevance” or “ability to work on Colombian issues” or “ability to work on local concerns,” etc.) is measured by the extent to which teams work on R+D projects and/or write journal articles that take into account Colombia either as unit of analysis, as ‘laboratory,’ or as the focus of their research activity. This is observed by the use (or lack thereof) the word “Colombia” in the title of their research projects or journal articles, or in their corresponding abstracts.

3.1.2 International Research Collaboration

The operational definition of International Research Collaboration (IRC) is threefold: IRC as co-authored work, IRC as foreign researchers affiliated with Colombian research teams, and IRC as foreign funding to team R&D projects.

1. IRC as co-authored work that involves at least one researcher with a contact address outside Colombia. Although co-authorship (local or international) is criticized for failing to capture the real breath of collaborative activities that do not end in publications, or because it counts as collaboration ‘honorary co-authors’ with little real contribution to the collaborative output (Harsanyi 1993; Katz and Martin 1997; Moed 2000; Laudel 2002; Cronin, Shaw et al. 2004; Yoshikane and Kageura 2004), it remains amply accepted, mostly because it eases analysis and does imply a relatively high level of actual collaboration (Beaver and Rosen 1979; Melin 1996; Bordons and Gomez 2000; Beaver 2001; Newman 2004). As Bordons and Gomez (2000) posit, defining research collaboration as

---

produce knowledge than on the quality or the relevance of their products themselves. The issue of relevance is analyzed differently here.

co-authorship eases bibliometric analyses, which are more reliable and offers several advantages when compared with other methods such as interviews or surveys. The reliability of bibliometric results can be verified by means of repeated analysis (Bordons and Gomez 2000), and its techniques enable the analysis of large amounts of data, producing cost-effective and statistically significant results (Cole 2000; Van Raan 2000).

2. IRC as foreign researchers affiliated with Colombian research teams. Given the limitations of co-authorship as a proxy for research collaboration, the extent to which a team has a member from another country is considered to better account for international research collaboration. Foreign students are excluded from this category.

3. IRC as foreign funding to R&D projects. For a developing country like Colombia, this source of collaboration is vital. Indeed, as could be established by the interviews with scientists in the framework of this dissertation, many research teams exist in Colombia thanks to foreign funding. Sometimes funding involves foreign researchers as well, and if successful, ends in co-authorship, but this is not always the case.

In addition, two types of partners are considered: partners from northern countries and partners from southern countries. Appendix B provides the classification used to operationalize these two types.

3.1.3 Research Team

For the purpose of this dissertation, a Research Team is defined as a) two or more people who claim they work together on common research problems or interests; b) are recognized by their home institution and Colciencias as such; c) work on at least one R&D project active during the period of observation; and d) produce research outputs jointly or independently that are attributed to the team's work. People affiliated with the team producing 'certified' research outputs or working as technicians are counted as part of the 'core' research team. Certified research outputs include research articles, literature reviews, books, book chapters, software, technical pilots, technical projects, prototypes, industrial designs, technical norms, masters theses directed, and PhD dissertations directed (Colciencias 2000a). Taking policymaking and managerial considerations, and based on the principle of national sovereignty, the focus of this dissertation is on the 'national' S&T system as represented by a set of research units affiliated with institutions located in a single country, in this case, Colombia. This, of course, does not deny the essentially international character of modern science and technology, a basic characteristic of its implicit universality. It also does not impose an artificial boundary since developing countries tend to be more 'locally-constrained' than developed countries which are more internationally-oriented (Wagner and Leydesdorff 2005).

The next section presents the data sources used to support the analyses done in the following chapters.

3.2 Data Sources

Cross sectional data of Colombian research teams containing information on their characteristics and activities performed between 2003 and 2005 comes from three data sources:

1. One, called GrupLac, registers information on the Colombian research teams as reported at the end of September 2005 by the team director to the Colombian Institute for

the Development of Science and Technology (COLCIENCIAS) within the framework of the national team registry and ranking policy. The dataset contains information on teams affiliated with four different types of institutions (Universities, Firms, Government Labs, and NGOs), working in six different disciplines (Natural Sciences, Agriculture Sciences, Medical Sciences, Engineering, Social Sciences, the Humanities, and Other Sciences), located in three different types of locations (Big Cities, Small Cities, and Mid-Size Cities); and affiliated with three different size of institutions (Big Institutions, Mid-Size Institutions, and Small-Institutions).

2. A second database, called CvLac, contains information on the individual team members and on their activities as reported before the end of the same month by the researchers in the framework of the same policy.

COLCIENCIAS has gathered these data periodically in electronic databases since 2000 as part of its strategy for encouraging research quality. Both the ranking and the registry are produced every two years around September and are designed to support funding decisions and encourage community interactions. Some local research universities, government institutions, and foreign funding agencies use this information to support their own funding decisions as well. These data are freely accessible through the Internet, and it is mostly updated every two years near the month of June by participating teams [9].

---

[9] www.colciencias.gov.co Look for the 'Scienti Platform' link in the institution's main webpage.

Team leaders and individual researchers complete these electronic surveys with information both about themselves, their group, and their R&D-related activities they perform. For the purpose of this study, the information used for the analyses refers to the activities performed between 2003 and 2005 by the teams as errors in reporting activities done the years before are more likely to occur due to potential lack of memory of the respondents of the 2005 survey.

Thus, based on these data, the total number of teams registered in 2005 was over 3340. This includes 'teams' with one member, teams created the same year the call was made, and teams reporting no R&D project active anytime between 2003-2005, however. As we can hardly say that there are teams of only one member or teams with no projects keeping its members working on common issues, and as it is likely that some 'teams' were created only to respond to Colciencias' call, we drop these so-called teams and base the analyses on the remaining 1889 research units that meet those criteria (see the sampling strategy in Appendix C).

These databases were used in the construction of one of the dependent variables analyzed here (team's bibliographic production), and provided part of the information used to create the second dependent variable (teams' ability to contribute to local knowledge). These databases were also used in the construction of the control variables related to the characteristics of the research teams (team size, team age, number of PhDs, number of projects, scientific field, sector, size of home institution, and size of city where the team is located), and in the creation of two of the independent variables related to their collaborative behavior (foreign researchers working at local teams and foreign funding to

R&D projects) and of part of the variables related to the origin of the partners (North and South).

The databases include information on 819 foreign researchers associated with 541 Colombian research teams working between 2003 and 2005. Although this directory does not include the information on all the researchers of foreign origin working in Colombia, one can argue that, given its unique characteristics described above, it does record more than 90% of them.

The same data sources provide basic data on 1902 projects in all scientific fields reported by the Colombian research teams that were funded by foreign institutions. Again, although they cannot be taken as comprising the universe of R&D projects done in Colombia with foreign funding during the period observed, they arguably represent the majority given the unique characteristics of the directory created by Colciencias[10].

3. A third data source comes from a query on the Web of Science (WOS)[11] comprising the references of 5491 journal articles published between 1998 and 2005 in all fields by researchers located in Colombia. This data source was not only useful for analyzing the characteristics and implications of the globalization process of the Colombian S&T community but also served for the construction of the second dependent variable (team contribution to local knowledge), the third type of international

[10] In deed, most of the research teams are registered at Colciencias as this registry is used as reference in the accreditation process performed periodically by the higher education institutions. Similarly, both local and international funding institutions rely on this registry to make funding decisions. It is then on the teams interest to report their ability to access foreign funding as that speaks well of them both for the accreditation process and their credibility in a competition for funding. This is particularly important for teams working in developing countries where most of their funding comes from extramural sources who, in turn rely on this registry.
[11] The Web of Science is a set of databases administered by the Institute of Scientific Information (ISI) and Thomson that indexes the most important journals in the world. It includes more than 8,700 journals indexed by the Science Citation Index, the Social Science Citation Index, and the Arts and Humanities Citation Index.

collaboration studied here (co-authorship), and the origin of the foreign partner. Based on this data source, international research collaboration is observed by looking at the authors' address. When an article includes an author from a country other than Colombia, this article is assumed to be the result of international research collaboration. For the purpose of this research, those reporting addresses in Colombia are assumed to be Colombian nationals, and those reporting addresses outside Colombia are assumed to be foreigners[12].

Not all scientific production done in collaboration by the Colombian S&T community is registered by the ISI's databases, however. In fact, there is no a single satisfactory way of observing international co-authorship of articles written by Colombians. Although the database SCOPUS indexes twice the number of journals as the WOS does, a preliminary analysis not shown here comparing both databases yielded that more than 95% of the international articles indexed by SCOPUS were also indexed by the WOS. Since the information required for the analyses intended was more complete in the WOS, its databases were chosen to perform the study shown here. The software VantagePoint developed at Georgia Tech and administered by Search Technology supported the analyses done with this dataset.

Table 3 presents the variables and the corresponding data sources used.

Some transformations of the information contained in the three databases were necessary to build the two datasets that supports this dissertation (See Appendixes B, D, and E).

---

[12] In fact, according to data on research teams from CvLac for 2005, less than 6% of researchers working in Colombia were from a foreign country. There is no information about the number of Colombian doing research overseas, however.

Table 3: Variables and Data Sources

Dependent, Independent, and Control Variables and Data Sources*

	Variable	Type of Variable: Count	Interval	Dummy	Data Source: GrupLAC	CVLAc	WOS
Dependent Vars.	Bibliographic Products	x			x		
	Keyword Colombia			x	x		x
Indep. Vars: IRC	Foreigners in Teams			x		x	
	International Funding			x		x	
	Co-Authorship with Internat Partner ***			x			x
Research Team Characteristics.	Team Size		x		x		
	Team Age		x		x		
	Researchers with PhD	x				x	
	Leader Studied Overseas**			x		x	
	Leader Speaks Other Language **			x		x	
	R&D Projects active		x		x		
Scientific Filed	Team 1st. Scientific Field			6x	x		
Home Institution	Sector of Operation			4x	x		
	R&D Size of Institution of Affiliation			3x	x		
Localization	CitySize of Team			3x	x		
Partner Charac.	Location of Partner****			2x			x

* Cross Sectional Data observed for the period 2003-2005
** Used for predicting International Research Collaboration only
*** Observed for 2001 and 2002 only
**** Used for those who collaborated internationally only

3.3 Interviews

In addition to the data described, and to better understand the determinants of international collaboration and the dynamics through which it creates its effects, 6 interviews with team members and 3 interviews with experts on Colombian research policy were conducted in the framework of this dissertation. The purpose of the interviews to the team members was to inquire about the motivations, ways, and results of their choice, that is, to collaborate or not to collaborate internationally, what factors affect their decisions, and how they perceive these factors to influence the way collaborating or not collaborating internationally affects team performance. In addition, for the purpose of verifying their ‘real’ existence as ‘teams’, a different set of questions related to team-working activities were asked to them.

The interviewees were selected based on their role in the research team, the discipline they specialized in, the institution they are affiliated with and the city where they live. Hence, 4 out of the 6 interviewees were team leaders; 1 of the interviewees works in health sciences, 1 in natural sciences, 1 in agricultural sciences, 1 in social sciences, 1 in the humanities, and 1 in engineering. 2 of the interviewees had never worked in collaboration with a foreign partner. 3 work in Bogota, 1 in Medellin, 1 in Cali and 1 in a small city.

These interviews added to the quantitative analyses intended in the sense that they were designed to provide complementary information to better understand the determinants, drivers, barriers, enablers, processes and impacts of international research collaboration on research team performance in Colombia. They inquired about motivations for collaboration, past experiences, main activities, roles and intangible results, as these were aspects not covered by the data available. These interviews were conducted by appointment after informed consent was obtained following the Georgia Tech’s IRB Guidelines for the protection of human research subjects[13]. The interview protocol is found in Appendix F.

On the other hand, the purpose of the interviews conducted to the experts on the local research policy was to assess the suitability of the model; the accuracy of the variable construction process; the plausibility of the findings and the correctness of their

[13] www.osp.gatech.edu

interpretation; the consistency, relevance and viability of the policies recommended, resulting from the findings obtained; and the generalizability of the conclusions to other countries they are familiar with. No structured interview protocol was used in this case, and the selection of the experts was based on the author's opinion regarding their professional and academic trajectory on the issues under study.

3.4 Models

This dissertation answers the following research questions: a) does collaborating internationally positively affect team performance? b) How does each type of collaboration impact team performance? and c) how different is the effect of collaborating with Northern and Southern countries?

Empirical studies already exist on the effects of research collaboration. The literature reviewed discusses the work done on the relation between research collaboration and research productivity. To the author's knowledge, no work has been done regarding the effects of international research collaboration on research team productivity and research team orientation in a developing country.

The difficulty of this kind of analysis is twofold: selection bias and endogeneity of research collaboration. Selection bias results as there is no random 'assignment' of teams to the 'treatment' group. In practice, international partners might collaborate only with those teams and in those R&D projects that are expected to generate new knowledge and technologies. For this reason, the inclusion of research collaboration in a linear regression will cause endogenous effects, which would lead to inconsistent and biased estimates if it is correlated with the error term. To estimate the "real" effect of international research collaboration, it is therefore necessary to address the basic question: How would the teams with international research collaboration have performed had they not collaborated with international partners? To the author's knowledge, no study on the impact of research collaboration, whether local or international, attempts to model this counterfactual situation.

Most of the studies surveyed do not pay attention to this kind of bias. The only exception found is the study by Lee and Bozeman (2005), which analyzed the effect of research collaboration on the productivity of 443 academic scientists in the USA, and controlled for reverse causality by using a 2SLS using cosmopolitanism as the instrumental variable (Lee and Bozeman 2005). Their work does not consider counterfactuals and comparable control groups to assess impact, however.

Two approaches are used to answer the research questions stated here: hypothesis testing using multiple regression models, and impact assessment using Propensity Score Matching. The following sections describe the ways these two approaches were used in this study.

3.4.1 Hypotheses Testing Using Multiple Regression Models

Between 2003 and 2005, 39% of the teams either hosted foreign researchers or worked on projects with foreign funding. These teams were on average larger, older, had more members with PhDs, were more active (as measured by the average number of projects done during the period observed), were more likely to work in the natural sciences, were less likely to work in the social sciences, were more likely to be associated with the non governmental sector, were less likely to work at universities, worked at bigger institutions (as measured by their R&D budget), and were more likely to be located in big cities in 2003 than the teams that did not collaborate internationally.

Table 4 shows the mean values of the interval level variables and what percentages have the value 1 for the dummy variables for collaborating and non collaborating teams (IRC and Non-IRC respectively). The third column shows the IRC and Non-IRC differences, with asterisks showing whether the differences are statistically significant at conventional levels. While collaborative activities and bibliographic production were observed for the years 2003 to 2005, team internal and overall characteristics were observed for up to 2003.

Table 4: Team Characteristics and Performance by Collaboration Status 2003-2005

	Variable	IRC	No-IRC	Difference
Internal Characteristics	Team Size in 2003	9.07	6.1	2.97***
	Team Age in 2003	8.85	5.67	3.18***
	Has PhDs in 2003	79%	46%	33***
	Total PhDs in 2003	2.4	0.85	1.55***
	Total Projects in 2003	8.22	4.18	4.04***
S&T Field	Natural Sciences	30%	19%	11***
	Agriculture Scs.	6%	6%	0
	Medical Sciences	13%	12%	1
	Social Sciences	12%	19%	-7***
	Humanities	25%	22%	-3
	Engineering	12%	14%	-2*
	Other/Multidisciplinary	5%	4%	1
Sector	Education Sector	88%	92%	-4***
	Business Sector	4%	3%	1
	Government Sector	4%	4%	0
	Other Sector (NGOs)	3%	1%	2***
Home Institution	Small Institution	14%	18%	-4***
	Mid-Size Institution	29%	44%	-15***
	Big Institution	57%	38%	19***
Location	Small City	1%	3%	-2***
	Mid-Size City	16%	25%	-9***
	Big City	82%	71%	11***
Performance	Tot. Bib. Prods 2003-5	13.68	5.79	7.89***
	Keyw. 'Colombia' 2003-5	47%	29%	18***
	Sample Size	736	1,153	

IRC & Non-IRC differences significant at:
* 0.1 level, ** 0.05 level, *** 0.01 level

More importantly, based on the data, those teams that collaborated internationally were more productive, as measured by the average number of bibliographic products done between 2003 and 2005, than those that did not collaborate internationally. In fact, collaborating teams produced on average around 14 bibliographic products, while non-collaborating teams produced on average around 6 products. According to a t-test, the difference of almost 8 products is statistically significant at the 0.01 level.

Similarly, collaborating teams were more likely to use 'Colombia' in their projects and products than non-collaborating teams: while 47% of the teams that collaborated internationally used Colombia in their projects or bibliographic products, only 29% of the teams that did not collaborate internationally used the country as the unit of analysis or object of their research processes. This accounts for a difference of around 18% in the odds of involving 'Colombia' between collaborating and non-collaborating teams. A t-test shows that this difference is significant at the 0.01 level.

However, other differences between teams that did and did not collaborate internationally may account for some of the differences in team performance. In fact, the problem with concluding and making generalizations based on the bivariate analysis presented here is that it ignores potential overlapping effects among the factors that may explain team output. This can mislead the understanding of the true role international research collaboration plays in this process. In fact, by not controlling simultaneously for variables that could have independent effects on the outcome variable we may overstate the real effects of the observed variable. In fact, it may be the case that the teams that work in the natural sciences and collaborate internationally are more productive than the other teams because the former tend to have more members with PhD than the others, or because they tend to be affiliated with bigger institutions, or are older, or bigger, for instance. Multivariate regression models analyze the relationship between an explanatory variable and an outcome variable while controlling for the effects of other variables. For that reason, we need to perform a series of more sophisticated analysis aimed at taking into account the nuances that surrounds the ways international research collaboration affects team performance.

To better understand the 'true' role international research collaboration plays in the performance of research teams, we need to use the appropriate model, that is, a model that takes into account a) the variables that may have an explanatory power on the outcome of interest, and b) the specific characteristics of the dependent variable itself.

The selection of variables depends on both the factors identified in the literature reviewed and the availability of information. Hence, based on the literature discussed earlier, team performance may be a function of team size, team age, composition, experience, dynamism, scientific specialization, sector where it works, institution it is affiliated with, and its geographical location among other variables not easily observable, such as internal cohesion, institutional constraints, government support, etc.

As for the characteristics of the dependent variable is concerned, two regression models are used to account for the impact of international collaboration on team output and on teams' ability to contribute to local knowledge: a Zero-Inflated Negative Binomial and a Logistic Regression, respectively.

3.4.1.1 Test of the Research Hypotheses Associated with Team's Production Using Zero-Inflated Negative Binomial Regressions

Team productivity, the first outcome variables analyzed, has a frequency distribution highly skewed to the left where, between 2003 and 2005, many teams had zero or small number of products, while very few teams produced a large number of bibliographical products. Hence, with a distribution far from yielding a nice bell-shaped graph typical to a normal distribution, our dependent variable is an account of the existence of sporadic team outputs typical of a count variable, that is, a variable that indicates how many times something has happened, as opposed to a continuous-type of variables.

According to Long and Freese (2001), using the Linear Regression Model (LRM) - which is designed to fit a normal distribution- to account for the effects of a given set of independent variables on a count dependent variable produces coefficients that are biased, inefficient and inconsistent. As the authors posit "Even though there are situations in which the LRM provides reasonable results, it is much safer to use models specifically designed for count outcomes" (Long and Freese 2001). The analysis of the Poisson distribution of the team output as well as the process done to decide what model to use is shown in Appendix G.

From the methodological perspective, and as shown in the appendix, the selection of the econometric model dramatically affects the accuracy of the findings and therefore the plausibility of the conclusions that can be drawn from them. The Poisson regression model (PRM) improves prediction of counts dependent variables by fitting better the data observed. It reduces under-prediction of zeroes, and allows heterogeneity among sample members regarding their production rate, which, as we saw, is drawn from a Poisson distribution. However, a Poisson regression model does not take into account overdispersion in the outcome variable. For this reason, estimates are inefficient and the standard errors are biased downward, resulting in spuriously large z-values and spuriously small p-values.

A negative binomial model (NBRM) improves upon the underprediction of zeroes in the Poisson regression model by controlling for overdispersion. However, since it assumes that every single team has a positive probability of producing any given number of bibliographic products, the model fails to satisfactorily account for excess of zeroes. In the real world, not all teams are potential producers of bibliographic outputs. Teams working at or for industry, or by contract to government agencies may be discouraged to produce bibliographic products. Others may fail to report bibliographic products in a given period because they lack resources or motivation or as a matter of chance. Both types of teams appear as being non-productive, however.

Zero-inflated count models (ZIP and ZINB) respond to this issue and allow for the possibility of considering different causes of unproductivity by increasing the conditional variance and the probability of zero counts. These models allow distinguishing between potentially-productive and always-unproductive teams. After comparing the four count models analyzed (PRM, NBRM, ZIP, and ZINB) using several standard criteria and tests, it became apparent that the zero-inflated negative binomial model not only addresses assumptions that make substantive sense but it also fits the data observed remarkably well. The Vuong test results reported in the appendix show that the zero-inflated version of the Negative Binomial model is favored over the standard NBRM in this study: Vuong Test = 5.31 (p=0.000).

Thus, based on the literature reviewed, the characteristics of the dependent variable, and the different probabilities of teams to be unproductive, chapter 5 discusses the results obtained by the Zero-Inflated Negative Binomial Model (ZINB).

This model uses team's bibliographic production as the dependent variable. It is measured by the total number of bibliographic products done between 2003 and 2005. It uses international research collaboration as the independent variable, which is represented by a dummy variable coded 1 if the team had foreign researchers and/or foreign funding between 2003 and 2005, zero otherwise. And, as control variables, it uses team size (an interval-level variable for the number of researchers and technicians the team had in 2003); team age (an interval-level variable for how long the team had been in existence in 2003); the total number of PhDs (represented by an interval-level variable for the number of members with PhD degree the team had in 2003); team dynamism (measured by an interval-level variable for the number of R&D projects the team had active in 2003); scientific field (represented by six dummy variables, with teams working in the natural sciences as the reference group); sector (represented by three dummy variables, with teams working in the academic sector as the reference group); size of the home institution (represented by two dummy variables, with teams affiliated with big institutions as the reference group); and city-size (represented by two dummy variables, with teams located in big cities as the reference group). (See descriptive statistics for these variables in Appendix H).

The Breusch-Pagan / Cook-Weisberg test was implemented to test for heteroskedasticity in the data. The test found that the 'constant variance of error term' assumption was violated. Although heteroskedasticity does not affect the parameter estimates as the coefficients are unbiased, it does bias the variance (and, thus, the standard errors) of the estimated parameters as the coefficients tend to be underestimated, therefore inflating z-scores and sometimes making insignificant variables appear to be statistically significant. To solve this problem, the Huber/White/sandwich estimator of variance is used in place of the traditional calculation. Therefore, chapter 5 only discusses the robust estimation results.

3.4.1.2 Test of the Research Hypotheses Associated with Teams' ability to Contribute to Local Knowledge Using Logistic Regressions

A binary regression model (BRM) is the most appropriate model to account for the impact of international collaboration on teams' ability to contribute to local knowledge, our second dependent variable, as it is a dummy variable. As Lewis (2003) notes, using Ordinary Least Square (OLS) regressions for dummy dependent variables, which by definition have either values of 1 or 0, gives a linear probability model that violates the assumption of a normal distribution of the error term. That is, as the value of an independent variable changes, the variance of the error term for that variable would also change, leading to heteroskedasticity. In such case, the OLS estimators of the regression coefficients may be unbiased but cannot be efficient. Furthermore, estimates of the standard errors of the regression coefficients would be biased, distorting confidence intervals and hypothesis tests. Moreover, with OLS, the residuals would lead to meaningless expected probabilities such as negative probabilities since OLS assumes that the impact of marginal change of the value of an independent variable remains constant along all range of the values.

According to Lewis' course notes, to use robust standard errors or weighted least squares "do not solve all the problems, such as probabilities outside the range between 0 and 1, and therefore do not solve the conceptual problem of independent variables having constant impacts up to a certain point, then no impact beyond" (Lewis 2003). If we use logit or probit we do not have these problems, since the impact of marginal change in the independent variable becomes increasingly non-linearly smaller as the probability gets closer to 0 or to 1, yielding a probability distribution curve S-shaped[14].

In this dissertation a logit model is preferred over a probit model mostly due to personal preferences as there is no objective reason one would choose one versus the other. Indeed, as Lewis posits, their coefficients are "nearly linear transformation of each other" (Lewis 2003). They do not provide meaningfully different conclusions. The main difference is that while probit analysis uses the normal cumulative distribution function, logit analysis is based on cumulative logistic distribution function.

Therefore, the model estimated using the research teams data to account for the impacts of international research collaboration on teams' ability to contribute to local knowledge is a logistic model that uses a dummy dependent variable coded 1 if the team published a bibliographic product or worked in an R&D project whose title or abstract included the word "Colombia", 0 otherwise. The independent variable, international research collaboration, and the control variables are measured the same way as in the ZINB model explained before.

3.4.2 Impact Assessment Using Counterfactuals

An even better method for assessing the impact of international research collaboration on research team performance is by comparing teams of similar characteristics in all relevant aspects and, in particular, on the probability of collaborating

[14] The mathematical structure of binary models is not explored here, as a discussion on the statistical model will be out of scope of this dissertation. See Long, S. (1997) Regression Models for Categorical and Limited Dependent Variables. Thousand Oaks, CA: Sage Publications.

internationally. In fact, collaborating teams may be more productive for the same reasons they collaborate internationally. That is, foreign partners may prefer to collaborate with those teams and in those R&D projects that are expected to generate new knowledge and outputs. This is technically called ‘endogeneity’ or reverse causality, and it may result from selection bias.

Hence, to estimate the “real” effect of international research collaboration, it is necessary to address the basic question: How would the teams collaborating internationally have performed had they not participated in a collaborative experience with international partners? To the author’s knowledge, no study on the impact of research collaboration (whether local or international) has attempted to model this counterfactual situation.

To solve this endogeneity problem, comparable groups both in terms of their internal characteristics and particularly in their propensity to collaborate internationally is used.

The use of tools to control for selection bias and endogeneity using comparison groups and propensity scores is not new in the S&T policy evaluation literature. Klette, Moen, and Grilches (2000) provide a comprehensive survey on the ways public subsidies affect firm productivity, private R&D investment, patent applications, fixed-asset investment, returns on capital, returns on sales, and growth of sales or employment (Klette, Moen et al. 2000). Focusing on crowding-out effects, Almus and Czarnitzki (2003) investigated the average causal effect of all public R&D schemes in Eastern Germany using a nonparametric matching approach to pay attention to the possible interdependence between public R&D funding and R&D performance of firms. To do that, the authors compared the potential outcome of a group of subsidized firms to a matched control group of nonsubsidized firms (Almus and Czarnitzki 2003).

Busom (2000), considered the problem of selection bias by applying a two-stage econometric treatment model in which the first stage consists of estimating a probit model on the participation probability in public funding programs and in the second stage the R&D activity is regressed on several covariates, including a selection term that accounts for the different propensities of firms to be publicly funded. This second equation is estimated separately for participants and nonparticipants. The difference in expected R&D expenditure of both groups is according to this approach the result of public funding (Busom 2000).

A similar approach is followed in this dissertation using team output and teams' ability to contribute to local knowledge as the dependent variables. Team characteristics, sector of affiliation, scientific field, size of home institution, and size of city or region of location are the control variables. The key independent variable is international research collaboration.

3.4.2.1 The Impact Assessment Framework: Propensity Score Matching

The Propensity Score Matching (PSM) is an approach to estimate causal "treatment" effects (Caliendo and Kopeining 2008). The PSM is useful to overcome the fundamental question in every evaluation attempt and address the possible occurrence of selection bias. We would like to know the difference between the participant's outcome (say, team's productivity) with and without treatment (International Research Collaboration).

We clearly cannot observe both outcomes for the same teams at the same time. And taking the mean outcome for non-collaborating teams as an approximation is not advisable, since collaborating and non-collaborating teams usually differ even in the absence of 'treatment.' This is what the selection bias is all about. In fact, motivated teams, with strong support from their home institution, and led by someone with long experience and reputation have a higher probability of collaborating internationally and being productive than comparable teams. The matching approach, which simulates an experimental context, is one possible solution to the self-selection problem.

Since we are interested in how teams would have performed had they not collaborated internationally [$\gamma i = Yi(1) - Yi(0)$] as a way of assessing the impact of IRC, and since only one of the potential outcomes is observed for each team, we need to construct the unobserved outcome or counterfactual to know the population average treatment effect. To do this, we need to find in a large group of non-collaborating teams those teams who are similar to the participants in all relevant pre-treatment characteristics (*X*). Once that is done, differences in outcomes of this well selected, and thus adequate control group and of 'participants' or counterfactuals, can be attributed to the collaborative activity.

Since conditioning on all relevant covariates is limited in the case of a high dimensional vector *X*, Rosenbaum and Rubin (1983) suggest the use of the so-called balancing scores *b(X)*, that is, a function of the relevant observed covariates *X* such that the conditional distribution given *b(X)* is independent of 'assignment into a treatment'

(Rosenbaum and Rubin 1983). That balancing score is the Propensity Score: the probability of collaborating internationally given observed characteristics X[15].

To estimate the population's average treatment effects by matching teams on their propensity scores we need to focus on the treated, and more precisely on the Average Treatment Effect on the Treated (ATT) parameter, which is defined as:

$$\gamma ATT = E(\gamma \mid D=1) = E[Y(1) \mid D=1] - E[Y(0) \mid D=1]$$

Where $D=1$ for participating teams; and $Y(1)$ and $Y(0)$ are the outcomes of participating and non-participating teams respectively.

As the counterfactual mean for those being treated - $E[Y(0) \mid D=1]$ – is not observed we have to invoke some identifying assumptions. According to Caliendo and Kopeining 2008, one possible identification strategy is to assume unconfoundedness, that is, given a set of observable covariates X which are not affected by treatment, potential outcomes are independent of treatment assignment. This is called the Conditional Independence Assumption (CIA) and it implies that selection is solely based on observable characteristics and that all variables that influence treatment assignment and potential outcome simultaneously are observed. We'll come back to this rather strong assumption later.

Besides the CIA, an additional requirement is the common support or overlap condition, which rules out the phenomenon of perfect predictability of $D \mid X$.

---

[15] The decision whether to apply PSM as opposed to Covariate Matching is not discussed here. See Zhao, Z. (2004). "Using Matching to Estimate Treatment Effects: Data Requirements, Matching Metrics, and Monte Carlo Evidence." The Review of Economics and Statistics **80**(1): 91-107. for Mahalanobis distance used to calculate similarity of two individuals in terms of covariate values, where the matching is done on these distances.

As the probability of collaborating given a set of team characteristics falls between 0 and 1, the common support condition (CSC) ensures that teams with the same *X* values have a positive probability of being both collaborating and non-collaborating teams. In other words, the CSC ensures that any combination of characteristics observed in the treatment group can also be observed among the control group.

Since the CIA holds and there is overlap between both groups, the PSM estimator for the Average treatment Effect on the Treated (ATT) can be written as:

$$\gamma ATT,PSM = EP(X) \mid D=1\{E[y(1) \mid D=1, P(X)] - E[y(0) \mid D=0, P(X)]\}$$

Where $P(X)$ is the Propensity Score $P(D = 1 \mid X)$; and $D=0$ for non-participating teams.

In other words, the PSM estimator is the mean difference in outcome over the common support, appropriately weighted by the propensity score distribution of participants.

3.4.2.2 Estimation Using International Research Collaboration (IRC) as the Treatment of Interest

Since the propensity score represents the discrete 'choice' of collaborating (or not) internationally, a logit or probit model can be used to estimate it. Either a logit or a probit model is preferred over a linear model due to the shortcomings of the latter model in terms of the unlikeliness of the functional form when the response variable is highly skewed and predictions that are outside the 0,1 bounds of probabilities.

The estimation of the probability of collaborating internationally is discussed in chapter 4. It uses the same set of variables plus two new variables that help to explain the determinants of international research collaboration: a) a dummy variable coded 1 if the team leader writes well in a second language, 0 otherwise; and b) a dummy variable coded 1 if the team leader studied overseas, 0 otherwise.

All of the variables considered are both unaffected by participation and are not influenced by the anticipation of participation. For theoretical or empirical reasons discussed, they all influence simultaneously the participation decision and the outcome variable. They credibly satisfy the CIA and justify the matching procedure. Their impact on the probability of collaborating internationally is discussed in chapter 4. For the purpose of the analyses done in chapters 5 and 6, these variables are used to compute the propensity score for matching.

To contrast the productivity and probability of involving Colombia of a collaborating team with productivities and probabilities of comparison group teams, the kernel matching algorithm is used[16]. The Kernel matching is a non parametric matching estimator that uses weighted averages of all teams in the control group to construct the counterfactual outcome. As Caliendo and Kopeining 2008 note, citing Smith and Todd (2005), "kernel matching can be seen as a weighted regression of the counterfactual outcome on an intercept with weights given by the kernel weights." Weights depend on the distance between each individual from the control group and the participant observation for which the counterfactual is estimated. (…) The estimated intercept provides an estimate of the counterfactual mean" (Smith and Todd 2005; Caliendo and Kopeining 2008).

---

[16] The technical details of this matching algorithm are not discussed here. See Imbens, G. Ibid."Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review." **86**: 4-29. According to Caliendo & Kopeining, since the sample is large enough, it doesn't matter what matching algorithm is used. Kernel seems to be the most popular in the literature.

Both a kernel function and a bandwidth parameter are used. As Caliendo and Kopeining 2008 warn, the selection of bandwidth values have implicit the trade-off that a high value yields a smoother estimated density function, therefore leading to a better fit and a decreasing variance between the estimated and the true underlying density function. On the other hand, underlying features may be smoothed away by a large bandwidth leading to a biased estimate. The opposite is true: a small bandwidth may decrease bias but increase variance which means decrease efficiency. The bandwidth choice is therefore a compromise between a small variance and an unbiased estimate of the true density function (Caliendo and Kopeining 2008). To avoid the risk of using observations that are bad matches, the common support condition is imposed.

To perform the analyses, the STATA module developed by Leuven and Sianesi (2003) is used to estimate the full model and test the balancing hypothesis using an iterative process to ensure that the estimated model is consistent with this requirement (Leuven and Sianesi 2003).

Since we do not condition on all covariates but on the propensity score, the matching procedure is checked to see if it is able to balance the distribution of the relevant variables in both the control and treatment group. This is done by comparing the situation before and after matching to see if there are any differences remaining after conditioning on the propensity score.

Finally, the bootstrapping procedure is used to test the statistical significance of treatment effects and to compute their standard errors in case analytical estimates are biased or unavailable. Each bootstrap draw consisted in the re-estimation of the results, including the estimation of propensity scores, common support, etc. The bootstrapping was repeated 999 times, which led to 999 bootstrap samples and 999 estimated average treatment effects.

The next chapters discuss about the determinants of team performance, and particularly about how international research collaboration affects team productivity and team orientation. However, before exploring such effects, it is worth investigating about the factors that explain who collaborates internationally in Colombia and who does not.

CHAPTER 4

DETERMINANTS OF INTERNATIONAL RESEARCH COLLABORATION

Since the main purpose of this dissertation is to investigate the role international collaboration plays on team performance in Colombia, knowing the factors affecting the choice of collaborating internationally will help us to understand the phenomena under study. Furthermore, it will help us to better design policies aimed at creating local S&T capabilities through the encouragement of the internationalization of the Colombian S&T community, or at reducing the negative effects derived from that process.

4.1 Model Specification

As discussed earlier, among the 1889 teams studied, 736 collaborated and 1153 did not. What factors explain the collaborative behavior? To answer to that question the following model is tested using logistic regressions.

Pr(IRC05=1) = F(Team Size, Team Age, Total PhDs in 2003, Total Projects in 2003, Total Bibliographic Products in 2003, Leader Writes Well in a Second Language, Leader Studied Overseas, Scientific Field, Sector, Size of Home Institution, Size of City of Location)

Thus, the choice of collaborating internationally may be a function of team's characteristics. That is, larger teams are perhaps more likely to collaborate internationally than smaller ones as each team member may act as a collaborating agent: more agents equal more opportunities for collaboration. Older teams may be more likely to collaborate than younger teams because of their longer exposure to the international scrutiny and the maturity attained in their field. The more PhD holders a team has, the more likely it is to collaborate internationally as team members with PhDs tend to be good counterparts of foreign scientists and engineers, and because they may have had international experience during their personal and professional career. Teams led by someone who writes well in a second language or who has studied overseas tend to be more likely to collaborate internationally than the other teams. In fact, writing in another language is 'a must' for those willing to interact with foreign partners coming from countries other than the Spanish speaking countries. Similarly, having studied overseas may help to establish international linkages that may result in collaboration with mentors, classmates and/or research associates located in the host country. Highly dynamic teams both in terms of the number of R&D projects active and of the number of bibliographic products done tend to engage more in international collaboration than teams that are less dynamic.

International collaboration may also be a function of the field the team specializes in. For instance, it is well known that while R&D projects in physics tend to be mostly collaborative, one can hardly find collaboration around projects on philosophy. Thus, using the UNESCO classification of the data, one can argue that teams working in the natural sciences may be more collaboration-prone than comparable teams working in the social sciences or the humanities given the idiosyncratic nature of the latter types of teams, which may keep them from working with scientists of different origins and epistemic grounds. In contrast, the teams working in the agriculture sciences may allegedly be more likely to collaborate internationally than comparable teams working in the natural sciences mostly because of the international recognition of the former teams attained thanks to their work on tropical agriculture.

International collaboration may also depend on the characteristics of their home institutions and of the sector they operate in. As collaborating with foreign partners requires both financial and institutional support, teams affiliated with big institutions may be more internationally oriented than teams affiliated with mid-size or small institutions in terms of their R&D budget. Competition among teams of the same institution may also help to explain collaborative behavior, and such competition is typical to big institutions. Teams working in the academic sector may be more likely to collaborate internationally than comparable teams working in the business sector. Allegedly, while the former type of teams tend to work on issues of public interest and therefore are expected to be more 'open' to interact with their peers from overseas, teams working in the business sector tend to work on issues with strategic value that may keep them from sharing information and interacting with foreign peers.

Finally, as discussed in the literature reviewed, international collaboration may also be affected by the characteristics of the environment teams are located. Hence, teams located in/or near big cities may be more likely to collaborate internationally than teams located in mid-size or small cities as the teams of the first group tend to have more opportunities to access valuable information on foreign peers, may be more visible given their greater participation in international workshops, and may engage more human resources of higher productivity than teams located in small cities.

The description of the data used is in Appendix H. The reason why a logit model is used as opposed to other models has been discussed in the previous chapter and it has to do with the characteristics of the dependent variable.

4.2 Who Collaborates Internationally in Colombia?

As Table 5 shows, larger teams, older team, teams with large numbers of doctorates, teams with many R&D projects active, and highly productive teams are more likely to collaborate internationally than comparable teams of smaller size, with fewer years in existence, with fewer PhD members, fewer projects active and less productive. Teams led by researchers able to write well in a second language, and teams led by someone who studied overseas in the past are more likely to collaborate than teams of similar characteristics led by someone without either capacities. Teams working in the medical sciences, the social sciences or in the engineering are less likely to collaborate internationally than comparable teams working in the natural sciences. Teams affiliated with large institutions are more likely to collaborate than comparable teams affiliated with small or mid-size institutions. And teams located in small cities are less likely to collaborate internationally than comparable teams located in big cities.

Table 5: Determinants of International Research Collaboration

Variable	Internat. Res. Coll.
Team Size in 2003	0.026
	(1.94)
Team Age in 2003	0.033**
	(2.83)
Total PhDs in 2003	0.253**
	(6.14)
Leader Writes Oth Langua	0.466**
	(3.94)
Leader Studied Overseas	0.453**
	(3.96)
Tot. Proj. in 2003	0.054**
	(4.11)
Tot. Bib. Prods. by 2003	0.008**
	(3.81)
Agrosciences	-0.397
	(1.57)
Medical Sciences	-0.486*
	(2.45)
Social Sciences	-0.366*
	(1.98)

Cont'd

Table 5 *Cont'd*

Variable	Internat. Res. Coll.
Humanities	-0.176
	(1.10)
Engineering	-0.579**
	(3.03)
Other Sciences	-0.140
	(0.47)
Business Sector	0.234
	(0.65)
Government	0.083
	(0.28)
Other Sector	0.601
	(1.44)
Mid. Home Inst.	-0.297*
	(2.33)
Small Home Inst.	-0.341
	(1.73)
Small City	-1.100
	(1.94)
Midsize City	-0.235
	(1.68)
Constant	-1.820**
	(9.17)
Observations	1889

Absolute value of z statistics in parentheses
* Significant at 5%; ** significant at 1%

Contrary to what one would expect, and as a Wald Test of joint effects shows, the sector where the team works does not appear to significantly affect the probability of collaborating internationally. In fact, there is a 53% probability that the observed results could have occurred by chance. Therefore, we can confidently conclude that the hypothesis that the effects of the sector variables are simultaneously equal to zero cannot be rejected.

As the model without the sector variables shows (see Table 6), the number of PhDs appears to be the variable with the greatest impact on the probability of collaborating internationally in Colombia, followed by the number of projects active, the past productivity of the team, and the characteristics of the team leader. In fact, a one standard deviation increase in the number of members with PhD increases team's odds of collaborating internationally by 75%, holding the other variables constant; and a one-unit increase in the number of members with PhD increases team's odds of collaborating by 29%, holding the other variables constant.

Interestingly, the odds that a team led by someone who writes well in a language other than Spanish collaborates are 1.61 times as high as that of teams with leaders who do not write well in a second language, holding the other variables constant. And having leaders who are able to write well in a second language increases the probability of collaborating by 11.2 percentage points, holding the other variables constant at their means.

Finally, as the table shows, teams led by someone who studied overseas in the past are more likely to collaborate internationally than comparable teams led by someone who did not study overseas in the past. Holding the other variables constant, the odds that a team led by someone who studied overseas in the past collaborates are 1.57 as high as that of teams led by people who did not studied overseas in the past, and holding the other variables constant at their means, the former type of teams increases the probability of collaborating internationally by 10.6 percentage points.

Table 6: Determinants of International Research Collaboration: Percentage Change in Odds

Internat. Res. Coll.	b	z	P>z	%	%StdX	SDofX
Team size in 2003	0.02745	2.034	0.042	2.8	17.0	5.7171
Team Age in 2003	0.03535	3.030	0.002	3.6	22.9	5.8373
Total PhDs in 2003	0.25248	6.146	0.000	28.7	74.7	2.2106
Leader Writes in Other Lang.	0.47462	4.023	0.000	60.7	26.7	0.4979
Leader Studied Overseas	0.44927	3.940	0.000	56.7	25.0	0.4972
Tot. Proj. in 2003	0.05508	4.203	0.000	5.7	43.9	6.6037
Tot. Bib. Prod by 2003	0.00777	3.846	0.000	0.8	40.6	43.8751
Agrosciences	-0.39975	-1.586	0.113	-33.0	-9.1	0.2392
Medical Sciences	-0.48066	-2.448	0.014	-38.2	-14.8	0.3325
Social Sciences	-0.38446	-2.088	0.037	-31.9	-13.2	0.3681
Humanities	-0.18291	-1.152	0.249	-16.7	-7.5	0.4284
Engineering	-0.59780	-3.143	0.002	-45.0	-18.4	0.3407
Other Sciences	-0.13717	-0.467	0.641	-12.8	-2.7	0.2014
Mid. Home Inst.	-0.28457	-2.263	0.024	-24.8	-12.9	0.4851
Small Home Inst.	-0.20575	-1.260	0.208	-18.6	-7.4	0.3738
Small City	-1.11583	-1.973	0.048	-67.2	-13.6	0.1310
Midsize City	-0.24056	-1.723	0.085	-21.4	-9.5	0.4138

Squared terms for team size, team age, total number of PhDs, total number of projects active and total number of bibliographic products are added to the model to see if there are curvilinear effects. In fact, according to one of the interviewees, “large teams sometimes experience free riding, that is, situations where when the team is too large, few people do the hard work while many get the merits. This situation ends by fatiguing those who do most of the work and affects internal cohesion. This lack of cohesion is sometimes reflected in the quality of the work done, and foreign institutions and foreign researchers perceive that tension.” Another interviewee claimed that, “when there are too many PhDs in a team, there tend to be too many ‘generals’ and too few ‘soldiers,’ which ends by increasing transaction costs of any collaborative enterprise.”

The exploration of this new model shows that such claims are not supported by the data, except for the number of bibliographic products done, which increases team's odds of collaborating internationally but a decreasing rate. The top number of products at which this positive trend reverses is outside of our data range, however. In fact, holding the other variables constant, the probability of collaborating increases with every additional product but once the team reaches a total of 132 products the probability starts to fall at an increasing rate.

Hence, the model analyzed predicted collaboration for 543 of the teams, of which 386 did collaborate and 157 did not. It predicted that 1346 did not collaborate, but 350 actually did.

The sensitivity of the model is 52.5%: it correctly predicted 386 of the 736 who collaborated. Also, the model is quite specific: 86.4% of those who did not collaborate were not predicted to collaborate (996/1153); 66.9% of those who were predicted to collaborate actually did collaborate, and 74.6% of those who were predicted not to collaborate did actually not collaborate.

More importantly, the model correctly classified 73.2% (386+996/1889), an improvement of about 12% compared to the null model (1153/1889=61.04%). By converting this to an adjusted count R2, we see that the number of errors in prediction drops from 736 to 507 (350+157), a decline of 31.1%.

In sum, based on the results obtained, team size, team age, team composition, leadership, productivity, discipline, institution of affiliation, and geographical location seem to affect the probability of collaborating internationally. In contrast, the sector where the team works is not significantly associated with the collaborative behavior. No significant curvilinear effects were found.

4.3 Factors Explaining Different Types of Collaboration

Based on the population data and using the full model with the sector variables included, the choice of hosting foreign funding depends mostly on the team's size, the number of PhDs, the characteristics of the team leader, the activities performed, the scientific discipline, and the size of the city where the team is located (see Table 7).

The choice of working with foreign funding depends on all the factors considered except team size and the size of the city where the team is located. This finding is confirmed by a Wald Test of the joint effects of the location variables.

More precisely, larger teams tend to host more foreign researchers than smaller teams of similar characteristics, but the size of the teams does not seem to affect the probability of working with foreign funding. Older teams tend to prefer working with foreign funding than younger teams, but team age is not associated with the choice of hosting foreign researchers. The number of PhDs is positively associated with both types of collaboration. Teams led by researchers able to write well in a second language or that studied overseas are more likely to collaborate internationally both through hosting foreign researchers and working with foreign funding than comparable teams. The number of projects active and the number of bibliographic products a team has is associated with the probability of working with foreign funding, but it is not significantly associated with hosting foreign researchers.

Teams working in the medical sciences, or in the engineering, are less likely to host foreign researchers than comparable teams working in the natural sciences. Teams working in the humanities are less likely to work with foreign funding than similar teams working in the natural sciences.

Teams working in the government sector or in the NGOs' sector are more likely to work with foreign funding than comparable teams affiliated with the academic sector. However, the sector where the team works does not seem to be associated with the probability of hosting foreign researchers. Teams affiliated with large institutions are more likely to work with foreign funding than comparable teams affiliated with small and mid size institutions, but the size of the home institution does not seem to be significantly associated with the probability of hosting foreign researchers.

Finally, the size of the city where the team is located also seems to affect the choice of hosting foreign researcher as opposed to the choice of working with foreign funding. In fact, teams located in mid-size cities are less likely to host foreign researchers than comparable teams located in large cities.

Table7: Factors Explaining the Choice of Hosting Foreign Researchers and of Working with Foreign Funding

Variable	Foreign Researchers	Foreign Funding
Team Size in 2003	0.048**	-0.008
	(3.67)	(0.59)
Team Age in 2003	0.006	0.028*
	(0.50)	(2.42)
Total PhDs in 2003	0.187**	0.164**
	(5.08)	(4.29)
Leader Writes Oth Langua	0.467**	0.467**
	(3.46)	(3.39)
Leader Studied Overseas	0.342**	0.395**
	(2.67)	(3.00)
Tot. Proj. in 2003	-0.018	0.065**
	(1.64)	(5.10)

Cont'd

Table 7 ***Cont'd***

Variable	Foreign Researchers	Foreign Funding
Tot. Bib. Prods. by 2003	0.000	0.011**
	(0.29)	(5.64)
Agrosciences	-0.482	-0.241
	(1.68)	(0.87)
Medical Sciences	-0.787**	0.028
	(3.46)	(0.13)
Social Sciences	-0.411*	-0.191
	(1.98)	(0.91)
Humanities	0.076	-0.436*
	(0.46)	(2.36)
Engineering	-0.711**	-0.258
	(3.28)	(1.24)
Other Sciences	-0.286	0.332
	(0.91)	(1.06)
Business Sector	0.182	0.344
	(0.48)	(0.83)
Government	-0.324	0.645*
	(0.90)	(2.01)
Other Sector	-0.104	1.687**
	(0.23)	(3.77)
Mid. Home Inst.	-0.263	-0.445**
	(1.82)	(3.04)
Small Home Inst.	-0.080	-0.762**
	(0.37)	(3.05)
Small City	-1.865	-0.162
	(1.82)	(0.28)
Midsize City	-0.391*	-0.022
	(2.34)	(0.14)
Constant	-1.958**	-2.427**
	(9.29)	(10.98)
Observations	1889	1889

Absolute value of z statistics in parentheses
* Significant at 5%; ** significant at 1%

The analysis of the factors affecting the choice of co-authoring with partners located overseas is done using the sample. In this case, the internal characteristics of the teams are excluded as they were observed after the co-authorship took place.

Hence, as shown in Table 8, teams working in the agricultural sciences or the engineering appear less likely to co-author with colleagues located overseas than comparable teams working in the natural sciences. Interestingly, teams working in the academic sector are less likely to co-author with colleagues located in foreign countries than comparable teams working in the business sector or in the government sector. This may suggest an important level of endogamy characteristic of the Colombian academic sector. Finally, teams affiliated with large institutions are more likely to co-author with partners located overseas than comparable teams affiliated with the small and midsize institutions. No significant effect of location is found. This is confirmed by a Wald Test of the joint effect of these variables not shown here.

Table 8: Factors Explaining the Choice of Co-authoring with Partners Located Overseas

Variable	Int. Co-Authorship in 2001-2
Agrosciences	-1.050**
	(2.69)
Medical Sciences	-0.322
	(1.25)
Engineering	-0.679**
	(2.59)
Other Sciences	-1.464
	(1.93)
Business Sector	2.325**
	(3.78)
Government	2.154**
	(4.66)

Cont'd

Table 8 *Cont'd*

Variable	Int. Co-Authorship in 2001-2
Other Sector	2.366*
	(2.40)
Mid. Home Inst.	-0.921**
	(3.82)
Small Home Inst.	-2.917**
	(5.19)
Small City	-0.468
	(0.59)
Midsize City	-0.403
	(1.54)
Constant	-0.392*
	(2.49)
Observations	672

Absolute value of z statistics in parentheses
* Significant at 5%; ** significant at 1%

4.4 Factors Explaining the Choice of Partners

Based on the population data, and as shown in Table 9, all the factors considered, except team size and team location, significantly affect team choice of collaborating with partners from the north. In contrast, the choice of collaborating with partners from the south seems to be associated with team size, the number of PhDs, the extent to which the team leader writes well in a second language, and team productivity only. The z-tests of the effects of individual variables and the Wald Tests of joint effects of the categorical variables confirm these findings.

Table 9: Factors Explaining the Choice of Collaborating with Partners from Northern and Southern Countries

Variable	Int. Res. w/ North	Int. Res. w/ South
Team Size in 2003	0.013	0.028*
	(0.94)	(2.04)
Team Age in 2003	0.037**	-0.004
	(3.13)	(0.34)
Total PhDs in 2003	0.267**	0.091*
	(6.54)	(2.40)
Leader Writes Oth Langua	0.338**	0.493**
	(2.63)	(3.25)
Leader Studied Overseas	0.483**	0.096
	(3.90)	(0.68)
Tot. Proj. in 2003	0.068**	0.002
	(5.18)	(0.18)
Tot. Bib. Prods. by 2003	0.007**	0.008**
	(3.50)	(4.72)
Agrosciences	-0.567*	0.010
	(2.08)	(0.03)
Medical Sciences	-0.419*	-0.380
	(2.03)	(1.58)
Social Sciences	-0.376	-0.139
	(1.89)	(0.61)
Humanities	-0.220	-0.069
	(1.31)	(0.36)
Engineering	-0.797**	-0.288
	(3.84)	(1.24)
Other Sciences	0.010	0.071
	(0.03)	(0.21)
Business Sector	0.270	0.288
	(0.71)	(0.70)
Government	0.307	-0.048
	(0.99)	(0.13)
Other Sector	1.051*	0.265
	(2.46)	(0.56)
Mid. Home Inst.	-0.436**	-0.164
	(3.15)	(1.02)
Small Home Inst.	-0.400	-0.235
	(1.85)	(0.92)

Cont'd

Table 9 ***Cont'd***

Variable	Int. Res. w/ North	Int. Res. w/ South
Small City	-0.653	-1.379
	(1.15)	(1.34)
Midsize City	-0.260	-0.141
	(1.69)	(0.79)
Constant	-2.146**	-2.435**
	(10.20)	(10.42)
Observations	1889	1889

Absolute value of z statistics in parentheses
* Significant at 5%; ** significant at 1%

4.5 Factors Explaining the Preference of Specific Combinations of Collaborative Activity and Partner

Based on the population data, and as shown in Table 10, the choice of hosting foreign researchers from the North depends mostly on the size of the team, the number of doctorates the team has, the characteristics of the leader, whether the team works in the natural science as opposed to working in the agricultural sciences, the medical sciences, the social sciences or in the engineering; whether it is affiliated with a large institution and whether it is located in a big city. The choice of hosting researchers from the South also depends on the size of the team, the number of PhDs it has, whether the team leader writes well in a second language, or whether it works in the natural sciences as opposed to working in the medical science. Receiving funding from the northern countries is associated with team age, the number of doctorates the team has, the characteristics of the team's leader characteristics, the dynamism of the team, the sector, and the size of the home institution. Finally, the choice of working with projects funded by southern countries is associated with how productive the team is, and whether it works in the multidisciplinary sciences as opposed to working in the natural sciences only.

More precisely, and as discussed before, team size positively affects the choice of hosting foreign researchers. However, it is slightly more important for explaining the choice of hosting researchers from the south than for explaining the choice of hosting researchers from the north. The difference of the effects of each variable can be seen by comparing the z-statistics in each model.

The opposite is true regarding the effects of having PhDs in teams. As the number of PhD holder increases, the probability of hosting foreign researchers increases, but it raises more for hosting researchers from the north than for hosting foreign researchers from the south, holding the other variables constant.

Teams led by someone who writes well a second language positively affects the probability of hosting foreign researchers, but it increases it more for hosting researchers from the south than from the north, holding the other variables constant.

Teams led by someone who studied overseas appear more likely to host foreign researchers than teams not led by someone who studied overseas, but this is mostly because this factor affects the choice of hosting researchers from the north and not from the south.

Teams working in the natural sciences are more likely to host foreign researchers than comparable teams working in the medical sciences, the social sciences and the engineering. However, this is mostly due to its higher probability of engaging researchers from the north than for its probability of engaging researchers from the south, which is not statistically significant. By contrast, the odds of hosting foreign researchers are higher among the teams working in the natural sciences than among the teams working in the medical science. These differences are statistically significant regarding both types of partners. In this case, the difference in the odds is also higher regarding the choice of hosting researchers from the north than of hosting foreign researchers from the south.

Finally, the higher probability of hosting foreign researchers among teams affiliated with big institutions or located in large cities compared to that of teams affiliated with mid-size institutions or being located in mid-size cities responds mostly to the higher probabilities of the former types of teams to host researchers from the north.

As for the factors affecting the choice of working with foreign funding is concerned, team age appears to affect positively the choice of working with foreign funding, but it affects more the choice of working with funding from the north than of working with funding from the south, holding the other variables constant.

By contrast, although the effects of having PhDs in teams positively affects the choice of working with foreign funding, it seems to affect positively more the choice of funding from the north than from the south, holding the other variables constant.

The extent to which a team has a leader who is able to write well in a second language or studied overseas in the past is more important for explaining the choice of working with foreign funding from the north than for explaining the choice of working with funding from the south (whose effects are not statistically significant).

The number of projects active a team has is important for explaining the choice of foreign funding. However, the effect is greater for explaining the choice of working with funding from the north. In contrast, the number of S&T products a team has is more important for explaining the choice of funding from the south than from the north, although it is also important for explaining the choice of receiving funding from the north.

Teams working in the other sciences or in the multidisciplinary sciences are more likely to work with projects funded by southern countries than teams working in the natural sciences. They are also more likely to work with funding from the south than with funding from the north.

Teams affiliated with the NGOs' are more likely to work with foreign funding than comparable teams affiliated with the academy, mostly because the former are more likely to work with funding from the north.

Finally, teams affiliated with large institutions are more likely to work with foreign funding than comparable teams affiliated with small and midsize institutions. However, the main difference is due to their likelihood of working with funding from northern countries.

No significant effects were found regarding the location variables on the probability of working with funding of any origin.

Table10: Factors Explaining the Choice of Different Combinations of Partners and Types of Collaboration

Variable	Researchers from North	Researchers from South	Funding from North	Funding from South
Team Size in 2003	0.040**	0.046**	-0.007	-0.021
	(2.72)	(3.12)	(0.50)	(1.01)
Team Age in 2003	0.017	-0.025	0.030*	0.018
	(1.35)	(1.56)	(2.52)	(1.21)
Total PhDs in 2003	0.206**	0.099*	0.191**	-0.011
	(5.08)	(2.33)	(4.93)	(0.24)

Cont'd

Table 10 ***Cont'd***

Variable	Researchers from North	Researchers from South	Funding from North	Funding from South
Leader Writes Oth Langua	0.386*	0.520**	0.467**	0.313
	(2.30)	(2.91)	(3.27)	(1.33)
Leader Studied Overseas	0.375*	0.146	0.445**	0.146
	(2.37)	(0.88)	(3.27)	(0.67)
Tot. Proj. in 2003	-0.010	-0.015	0.066**	0.026
	(0.77)	(1.06)	(5.14)	(1.86)
Tot. Bib. Prods. by 2003	-0.001	0.002	0.009**	0.011**
	(0.43)	(1.24)	(4.94)	(5.52)
Agrosciences	-0.875*	-0.285	-0.235	0.127
	(2.30)	(0.80)	(0.82)	(0.29)
Medical Sciences	-0.889**	-0.771**	0.052	0.168
	(3.24)	(2.58)	(0.24)	(0.51)
Social Sciences	-0.542*	-0.397	-0.150	0.164
	(2.11)	(1.49)	(0.69)	(0.47)
Humanities	0.087	-0.103	-0.355	-0.461
	(0.45)	(0.48)	(1.87)	(1.36)
Engineering	-1.266**	-0.353	-0.246	-0.225
	(4.16)	(1.35)	(1.15)	(0.62)
Other Sciences	-0.079	-0.647	0.272	0.897*
	(0.23)	(1.44)	(0.84)	(2.20)
Business Sector	0.268	0.330	0.263	0.572
	(0.59)	(0.69)	(0.62)	(0.90)
Government	-0.103	-0.406	0.623	0.722
	(0.24)	(0.84)	(1.91)	(1.45)
Other Sector	0.251	-0.229	1.761**	1.245
	(0.51)	(0.37)	(3.92)	(1.90)
Mid. Home Inst.	-0.460*	-0.072	-0.447**	-0.463
	(2.53)	(0.39)	(2.94)	(1.77)
Small Home Inst.	-0.092	-0.118	-0.656*	-0.810
	(0.34)	(0.41)	(2.57)	(1.80)
Small City	-1.231		-0.009	-0.073
	(1.19)		(0.02)	(0.07)
Midsize City	-0.519*	-0.200	-0.088	0.049
	(2.34)	(0.96)	(0.52)	(0.18)
Constant	-2.499**	-2.541**	-2.648**	-3.492**
	(9.89)	(9.43)	(11.55)	(10.16)
Observations	1889	1856	1889	1889

Absolute value of z statistics in parentheses
* Significant at 5%; ** significant at 1%

4.6 Conclusions

Table 11 summarizes the findings on the effects of the variables studied on the collaborative behavior in Colombia. It shows the positive, the negative or the non-effects (at he 0.05 level) of each variable on the type of collaboration studied, the type of partner involved, and the preference for a specific combination of collaborative activity and partner's origin. In particular, it shows that the number of doctorates a team has and the characteristics of its leader are the variables with the stronger explanatory power on team's decision to collaborate internationally all types and origins considered.

Table 11: Summary Table: Determinants of International Research Collaboration in Colombia

Variable	IRC	Type of Collaboration			Type of Partner		Type of Collaboration and Partner			
		Foreign Researchers	Foreign Funding	Co-Authorship	North	South	Researchers from North	Researchers from South	Funding from North	Funding form South
Team size in 2003	+	+	No Sig	?	No Sig	+	+	+	No Sig	No Sig
Team Age in 2003	+	No Sig	+	?	+	No Sig	No Sig	No Sig	+	No Sig
Total PhDs in 2003	+	+	+	?	+	+	+	+	+	No Sig
Leader Writes in Other Lang.	+	+	+	?	+	+	+	+	+	No Sig
Leader Studied Overseas	+	+	+	?	+	No Sig	+	No Sig	+	No Sig
Tot. Proj. in 2003	+	No Sig	+	?	+	No Sig	No Sig	No Sig	+	No Sig
Tot. Bib. Prod by 2003	+	No Sig	+	?	+	+	No Sig	No Sig	+	+
Agro sciences	No Sig	No Sig	No Sig	-	-	No Sig	-	No Sig	No Sig	No Sig
Medical Sciences	-	-	No Sig	No Sig	-	No Sig	-	-	No Sig	No Sig
Social Sciences	-	-	No Sig	?	No Sig	No Sig	-	No Sig	No Sig	No Sig
Humanities	No Sig	No Sig	-	?	No Sig	No Sig	No Sig	No Sig	No Sig	No Sig
Engineering	-	-	No Sig	-	-	No Sig	-	No Sig	No Sig	No Sig
Other Sciences	No Sig	No Sig	No Sig	No Sig	No Sig	No Sig	No Sig	No Sig	No Sig	+
Business Sector	No Sig	No Sig	No Sig	+	No Sig	No Sig	No Sig	No Sig	No Sig	No Sig
Government	No Sig	No Sig	+	+	No Sig	No Sig	No Sig	No Sig	No Sig	No Sig
Other Sector	No Sig	No Sig	+	+	+	No Sig	No Sig	No Sig	+	No Sig
Small Home Inst.	No Sig	No Sig	-	-	No Sig	No Sig	No Sig	No Sig	-	No Sig
Mid. Home Inst.	-	No Sig	-	-	-	No Sig	-	No Sig	-	No Sig
Small City	-	No Sig	No Sig	No Sig	No Sig	No Sig	No Sig	No Sig	No Sig	No Sig
Midsize City	No Sig	-	No Sig	No Sig	No Sig	No Sig	-	No Sig	No Sig	No Sig

The understanding of the determinants of international research collaboration and of the different ways it is conceived as well as of the choice of partners helps to better design public policies oriented at exploiting the benefits derived from collaborating internationally or at reducing the negative effects that may result from it. The policy implications of the results shown here are discussed in the last chapter of the dissertation as they would depend on the effects found to have international collaboration on team performance. Hence, the next chapters discuss the results obtained using the models described in chapter 3 to give an account of the determinants of research team productivity and research team orientation respectively. In particular, they focus on the ways international research collaboration -as expressed by the modalities studied so far and involving the types of partner identified -affects team output and teams' ability to contribute to local knowledge. They also discuss the findings of the propensity score matching approach used to assess the impact of international research collaboration while controlling for selection bias.

CHAPTER 5

INTERNATIONAL COLLABORATION AND RESEARCH TEAM OUTPUT IN COLOMBIA

5.1 To What Extent Does International Collaboration Affect Team Output?

In chapter 3 and in the appendix we concluded that, given the characteristics of the dependent variable, the Zero-Inflated Negative Binomial Model was the preferred model to study the determinants of team output in Colombia. Given the relatively complexity of this model, a brief explanation of the ways its coefficients should be interpreted is worth doing before we discuss the findings.

As Long and Freese (2001) acknowledge, when interpreting zero inflated models, it is easy to be confused by the direction of the coefficients. The first portion of the Stata output, which in this study is reported in the first four columns, contains coefficients for those in the Not Always-0 Group. This group comprises those teams who have the opportunity to report bibliographic products during the period of observation. The coefficients can be interpreted in the same way as coefficients from the OLS, the PRM or the NBRM models. The second portion of the Stata output, which in this study is reported in the latter three columns, contains coefficients for the log-odds of being in the Always-0 Group of teams compared to the Not Always-0 Group. As explained earlier, a team is in the Always-0 group if it is not allowed to report bibliographic products due to structural constraints (it works for industry on cutting edge technologies or works for a government agency on issues not to be made public, etc.) or due to conjunctural reasons (it did not have the chance to report bibliographic products for a given reason). These coefficients can be interpreted just as the coefficients for a binary logit model. When the same variables are included in both equations (because they are judged to be important for explaining team output as well as for explaining the impossibility of teams to report bibliographic products, or due to the lack of theory like in the case of this study), the signs of the corresponding coefficients from the binary equation are often in the opposite direction of the coefficients for the count equation. Hence, while the first columns help to predict number of bibliographic products so that a negative coefficient would indicate lower productivity, the latter columns helps to predict membership in the group of teams that always has zero counts so a positive coefficient implies lower productivity.

Thus, using a Zero-Inflated Negative Binomial model to account for the effects of research collaboration on team output in Colombia, we find that teams that collaborate internationally, that are large in size, that have many PhDs, and that report many projects active tend to be more productive than comparable teams that do not collaborate internationally, that are of a small size, that have few or no PhD members, and that report few projects active. Teams working in the humanities are less productive than comparable teams working in the natural sciences. Teams affiliated with small institutions are less productive than comparable teams affiliated with big institutions (see Table 12).

Collaborating internationally or having many research projects active reduces the odds of reporting no bibliographic products. In contrast, larger teams, older teams, teams working in the medical sciences, and teams affiliated with the business or the government sector, are more likely than comparable teams of smaller size, younger teams, teams working in the natural sciences, and teams affiliated with the academic sector to be in the always-0 group of non-productive teams.

Table 12: Factors Affecting Team Output: ZINB

	Tot. Bib. Prods. 2003-5			Always0		
	Coef.	z	P>|z|	Coef.	z	P>|z|
Internat. Res. Coll.	0.305	4.38	0.000	-0.824	-3.13	0.002
Team size in 2003	0.024	2.71	0.007	0.052	2.49	0.013
Team Age in 2003	0.011	1.74	0.082	0.041	2.22	0.027
Total PhDs in 2003	0.075	3.27	0.001	-0.075	-1.29	0.197
Tot. Proj. in 2003	0.044	7.54	0.000	-0.064	-2.11	0.035
Agrosciences	-0.098	-0.55	0.584	0.786	1.78	0.075
Medical Sciences	0.028	0.26	0.794	0.822	2.20	0.028
Social Sciences	0.003	0.03	0.975	0.559	1.37	0.170
Humanities	-0.181	-1.98	0.047	0.288	0.74	0.459
Engineering	0.017	0.16	0.871	0.216	0.53	0.598
Other Sciences	-0.169	-1.09	0.274	0.178	0.32	0.750
Business Sector	-0.383	-1.71	0.088	1.198	2.33	0.020
Government	0.205	1.22	0.222	1.256	3.07	0.002
Other Sector	0.300	1.70	0.090	0.579	0.92	0.355
Mid. Home Inst.	0.036	0.51	0.609	-0.373	-1.52	0.129
Small Home Inst.	-0.268	-2.37	0.018	-0.328	-0.82	0.415
Small City	0.131	0.60	0.549	0.521	0.79	0.432
Midsize City	-0.139	-1.77	0.077	-0.363	-1.20	0.230
Constant	1.521	14.14	0.000	-1.957	-4.83	0.000
/lnalpha				-0.020	-0.33	0.744
alpha				0.980		
Observations: 1889						

International research collaboration is the explanatory variable with the greatest impact on team productivity right after the number of R&D projects active and the number of PhDs a team has. In fact, a one standard deviation increase (not shown in the table) in international collaboration increases team's expected productivity count by 16%, holding the other variables constant (the effects of number of projects active and of members with PhDs are 34% and 18% respectively). The expected count of bibliographic products of collaborating teams is 36% higher than that of non-collaborating teams, and their odds of being in the always zero group of unproductive teams are 56% lower than comparable teams that do not collaborate internationally. Measured in terms of discrete changes, and holding all other variables constant at their means, collaborating internationally increases expected productivity count by 3.14 bibliographic products.

On the other hand, regarding the control variables, team size increases the expected team's rate of bibliographic products, but curiously, an additional team member increases team's odds of being in the always zero group. This contradictory result may suggest the presence of quadratic effects of team size, that is, it might be the case that team size increases productivity but at a decreasing rate. This hypothesis is explored later.

The data does not support the claim found in the literature that team age increases team productivity (Harrison, Price et al. 2002; Rey-Rocha, Martin-Sempere et al. 2002), such effects are not confirmed by the data. As claimed by Cohen 1991, team age is not associated with team productivity. Instead, in the Colombian case, a one unit increase of team age increases team's odds of being in the non-productive group of teams, once we hold the other variables constant.

The number of doctorates increases team's expected bibliographic production, but it does not affect the odds of being in the always-0 group of teams. As one would have thought, the number of R&D projects active increases team's expected rate of bibliographic products and decreases its odds of being in the always-0 group. Teams working in medical sciences have similar expected rate of production of comparable teams working in the natural sciences, but they are more likely of being in the unproductive group of teams than those working in the natural sciences. Interestingly, and contrary to what one would expect, teams working in the engineering are not less productive or more likely of being in the Always-0 group than teams of comparable characteristics working in the natural sciences.

As expected, teams affiliated with the business sector are less productive and are more likely of not having the opportunity to report bibliographic products than similar teams affiliated with the academic sector. Teams affiliated with the government sector, as compared to teams affiliated with institutions working in the academic sector, do not have different expected rate of production, however, and as expected, these teams are more likely than comparable teams affiliated with the education sector of not having the opportunity to report bibliographic products. As one would think, teams affiliated with small institutions are less productive than teams affiliated with big institutions, but these teams are not more likely than comparable teams affiliated with big institutions to report zero bibliographic products.

Finally, contrary to extant literature (see literature review), the size of the urban agglomerate where the team is located does not seem to affect its production nor its likelihood of reporting zero counts once we hold the other variables constant. In fact, a Wald Test performed on the joint effects of the variables associated with team location on team productivity shows that, holding all other variables constant, there is a 26% probability that the observed results could have occurred by chance (Prob > chi2 = 0.2649). Therefore we conclude that location is not associated with team output. Furthermore, the measures of fit shown in Table 13 below, allows us to confidently conclude that the model without these location variables (called 'current' model in the table) is much better than the full model with all the variables considered. The difference of 24.191 in BIC' provides very strong support for the 'restricted' model.

Table 13: Measures of Fit to Compare Models With and Without the Location Variables

	Current	Saved	Difference
Model:	zinb	zinb	
N:	1889	1889	0
Log-Lik Intercept Only	-5889.003	-5889.003	0.000
Log-Lik Full Model	-5631.601	-5628.609	-2.992
D	11263.202(1854)	11257.218(1850)	5.984(4)
LR	514.803(32)	520.787(36)	5.984(4)
Prob > LR	0.000	0.000	0.200
McFadden's R2	0.044	0.044	-0.001
McFadden's Adj R2	0.038	0.038	0.000
ML (Cox-Snell) R2	0.239	0.241	-0.002
Cragg-Uhler(Nagelkerke) R2	0.239	0.241	-0.002
AIC	6.000	6.001	-0.001
AIC*n	11333.202	11335.218	-2.016
BIC	-2723.008	-2698.817	-24.191
BIC'	-273.401	-249.210	-24.191
BIC used by Stata	11527.235	11551.427	-24.191
AIC used by Stata	11333.202	11335.218	-2.016

Difference of 24.191 in BIC' provides very strong support for current model.

The agglomeration effects extensively claimed in the literature seem therefore not to be confirmed by the data in the Colombian case. Further investigation is necessary. We will come back to this point later.

To explore if there are curvilinear effects (Lewis 2002) of team size, team age, number of PhDs, and number of projects active, four quadratic variables are added to the model without the location variables. According to the regression, team size, team age, and number of projects active increase the expected number of bibliographic products but at a decreasing rate. In fact, holding all other variables constant, every additional team member increases expected team productivity, but once the team reaches a size greater than 16 members, team output begins to decrease at an increasing rate with every additional team member. This finding is consistent with that suggested by Qurashi 1991 and Qurashi 1993 who also found curvilinear effects of team size in the US, UK, Pakistan, Bangladesh, and Greece with peaks between 6 and 46 members.

Interestingly, once the team reaches 20 years old, its output begins to fall at an increasing rate with every additional year of team age, holding the other variables constant. Finally, as the number of projects active rises, team productivity increases but once the team reaches a top of 46 projects active, the number of bibliographic products decreases at an increasing rate with every additional project, holding the other variables constant. All these top values are within the data range. Therefore, we conclude that there are curvilinear effects of team size, team age and number of projects active but not of number of doctorates a team has.

The comparison of the models with and without the quadratic variables through an LR Chi2 test shows that the model with the quadratic variables is preferred over the model without them. For this reason, we report the results obtained using the model with the quadratic variables in the study of the effects of different types of collaboration activities and of partners. Before we do that, let’s first analyze the overall effects of international research collaboration on research team output using control groups.

5.2 Overall Impact of International Research Collaboration on Team Output in Colombia

Using control groups to test the effects of international research collaboration on team output confirms the fact that collaborating internationally contributes importantly to local S&T capabilities. As explained in chapter 3, the control group is constructed based on the characteristics of the teams and their probabilities to collaborate internationally. As discussed in chapter 4, these characteristics include team size, team age, team composition, team leader characteristics, team productive trajectory, team activity, the discipline they work in, the sector they are affiliated with, the characteristics of their home institution, and their location.

Therefore, the impact assessment done using the Propensity Score Matching approach explained earlier shows that, as depicted in Table 14, the average treatment effects on the treated yields a difference of 2.08 bibliographic products in favor of those teams that collaborate internationally versus those that do not. A difference of 2.08 is significantly large even in the most conservative scenario of a bandwidth of 0.01. If we increase the bandwidth to 0.05 and 0.1, the difference in productivity between collaborating and non-collaborating reaches 2.4 and 3.29 products respectively, a statistic far from the 7.89 reported by the simple t-test model based on the unmatched sample shown earlier and on the output below, however it is still an important difference[17].

Table 14: Team Output using PSM and International Research Collaboration as the Treatment Variable

Sample	Treated	Controls	Difference	S.E.	T-stat
Unmatched	13.67	5.79	7.89	0.62	12.62
ATT	12.86	10.78	2.08	0.77	2.71
ATU	5.86	7.07	1.21		
ATE			1.55		

Note: ATT: Average Treatment Effect on the Treated; ATU: Average Treatment Effect on the Untreated; and ATE: Average Treatment Effect on the Population.

[17] After imposing the common support condition, 18 of the 736 teams that collaborated internationally fell outside the common support region because their propensity score was higher than the maximum propensity score of the non-collaborating teams. Hence, these 18 cases were discarded, and the analyses of the average treatment effect (ATE) and of the average treatment effect of the untreated (ATU) are done on 1,846 teams out of the 1,889 teams of the sample.

To assess the matching quality Table 15 below shows that we significantly reduced the differences between the characteristics on the teams with the matching procedure, turning the treated and the control groups significantly similar, which makes them comparable in all relevant aspects. For example, the difference in the number of PhD holders between teams that collaborated versus those that did not falls substantially from 1.553 to 0.008 once we used the matched teams, reducing the bias 99.5%. A 95.2% reduction of the bias is achieved by the matching procedure regarding the difference in team size between collaborating and non-collaborating teams: the difference of 2.971 members between unmatched teams falls to 0.143 between matched teams.

Table 15: Assessment of the Matching Quality: PSM-Research Team Output

Variable	Sample	Mean Treated	Mean Control	%bias	% reduct bias
Internat. Res. Coll.	Unmatched	1	0	.	
	Matched	1	0	.	.
Team size in 2003	Unmatched	9.072	6.1006	50.7	
	Matched	8.5891	8.4454	2.5	95.2
Team Age in 2003	Unmatched	8.8478	5.6644	54	
	Matched	8.5961	8.9457	-5.9	89
Total PhDs in 2003	Unmatched	2.4035	0.84996	69.1	
	Matched	2.1825	2.1745	0.4	99.5
Leader Writes in Other Lang.	Unmatched	0.70245	0.4484	53.2	
	Matched	0.69916	0.71552	-3.4	93.6
Leader Studied Overseas	Unmatched	0.66712	0.48222	38	
	Matched	0.66156	0.66606	-0.9	97.6
Tot. Proj. in 2003	Unmatched	8.2215	4.1752	59.6	
	Matched	7.6031	7.161	6.5	89.1
Tot. Bib. Prod by 2003	Unmatched	49.432	18.162	70.6	
	Matched	45.213	45.396	-0.4	99.4
Agrosciences	Unmatched	0.06386	0.05898	2	
	Matched	0.06546	0.07271	-3	-48.5

Cont'd

Table 15 ***Cont'd***

Variable	Sample	Mean Treated	Mean Control	%bias	% reduct bias
Medical Sciences	Unmatched	0.13315	0.12229	3.3	
	Matched	0.12953	0.12152	2.4	26.3
Social Sciences	Unmatched	0.11821	0.18907	-19.7	
	Matched	0.12117	0.10894	3.4	82.7
Humanities	Unmatched	0.22283	0.25412	-7.3	
	Matched	0.22423	0.20461	4.6	37.3
Engineering	Unmatched	0.11821	0.14397	-7.6	
	Matched	0.12117	0.1357	-4.3	43.6
Other Sciences	Unmatched	0.04755	0.03903	4.2	
	Matched	0.04318	0.03643	3.3	20.9
Mid. Home Inst.	Unmatched	0.2894	0.43539	-30.7	
	Matched	0.29666	0.28139	3.2	89.5
Small Home Inst.	Unmatched	0.1413	0.18474	-11.8	
	Matched	0.13928	0.12638	3.5	70.3
Small City	Unmatched	0.00543	0.02515	-16.1	
	Matched	0.00557	0.0026	2.4	84.9
Midsize City	Unmatched	0.1644	0.25412	-22.2	
	Matched	0.16574	0.17642	-2.6	88.1

Appendix I shows the bootstrapping procedure used to test the statistical significance of treatment effects and to compute their standard errors in case analytical estimates are biased or unavailable. Each bootstrap draw consisted in the re-estimation of the results, including the estimation of propensity scores, common support, etc. The bootstrapping was repeated 999 times, which led to 999 bootstrap samples and 999 estimated average treatment effects. Based on the bootstrap results obtained (observed coefficient = 2.08, z = 1.96, P>|z| = 0.050), we confirm our finding and confidently conclude that *international research collaboration is a strong factor affecting research team output in Colombia*. However, to better understand the ways international research collaboration affects team output, it is necessary to study the effects of each type of collaboration on team performance. That is the purpose of the next section.

5.3 Type of Collaboration and Team Output

In the previous sections we saw that collaborating internationally positively affects research team output. International Research Collaboration was defined in that analysis as a dummy variable coded 1 if the team had a foreign member or if it had projects with foreign funding, zero otherwise (that is, none of these two types of collaboration were present). In this section, we answer the questions: a) how does working with foreign funding affect team output? And b) how does working with foreign researchers affiliated with teams affect team bibliographic production? The third type of collaboration considered, co-authoring with partners located overseas, is studied in the following section using a smaller sample.

5.3.1 Effects of Hosting Foreign Researchers and Having Foreign Funding on Team Output

The analyses use the same data, and the international research collaboration variables are measured as two dummy variables: hosting foreign partners is coded 1 for teams reporting members from foreign origin working between 2003 and 2005, 0 otherwise; and foreign funding is coded 1 for teams reporting projects with foreign funding between 2003 and 2005, and 0 otherwise. Summary statistics are reported in Appendix J.

Using the ZINB model without the location variables, considering quadratic effects, and analyzing independently both types of international collaboration, we find that having foreign funding affects team productivity positively more than having foreign researchers. In fact, as shown in Table 16, which omits the results obtained concerning the control variables, having foreign funding increases team bibliographic production by nearly 33%; and it decreases the odds of being in the always-0 group by nearly 71%, holding the other variables constant. Holding the other variables constant at their mean, collaborating through foreign funding increases expected productivity count by 3.27 bibliographic products. By contrast, and surprisingly, having foreign researchers working at Colombian teams also appears to have a positive effect on team performance, but we cannot reject the null that such result is due to chance.

Table 16: Foreign Researchers, Foreign Funding, and Team Output

	Tot. Bib. Prods. 2003-5*		Always0**	
	%	P>|z|	%	P>|z|
Foreign Researchers	12.3	0.098	-6.8	0.771
Foreign Funding	32.5	0.000	-71.2	0.000
ln alpha	-0.08868			
alpha	0.91514	SE(alpha) = 0.06888		

% = percent change in expected count for unit increase in X
P>|z| = p-value for z-test
* Count Equation: Percentage Change in Expected Count for Those Not Always 0
** Binary Equation: Factor Change in Odds of Always 0

5.3.2 Effects of Co-Authoring with Scientists and Engineers Located Overseas

To better assess the impact of different types of collaborative activities we added the third modality of international collaboration discussed earlier: Co-Authoring journal articles. To do the analysis, we relied on a new dataset: one where the co-authorship of papers was observed. As explained earlier, this dataset was created based on the articles published by Colombian scientists and technicians between 1998 and 2005 indexed by the Web of Science and Scopus.

A sample of 672 teams was randomly selected. The selection criteria included a) teams with at least two members working together by 2003, b) at least one research project active between 2003 and 2005, c) not working in the social sciences or the humanities, and d) were created by March 2004. The reasons why the social sciences and the humanities were excluded are both practical and epistemological. Regarding the latter, it is commonly accepted that the main products of the teams working in these areas are not journal articles, but rather books. Since we are interested on the effects of different types of collaboration activities on team productivity as measured by the total number of bibliographic products, focusing on journal articles would go against those teams. The practical reason is that, given the large number of researchers these teams usually have, the searching and assigning process would have been highly costly. The data is described in more detail in Appendix K.

An important improvement is done to the model as a time lag is introduced between the new collaboration variable considered, co-authoring, and the outcome variable. Thus, a new dummy variable is used to account for international co-authorship taking place in 2002 or before, while team productivity is observed to have been produced between 2003 and 2005.

As Table 17 shows, co-authoring with foreign partners also positively affects team productivity. It increases team output by 39%, holding the other variables constant. Co-authoring increases the expected productivity by 2.91 bibliographic products, holding the other variables constant at their means (not shown in the table). By contrast, collaborating internationally through co-authoring does not seem to be associated with the probability of being (or not being) in the unproductive group of teams.

Table 17: Co-authorship with Colleagues Located Overseas and Team Output

	Tot. Bib. Prods. 2003-5*		Always0**	
	%	P>|z|	%	P>|z|
Co-Authorship in 2001-2	39.2	0.011	-2.3	0.992
Team size in 2003	4.0	0.013	14.8	0.553
Team Age in 2003	0.9	0.295	-11.7	0.377
Total PhDs in 2003	7.5	0.015	-49.9	0.531
Tot. Proj. in 2003	2.6	0.002	-59.8	0.200
Agrosciences	8.8	0.683	575.8	0.192
Medical Sciences	-16.4	0.209	-85.6	0.774
Engineering	-1.3	0.916	44.5	0.754
Other Sciences	-23.7	0.164	1954.5	0.196
Business Sector	-54.2	0.031	1363.1	0.282
Government	13.0	0.548	441.7	0.324
Other Sector	55.4	0.199	-100.0	0.001
Small Home Inst.	-21.1	0.215	6.4	0.944
Mid Home Inst.	0.2	0.987	-99.3	0.133
Small City	65.8	0.121	5313.2	0.192
Midsize City	-6.9	0.563	3.6	0.967
ln alpha	0.28807			
alpha	1.33385	SE(alpha) = 0.14822		

% = percent change in expected count for unit increase in X
P>|z| = p-value for z-test
* Count Equation: Percentage Change in Expected Count for Those Not Always 0
** Binary Equation: Factor Change in Odds of
Always 0

Although the two datasets are not comparable, the results for the control variables drawn from the sample mirror the direction of the effects of most of the control variables analyzed using the whole population. Similarly to the findings of the analysis of the larger dataset, here team size, education, and activity appear to have positive effects on team output. Here, as before, team age does not seem to be related to team productivity. Teams affiliated with business are less productive than teams affiliated with academic institutions, but contrary to the finding based on the larger dataset, belonging to a team affiliated with the business sector does not affect the odds of being in the always-0 group. In this model, scientific field does not seem to matter and only the teams affiliated with the NGOs sector are significantly less likely of being in the Always-0 group of teams.

To understand further the effects of international research collaboration, we need to look at the characteristics of the partners. That is the purpose of the next section.

5.4 North-South and South-South Collaboration and Team Output

In this section we are interested on knowing whether the effects of collaborating with northern countries differs from collaborating with southern countries and on answering to the questions of, if so, how does collaboration resulting from foreign funding from the North differ from similar type of collaboration from the South? And how does having research members from the North differ from having research members from the South?

Using the larger sample again without the location variables and including the curvilinear effects discussed earlier, we use de ZINB model with the origin of the partners as the key independent variables. Collaboration with partners from the North is observed as a dummy variable coded 1 for the teams that either had foreign researchers or foreign funding between 2003 and 2005 whose origin was a Northern country (see classification of countries in Appendix B), 0 otherwise. Collaboration with partners from the South is represented by a dummy variable coded 1 for the teams that either had foreign researchers or foreign funding between 2003 and 2005 whose origin was a Southern country, 0 otherwise. Hosting people from the North is a dummy variable coded 1 for the teams that had foreign researchers from the North between 2003 and 2005, zero otherwise. Hosting people from the South is a dummy variable coded 1 for the teams that had foreign researchers from the South between 2003 and 2005, zero otherwise. Working on projects funded by institutions from the North is represented by a dummy variable coded 1 for the teams that had foreign funding from the North between 2003 and 2005, zero otherwise. Having funding for R&D projects from countries from the South is a dummy variable coded 1 for the teams that had foreign funding from the South between 2003 and 2005, zero otherwise. Summary statistics are reported in Appendix J.

Table 18 below shows another surprising result: for Colombian teams, the impact of collaborating with the South is greater than that of collaborating with the North, holding the other variables constant. In fact, collaborating with the South increases bibliographic production by 46%, whereas collaborating with the North is not statistically significant, holding the other variables constant. Collaborating with southern countries increases expected productivity count by 3.35 bibliographic products, holding the other variables constant at their means (not shown in the table).

Table 18: Team Output: Percentage Change in Expected Count by Type of Partner

	Tot. Bib. Prods. 2003-5*		Always0**	
	%	P>|z|	%	P>|z|
Int. Res. with North	10.8	0.140	-58.6	0.001
Int. Res. with South	45.8	0.000	-17.1	0.519
ln alpha	-0.11376			
alpha	0.89247	SE(alpha) = 0.06776		

% = percent change in expected count for unit increase in X
P>|z| = p-value for z-test
* Count Equation: Percentage Change in Expected Count for Those Not Always 0
** Binary Equation: Factor Change in Odds of
Always 0

However, although collaborating with the South appears to have greater impact on team productivity, it does not seem to reduce the odds of being in the always-0 group. In contrast, collaborating with partners from the North reduces the odds of being in the always-0 group by nearly 59%, holding the other variables constant. This is statistically significant at the 0.001 level.

What difference does it make collaborating with people from the North or having funding from the North as opposed to collaborating with people from the South or having funding from the South? These are the questions discussed in the next section.

5.5 Type of Collaboration, Type of Partner, and Team Output

5.5.1 Regression Analysis

Table 19 below reveals several interesting findings: having projects funded by institutions from the South, and hosting people from the South increase bibliographic production more than having funding from the North or hosting researchers from the North. In fact, holding all other variables constant, having funding from the South increases the number of bibliographic products by 52%; hosting researchers from the South increases it by 32%; and having foreign funding from the North increases it by 20%. Holding the other variables constant at their means, having foreign funding from southern countries increases expected productivity counts by 4.98 bibliographic products; having foreign funding from northern countries increases it by 2.21, and hosting foreign researchers from southern countries increases it by 1.91.

Surprisingly, having researchers from the North is not associated with team productivity in any meaningful way (in fact, it appears negatively associated! but this is not statistically significant). Equally surprising, whereas having funding from the North reduces the odds of being in the always-0 group (it decreases it by nearly 60%, holding the other variables constant); having funding from the south or hosting foreign researchers regardless of their origin is not significantly associated with the odds of being (or not being) in the always-0 group of teams.

Table 19: Team Output: Percentage Change in Expected Count by Type of Collaboration and Type of Partner

	Tot. Bib. Prods. 2003-5*		Always0**	
	%	P>|z|	%	P>|z|
People from North	-7.0	0.348	-27.1	0.306
People from South	32.1	0.002	30.7	0.351
Funding from North	20.0	0.014	-59.9	0.004
Funding from South	51.8	0.000	-78.7	0.079
ln alpha	-0.12775			
alpha	0.88008	SE(alpha) = 0.06541		

% = percent change in expected count for unit increase in X

P>|z| = p-value for z-test

* Count Equation: Percentage Change in Expected Count for Those Not Always 0

** Binary Equation: Factor Change in Odds of Always 0

5.5.2 Non-Parametric Analysis

To assess the effect of different types of collaborative activities and different types of partners using the PSM approach, North and South foreign researchers and foreign funding are taken as the treatments of interest. The analyses are done based on the larger database. Table 20 summarizes the analyses done and shows the differences in productivity between collaborating and non-collaborating teams before and after the matching procedure using each of the treatment variables studied. As the table shows, the effect of collaborating through projects funded by foreign institutions is greater than the effect of collaborating through hosting foreign researchers. This is consistent with the finding discussed earlier using the hypothesis testing approach implemented through the parametric models. Then, as now, collaborating with foreign researches does not seem to be associated with team output.

Similarly to the findings reported earlier, the matching technique also shows that collaborating with the South yields greater effects than collaborating with the North, and that collaborating with partners from the North does not seem to be associated with research team output.

Finally, and again consistent with the findings obtained using the ZINB model, the PSM analysis provides evidence in the sense that collaborating through projects with the South yields the greatest positive effects on team productivity. However, collaborating with foreign researchers both from the North and the South appear in this analysis not having an impact on team output.

Table 20: Type of Collaboration, Type of Partner, and Team Output –PSM

	Team Productivity: Population					
Treatment bwidth: 0.01	**Difference Unmatched**	**T-statistic**	**Difference Matched ATE**	**Difference Matched ATT**	**T-statistic**	**On Support**
Internat. Res. Coll.	7.89	12.62	1.55	2.08	2.71	1846
Foreign Researchers	2.68	3.48	0.69	-0.75	-0.92	1834
Foreign Funding	10.51	15.39	2.34	4.15	4.00	1747
Co-Authorship *	8.97	7.45	2.51	4.20	2.29	629
Int. Res. with North	8.17	12.41	0.98	0.91	1.02	1824
Int. Res. with South	8.40	9.97	2.18	2.97	2.51	1828
People from North	2.57	2.77	-0.41	-0.64	-0.65	1736
People from South	3.26	3.17	1.59	0.39	0.38	1759
Funding from North	10.20	14.43	1.83	3.10	2.82	1740
Funding from South	15.29	12.16	3.46	5.71	2.73	1852

Source: Silac 2005. Author: Gonzalo Ordonez

* Analysis done on 672 Teams

5.6 Summary of Findings and Conclusions

Several findings emerge from the analysis of the data and relevant conclusions can be drawn from both the methodological and the theoretical perspectives.

As expected, research team output depends on the size of the team, its age, the level of education of its members, the number of projects active, the scientific field it specializes in, the sector it works in, and the size of the institution with which it is affiliated.

Regarding the effects of team size, the findings are consistent with that of Adams, Black et al. who studied 2.4 million scientific papers written by research teams in 110 top U.S. research universities over the period 1981-1999 (Adams, Black et al. 2005). The effects of team size are not linear, however. In this sense, the finding of curvilinear effects of team size is consistent with that of Qurashi who studied research groups in the UK, the USA, Bangladesh, Pakistan, and Greece (Qurashi 1984; Qurashi 1991; Qurashi 1993). This finding also supports the claim by Landry and Amara regarding the large transaction costs a team may face due to the large number of members involved (Landry and Amara 1998).

As for the effects of team age is concerned, the findings are consistent with Harrison, Price et al. who studied 144 student project teams (Harrison, Price et al. 2002) and concluded that time serves as a medium for collaboration in teams, allowing members to exchange personal and task-related information, and that, as time passes, increasing collaboration weakens the effects of surface-level (demographic) diversity on team outcomes but strengthens those of deep-level (psychological) diversity, which affects team performance. They are also consistent with Rey-Rocha, Martin-Sempere et.al. who studied the research performance of Spanish senior university researchers in Geology to investigate the effect of team consolidation on individual productivity and found that the number of researchers within the team that reached a stable job position positively affected research productivity (Rey-Rocha, Martin-Sempere et al. 2002).

The curvilinear effects of team age found in this study is rather new and it is an account of the fact that once a team reaches certain “maturity” it tends to ‘rest on its laurels,’ as the popular expression goes. Similarly novel is the finding of the curvilinear effects found of the number of R&D projects active a team has, suggesting the fact that, by trying to do more, a team may risk doing less, due to the implicit cost of managing too many projects at the same time. No curvilinear effects were found regarding the number of PhDs a team has, however, which implies that having one more PhD member is always good for the team.

In contrast to abundant literature from geography economists (Saxenian 1994; Acs, de la Mothe et al. 1996; Landry and Amara 1998; Malo and Geuna 2000; Scott 2001; Liang and Zhu 2002; Stolpe 2002; Casper and Karamanos 2003; McKelvey, Alm et al. 2003; Zitt, Ramanana-Rahary et al. 2003; Bonaccorsi and Daraio 2005), it seems that, at least in the Colombian case, the city or agglomerate where the team is located does not matter. Further investigation is needed to understand how things work at different sizes of the cities where teams are located. In this sense, and according to one of the interviewees, “today, and increasingly, communication costs are making scientific interaction much easier than in the past. Given its strategic location [Colombia is one of the hubs of the submarine cables that connects the rest of South American Countries], the country has one of the largest penetrations rate of the Internet service in the region. More and more researchers are able to work with colleagues located in different areas of the country and of the world."[18]

Finally and more importantly considering the main research question of this dissertation, it is apparent that international research collaboration is a strong predictor of team output. Collaborating or not collaborating may make the difference among teams of the same internal characteristics, same discipline, same sector or same characteristics of the institution of affiliation.

Type of collaboration affects team productivity in different ways, however. As shown in the summary Table 21, while leveraging foreign funding increases team productivity by 33%, and between 3.3 and 4.2 bibliographic products (depending on the bandwidth one chooses to take), hosting foreign researchers is not statistically associated with team productivity. Co-authoring with partners located overseas also appears to have positive and significant effects on team output in Colombia. It increases it by nearly 40% and by between 2.9 and 4.6 bibliographic products.

The effects of international research collaboration on team's productivity also depend on the type of partner the team collaborates with. Thus, teams collaborating with partners from the South are 46% more productive than comparable teams not collaborating with partners from the South. In fact, teams collaborating with the South produce between 3.4 and 4 bibliographic products. In contrast, and contrary to the research hypothesis, collaborating with partners from the North does not seem to affect team productivity.

[18] Translation from Spanish by the author.

Hence, the data shows that collaborating with partners from the South yields greater effects than collaborating with partners from the North. Although these findings contradict some of the hypotheses stated, they make sense. According to one of the scientists interviewed, whereas funding from northern countries are sometimes donations where the supporting institution does not expect to get anything from their funding and therefore does not require the publication of research results, funding from southern countries often involves the matching of local funds and the research they support are commanded for specific purposes; therefore they require the production of bibliographic products. Whereas in the first case the partners do not share the same interests, in the latter both partners do.

More interestingly, as the table shows, different combinations of type of collaboration and origin also yield different effects on team output. Hence, funding form southern countries appears contributing more on team productivity than funding from northern countries and than hosting foreign researchers from southern countries. Hosting researchers from northern countries does not seem to be associated with team output in Colombia.

Table 21: Summary Table: International Research Collaboration and Team Output: ZINB and PSM

	ZINB			PSM			
	%	Count	P>|z|	0.01	T-stat	0.05	T-stat
Internat. Res. Coll.	29	2.66	0.000	2.08	2.71	2.4	3.18
Foreign Researchers	12	0.94	0.098	-0.75	-0.92	-0.23	-0.29
Foreign Funding	33	3.27	0.000	4.15	4.00	4.13	4.07
Co-Authorship *	39	2.91	0.011	4.20	2.29	4.63	2.6

Cont'd

Table 21 *Cont'd*

	ZINB			PSM			
	%	Count	P>|z|	0.01	T-stat	0.05	T-stat
Int. Res. with North	11	1.57	0.140	0.91	1.02	1.68	1.95
Int. Res. with South	46	3.35	0.000	2.97	2.51	4.00	3.45
People from North	-7	-0.24	0.348	-0.64	-0.65	-0.02	-0.02
People from South	32	1.91	0.002	0.39	0.38	1.84	1.87
Funding from North	20	2.12	0.014	3.10	2.82	3.20	3.00
Funding from South	52	4.99	0.000	5.71	2.73	8.7	4.23

Observations: 1889

* Analysis done on 672 Teams

The reasons why collaborating with foreign researchers associated with Colombian teams does not affect team output are not clear. As discussed earlier, it is probably because either the teams cannot get the most they can of their partners knowledge and experience; because foreign researchers cannot exploit all their potential given the material, resource or cognitive constraints they face in Colombia; or because they are overwhelmed with day-to-day issues they have to deal with in Colombia as foreigners.

In addition, the effects of international research collaboration on team productivity observed may be hiding locking-in effects. That is, while collaborative teams may increase their production as a result of collaborating internationally, they may also experience a fall in their productivity as a result of having to manage their time in a collaborative activity. In fact, non-collaborative teams may have more time for producing bibliographic products than collaborative teams. Since both effects may happen simultaneously and therefore it's impossible to disentangle them with the data available, the positive net effect of collaborating internationally found is assumed to result from the fact that the benefits of collaborating outweigh the costs of doing so.

Finally, contrasting the direction of the effects hypothesized in chapter 2 with that actually observed in this chapter one can conclude that a) the “transaction-cost” argument supporting the hypothesis that hosting foreign researchers may entail negative effects on team productivity is not supported by the data. However, the hypothesis cannot be rejected either as no effect was found to be statistically significant; b) the hypothesis of the positive effects of receiving foreign funding based on the “linear-model” argument is confirmed by the data; c) the hypothesis suggesting that co-authoring with colleagues located overseas based on the “more-is-better” argument is also supported by the data; and d) the hypothesis suggesting that collaborating with partners from northern countries have positive effects on team productivity cannot be accepted nor rejected as no effect was found to be statistically significant in either direction. No hypothesis was made regarding the effects of working with partners from the South. Surprisingly, it was found to be positively and strongly related to team output in Colombia. The policy implications of these findings are discussed further in chapter 7. Table 22 summarizes this comparison.

Table 22: Summary of Research Hypotheses and of the Results Obtained Concerning Research Team Output in Colombia

Dependent Variable / Indep. Variable	**RTPC** Hypothesized Effect	Observed Effect	Hypothesis Confirmed?
IRC	Positive	Positive	Yes
Foreign Researchers	Negative	No Significant	Maybe
Foreign Funding	Positive	Positive	Yes
Co-Authorship	Positive	Positive	Yes
North	Positive	No Significant	Maybe
South	-	Positive	-

CHAPTER 6

INTERNATIONAL COLLABORATION AND RESEARCH TEAMS' ABILITY TO CONTRIBUTE TO LOCAL KNOWLEDGE IN COLOMBIA

Between 2003 and 2005 only 681 teams, 36% of the 1889 studied, appeared to be working on projects or produced results that used 'Colombia' as their unit of analysis, their 'laboratory,' or their main object under study. The questions to be answered here are therefore, what factors explain such performance? Did collaborating internationally affect such behavior? If so in what sense? This chapter discusses about the factors affecting team's contribution to local knowledge as reflected by the extent to which the team worked in an R&D project or wrote a bibliographic product whose title or abstract contained the word 'Colombia.'

To do this, the research hypotheses stated earlier are tested using logistic regression models, and the overall impact of international research collaboration on teams' ability to contribute to local knowledge is assessed using control groups and the Propensity Score Matching approach.

The population database supports the analyses of the effects of hosting foreign researchers and of working on projects with foreign funding on the probability of research teams of using 'Colombia' in the title or abstract of their R&D projects and products. It also supports the analyses of the effects of collaborating with partners from the North and the South. The sample data supports the analysis of the effects of co-authoring with colleagues located overseas on the extent to which the teams published journal articles whose title or abstract contained the word 'Colombia.'

Section 6.1 discusses the ways international research collaboration and team characteristics affect teams' ability to contribute to local knowledge. It analyzes the results of the parametric study and of the matching procedure. Section 6.2 discusses the overall impact of international collaboration on teams' ability to contribute to local knowledge using control groups. Section 6.3 discusses the effects of different types of collaboration. Section 6.4 analyses the effects of different types of partners and compares the effects of different combinations of collaboration activity and type of partners. To do that, it uses both multiple regression and non-parametric analyses. Section 6.5 discusses some conclusions based on the results obtained.

6.1 International Research Collaboration, Team Characteristics and Team's Capability to Contribute to Local Knowledge

As shown in Table 23, which is built with the results of the logistic regression, collaborating internationally, having many projects active, being highly productive, working in the social sciences, or working in the NGOs' sector increase a team's odds of working on projects or products that use Colombia in their research process as compared to those of similar characteristics that do not collaborate internationally, have fewer projects active, fewer bibliographic products, work in the natural sciences, or work in the academic sector.

Surprisingly, and contrary to what one would expect, each additional year in operation and each additional member with PhD decreases team's odd of working on issues involving Colombia. As expected, each additional project active, and each additional product a team has increases team's odds of contributing to local knowledge. Similarly, as expected, the odds of working on research projects or producing bibliographic products that use 'Colombia' are larger among teams working in the social sciences as among comparable teams working in the natural sciences; and the odds of involving Colombia in the research process are lower for teams working in the engineering areas than for those of comparable characteristics working in the natural sciences.

Interestingly, the odds of working on projects or products involving Colombia of a team affiliated with the NGO's sector are larger compared to those of similar teams affiliated with the academic sector; and the odds of using 'Colombia' in the research process of a team affiliated with a small institution are lower than that of comparable teams affiliated with large institutions.

The table also shows that team size and team location do not seem to have a significant explanatory power of team's contribution to local knowledge, once we hold all other variables constant.

Table 23: Factors Affecting Team Contribution to Local Knowledge: Logit

Variable	'Colombia' in Prod or Proj
Internat. Res. Coll.	0.347**
	(2.94)
Team Size in 2003	0.001
	(0.12)
Team Age in 2003	-0.039**
	(3.50)
Total PhDs in 2003	-0.078*
	(2.28)
Tot. Proj. in 2003	0.062**
	(5.24)

Cont'd

Table 23 *Cont'd*

Variable	'Colombia' in Prod or Proj
Tot. Bib. Prods. by 2003	0.016**
	(8.20)
Agrosciences	0.297
	(1.23)
Medical Sciences	0.107
	(0.57)
Social Sciences	0.783**
	(4.59)
Humanities	0.100
	(0.64)
Engineering	-0.675**
	(3.35)
Other Sciences	0.227
	(0.82)
Business Sector	0.156
	(0.45)
Government	0.478
	(1.66)
Other Sector	1.011*
	(2.51)
Mid. Home Inst.	-0.026
	(0.21)
Small Home Inst.	-0.381*
	(2.01)
Small City	0.240
	(0.61)
Midsize City	-0.178
	(1.34)
Constant	-1.258**
	(7.55)
Observations	1889

Absolute value of z statistics in parentheses
* Significant at 5%; ** significant at 1%

A Wald Test of the joint effects of team location reveals that there is a 31% probability that the observed results could have occurred by chance (Prob > chi2 = 0.3064). We therefore can drop these variables from the model.

Excluding team size and the location variables, and considering only the variables with statistically significant effects, Table 24 shows that, holding all other variables constant, the odds of a team working in research involving 'Colombia' are 1.42 times as large for collaborating teams as for non-collaborating teams. In fact, holding all other variables constant at their means, collaborating internationally increases team's probability of contributing to local knowledge by 8.1 percentage points.

Table 24: Team Contribution to Local Knowledge: Percentage Change in Odds

	'Colombia' in Prod or Proj	
	%	%StdX
Internat. Res. Coll.	42.0	18.7
Team Age in 2003	-3.7	-19.7
Total PhDs in 2003	-7.2	-15.2
Tot. Proj. in 2003	6.4	50.7
Tot. Bib. Prod by 2003	1.6	104.0
Social Sciences	120.8	33.8
Engineering	-48.7	-20.3
Other Sector	178.5	15.7
Small Home Inst.	-31.0	-13.0

% = percent change in odds for unit increase in X
%StdX = percent change in odds for SD increase in X
SDofX = standard deviation of X

An exploration of quadratic effects shows that team's probability to contribute to local knowledge increases with every additional bibliographic product but at a decreasing rate. The peak at which the positive trend starts to change (160 products) falls beyond our data range, however. In contrast, for each doctorate a team has its odds of involving Colombia in their research process decreases but at a decreasing rate: once the team reaches 9 members with PhD, its odds to involve Colombia raises at an increasing rate (see Table 25). 9 is within our data range.

Table 25: Team Contribution to Local Knowledge: Percentage Change in Odds-Curvilinear Effects

'Colombia' in Prod or Proj	b	z	P>z	%	%StdX	SDofX
Internat. Res. Coll.	0.26497	2.184	0.029	30.3	13.8	0.4878
Total PhDs in 2003	-0.14477	-3.411	0.001	-13.5	-27.4	2.2106
Total PhDs in 2003^2	0.00794	4.031	0.000	0.8	53.0	53.5865
Tot. Bib. Prod by 2003	0.03857	11.094	0.000	3.9	443.1	43.8751
Tot. Bib. Prod by 2003^2	-0.00012	-8.293	0.000	-0.0	-70.3	9945.4326

b = raw coefficient
z = z-score for test of b=0
P>|z| = p-value for z-test
% = percent change in odds for unit increase in X
%StdX = percent change in odds for SD increase in X
SDofX = standard deviation of X

The comparison of both models (with and without the squared terms) through an LR test shows that the improvement (63.44) is greater than the critical value of the chi-square distribution for 2 degrees of freedom at the 0.01 significance level (9.21). We therefore keep these quadratic variables in the model in further analyses.

The next chapter discusses a better way to assess the impact of international research collaboration on the contribution research teams make to local knowledge.

6.2 Assessment of the Effects of International Research Collaboration on Team Contribution to Local Knowledge

As explained in chapter 3, the use of control groups created using the Propensity Score Matching approach allows us to compare the results of collaborating and non collaborating teams based on similar grounds, in this case, their probability of collaborating internationally given their own characteristics. To do this, as was done in chapter 5, the variables for the matching algorithm include the control variables found to have a significant effect on team's choice of collaborating internationally, including the variables regarding team leader characteristics. In this model the variables related to the location of the research team are included back again as the assumption that larger cities offers better opportunities for teams to collaborate internationally than smaller cities has been statistically confirmed in chapter 4.

Thus, Table 26 shows that using a bandwidth of 0.01, the average treatment effect on the treated yields a difference of 8% in the probability of contributing to local knowledge in favor of those teams that collaborate internationally over those that do not. This difference in the odds is smaller to the one observed comparing collaborating and non-collaborating teams using an unmatched sample (17.3%), but it remains significant after the matching algorithm is applied.

Table 26: Team Contribution to Local Knowledge: PSM

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
'Colombia' in Prod or Proj.	Unmatched	. 466032609	.293148309	1728843 .	022313464	7.75
	ATT	.45821727	.380004893	.078212378	.030438146	2.57
	ATU	.29787234	.354063658	.056191317		
	ATE			.064756389		

Table 27 shows the accuracy of the matching procedure. Where the differences in characteristics were statistically significant in the unmatched sample, they became statistically insignificant in the matched sample after the matching algorithm was applied (see the p>|t| column).

Table 27: Assessment of the Matching Quality: PSM-Research Team Contribution to Local Knowledge

Variable	Sample	Mean Treated	Mean Control	t-test t	t-test p>|t|
Internat. Res. Coll.	Unmatched	1	0	.	.
	Matched	1	0	.	.
Team size in 2003	Unmatched	9.072	6.1006	11.39	0.000
	Matched	8.5891	8.4454	0.44	0.659
Team Age in 2003	Unmatched	8.8478	5.6644	11.99	0.000
	Matched	8.5961	8.9459	-0.93	0.351
Total PhDs in 2003	Unmatched	2.4035	0.84996	15.85	0.000
	Matched	2.1825	2.1746	0.07	0.944
Leader Writes in Other Lang.	Unmatched	0.70245	0.4484	11.16	0.000
	Matched	0.69916	0.71552	-0.68	0.496
Leader Studied Overseas	Unmatched	0.66712	0.48222	8.01	0.000
	Matched	0.66156	0.66608	-0.18	0.856
Tot. Proj. in 2003	Unmatched	8.2215	4.1752	13.61	0.000
	Matched	7.6031	7.161	1.17	0.241
Tot. Bib. Prod by 2003	Unmatched	49.432	18.162	16.11	0.000
	Matched	45.213	45.397	-0.07	0.946
Agrosciences	Unmatched	0.06386	0.05898	0.43	0.665
	Matched	0.06546	0.07273	-0.54	0.587
Medical Sciences	Unmatched	0.13315	0.12229	0.69	0.489
	Matched	0.12953	0.12148	0.46	0.646
Social Sciences	Unmatched	0.11821	0.18907	-4.10	0.000
	Matched	0.12117	0.10875	0.74	0.461
Humanities	Unmatched	0.22283	0.25412	-1.55	0.122
	Matched	0.22423	0.20466	0.90	0.366
Engineering	Unmatched	0.11821	0.14397	-1.60	0.109
	Matched	0.12117	0.13587	-0.83	0.406
Other Sciences	Unmatched	0.04755	0.03903	0.90	0.370
	Matched	0.04318	0.03645	0.65	0.515
Mid. Home Inst.	Unmatched	0.2894	0.43539	-6.45	0.000
	Matched	0.29666	0.28142	0.64	0.525
Small Home Inst.	Unmatched	0.1413	0.18474	-2.47	0.014
	Matched	0.13928	0.12631	0.72	0.470
Small City	Unmatched	0.00543	0.02515	-3.20	0.001
	Matched	0.00557	0.00271	0.84	0.400
Midsize City	Unmatched	0.1644	0.25412	-4.62	0.000
	Matched	0.16574	0.17646	-0.54	0.590

Finally, Appendix M shows that, based on the bootstrapping performed through the re-estimation of the results 999 times to test the statistical significance of the findings, we can confidently conclude that *international research collaboration positively affects team's contribution to local knowledge* (Observed Coefficient = .078314, $z = 2.40$, and $P>|z| = 0.016$).

We saw in the previous chapter that the ways international collaboration take place seems do matter for explaining team output. The next section discusses whether these effects are also important for explaining team contribution to local knowledge.

6.3 Type of Collaboration and Team Contribution to Local Knowledge

As in the previous chapter, international research collaboration is measured here in three different ways: hosting foreign researchers, receiving foreign funding, and co-authoring with partners located overseas. While the analysis of the first two types of collaboration is done using the population and the same dependent variable used above, the analysis of the latter is done on the smaller sample and tests the effects of co-authoring on team's probability of publishing an article whose title or abstract contained the word 'Colombia' anytime between 2003 and 2005. In this sample co-authorship is measured for the years 2001 and 2002.

6.3.1 Foreign Members, Foreign Funding, and Teams' Ability to Contribute to Local Knowledge

Similar to the finding regarding team output, Table 28 shows that hosting foreign researchers is not significantly associated with teams' ability to contribute to local knowledge once we hold all other variables constant. In contrast, working on projects with foreign funding increases teams' odds of using 'Colombia' in their research process.

In fact, holding all other variables constant, the odds of a team working in research involving ‘Colombia’ are 1.4 times larger for those working on projects with foreign funding than for those that do not. And, holding all other variables constant at their means, having foreign funding increases team’s probability of contributing to local knowledge by 8%.

Table 28: Foreign Researchers, Foreign Funding and Team Contribution to Local Knowledge

'Colombia' in Prod or Proj	b	Z	P>z	%	%StdX	SDofX
Foreign Researchers	-0.03436	-0.252	0.801	-3.4	-1.4	0.4109
Foreign Funding	0.35973	2.652	0.008	43.3	17.1	0.4381

6.3.2 Co-Authorship and Teams’ Ability to Contribute to Local Knowledge

If we account for the to contribute to local knowledge by the extent to which a team used Colombia in the title or abstract of a journal article published between 2003 and 2005, we find that, as shown in Table 29, and holding the other variables constant, the odds of doing so are 2.21 times larger for those co-authoring with a partner located overseas than for those that do not. This finding is statistically significant at the 0.01 level. Compared to the effects of the other types of collaboration, co-authoring appears to be the one with the greatest impact. In fact, holding all other variables constant at their means, co-authoring with partners located in other countries increases the probability of contributing to local knowledge by 9%. This result has to be interpreted with caution as the two samples used are not quite comparable, however.

Table 29: Co-Authorship with Colleagues Located Overseas and Team Contribution to Local Knowledge

'Colombia' in Prod or Proj	b	z	P>z	%	%StdX	SDofX
Co-Authorship in 2001-2	0.79432	2.784	0.005	121.3	40.1	0.4244

6.4 Type of Partner and Team Contribution to Local Knowledge

Contrary to the effects of working with partners from the South on team output, the effects of collaborating with partners from that origin does not seem to be significantly associated with team contribution to local knowledge. In this case, collaborating with partners from northern countries appears to have greater effects on teams' ability to contribute to local knowledge than collaborating with partners from southern countries. In fact, as shown in Table 30, holding the other variables constant, the odds of a team involving 'Colombia' in its research activities are 1.5 times larger for those working with partners from the North than for those that do not. And working with partners from northern countries increases team's probability of contributing to local knowledge by 10%, holding the other variables constant at their means.

Table 30: Type of Partner and Team Contribution to Local Knowledge

'Colombia' in Prod or Proj	b	z	P>z	%	%StdX	SDofX
Int. Res. with North	0.41782	3.254	0.001	51.9	21.4	0.4638
Int. Res. with South	0.00754	0.049	0.961	0.8	0.3	0.3671

6.4.1 Partner's Origin, Type of Collaboration, and Teams' Ability to Contribute to Local Knowledge

As discussed above, hosting foreign researchers, regardless of their origin, is not significantly associated with teams' ability to contribute to local knowledge. In contrast, working with foreign funding does appear to affect, but how much does it matter the origin of such funding?

Base on the data, and as shown in Table 31, the odds of a team using 'Colombia' in its research activities are 1.5 times larger for those working with projects funded by partners from the North than for those that do not, holding the other variables constant. This finding is statistically significant at the 0.01 level. In fact, holding the other variables constant at their means, working with projects funded by partners from northern countries increases team's probability of contributing to local knowledge by 9%.

In contrast, the odds of a team involving 'Colombia' in its research projects and publications are not statistically significantly larger for those working with projects funded by partners from the South than for those that do not.

Table 31: Type of Partner, Type of Collaboration and Team Contribution to Local Knowledge

Colombia' in Prod or Proj	b	z	P>z	%	%StdX	SDofX
People from North	0.13350	0.825	0.409	14.3	4.7	0.3412
People from South	-0.19289	-1.059	0.290	-17.5	-5.8	0.3084
Funding from North	0.40522	2.886	0.004	50.0	18.8	0.4261
Funding from South	0.21284	0.917	0.359	23.7	5.3	0.2430

6.5 Conclusions

Based on the two approaches used, that is, the logistic regression and the non-parametric models, we found that collaborating internationally is positively associated with team' ability to contribute to local knowledge. Furthermore, we found that both, type of collaboration activity and type of partner do matter at explaining the effects of international research collaboration on teams' ability to contribute to local knowledge in Colombia.

In this sense, we found that co-authoring with partners located overseas and receiving foreign funding positively affect team performance. We also found that, similar to the factors affecting research team output, hosting foreign researchers does not seem to have a significant effect on team research orientation.

Contrary to the findings regarding the effects of collaborating with partners from the South on research team output, collaborating with such partners is not significantly associated with team contribution to local knowledge. In this case working with partners from northern countries appears to have large effects on teams' ability to contribute to local knowledge once we hold all other variables constant.

Finally, working with projects funded by northern countries appears to have the greatest effect on team contribution to local knowledge. Table 32 summarizes these findings.

Table 32: Summary Table: International Research Collaboration and Team Contribution to Local Knowledge: Logit and PSM

	Logit			PSM (%)			
	% (1)	%(2)	P>|z|	0.01	T-stat	0.05	T-stat
Internat. Res. Coll.	30.3	6.0	0.029	7.8	2.57	8.0	2.73
Foreign Researchers	-3.4	-0.7	0.801	-0.2	-0.08	0.5	0.16
Foreign Funding	43.3	8.3	0.008	12.3	3.56	13.3	4.08
Co-Authorship *	121.3	9.0	0.005	16.0	3.11	11.8	2.46
Int. Res. with North	51.9	9.6	0.001	11.9	3.57	11.2	3.60
Int. Res. with South	0.8	0.2	0.961	2.1	0.64	4.3	1.31
People from North	14.3	3.1	0.409	2.9	0.80	4.4	1.26
People from South	-17.5	-4.2	0.290	-2.8	-0.76	-0.5	-0.15
Funding from North	50.0	9.4	0.004	12.8	3.58	14.2	4.24
Funding from South	23.7	4.9	0.359	9.9	1.98	17.2	3.58

Observations: 1889

* Analysis done on 672 Teams

%(1): Percentage Change in Odds

%(2): Changes in Predicted Probabilities for 'Colombia' in Prod or Proj

Contrasting the hypotheses proposed in chapter 2 with the results obtained in this chapter we conclude that a) the “commitment” argument suggested to sustain the hypothesized positive effect of hosting foreign researchers on teams’ ability to contribute to local knowledge is not supported nor rejected as the effect found is not statistically significant; b) the “opportunity” argument used to hypothesize that working with foreign funding may negatively affect teams’ ability to contribute to local knowledge is rejected as the effect found is positive and statistically significant; c) the “outsourcing” argument used to hypothesize a negative effects of co-authoring on the probability of teams to involve Colombia in their research process is rejected for the same reason; and d) the “complementarity-based-on-epistemological-similarity” argument used to suggest that collaborating with partners from southern countries may positively affect team orientation cannot be rejected nor confirmed as the effects found were not statistically significant. Table 33 summarizes this comparison.

Table 33: Summary of Research Hypotheses and of the Results Obtained Concerning Research Teams' Ability to Contribute to Local Knowledge in Colombia

Dependent Variable / Indep. Variable	**RTCLK** Hypothesized Effect	Observed Effect	Hypothesis Confirmed?
IRC	Negative	Positive	No
Foreign Researchers	Positive	No Significant	Maybe
Foreign Funding	Negative	Positive	No
Co-Authorship	Negative	Positive	No
North	-	Positive	-
South	Positive	No Significant	Maybe

The next chapter discusses further the policy implications of the findings reported here.

CHAPTER 7

OVERALL THEORETICAL AND POLICY IMPLICATIONS

This study attempts to contribute to the policy debate and study of the determinants of local S&T capabilities, and particularly on the role international research collaboration plays on the performance of research teams in developing countries, using Colombia as a case study. In particular, this dissertation provides elements useful to increase current understanding of the determinants of team output and of teams' ability to contribute to local knowledge. In so doing, it uses econometric approaches for testing the research hypotheses supporting the theoretical model proposed, and for assessing the overall impact of research collaboration on the performance of research teams while controlling for endogeneity and selection bias.

In particular, the findings of this study contribute to current literature on research policy, research evaluation, studies of S&T and development, foreign policy, and sociology of science and technology among other areas of research. In fact, the Colombian case is useful for understanding the role international collaboration plays (and could play) in developed countries. Despite the unique characteristics of the country in several respects, those aspects associated with the characteristics of its national science and technology system (structure, dynamic and performance) and the ways the country enters into the global arena are arguably similar to most countries classified as 'peripheral countries' (highly specialized in and dependent on few primary goods, highly politically and economically dependent on few developed countries, showing high levels of internal inequality, small markets, weak democracies, and showing slow progress at meeting the millennium development goals).

7.1 The Results and their Implications

7.1.1 Three Publication Increase between Collaborating and Non-Collaborating Teams

The results show that research team output and teams' ability to contribute to local knowledge depend in part on team internal characteristics including its size, its age, the level of education of its members, and their R&D processes, as well as on the field it specializes in, the sector where it performs its activities, the characteristics of its home institution and the characteristics of the city it is located.

More importantly, the study shows that international research collaboration significantly affects the Colombian S&T system in a positive way. In fact, collaborating internationally nearly doubles average team productivity. It increases team production by almost 3 bibliographic products. In other words, the teams that had the opportunity to produce bibliographic products and that reported zero productivity during the period observed (which as we showed before are numerous) would have produced up to 9 bibliographic products had they collaborated internationally. This is a relatively large contribution since the median number of bibliographic products of the entire distribution observed was 4 bibliographical products. Had all the teams collaborated, the median articles produced in Colombia would have been much larger than what currently is.

7.1.2 Seven Percent Increase in the Odds of Involving Colombia in the Research Process between Collaborating and Non-Collaborating Teams

International collaboration also increases the odds of involving Colombia in team's research processes by between 6% and 8%, holding all other variables constant. Moreover, the teams that reported not having involved Colombia in their research process during the period observed would have increased their probability of doing so at least by 17% had they collaborated internationally.

In sum, teams that could have collaborated and chose not to do so assumed high opportunity costs that are not only burdensome for themselves in an increasingly integrated and competitive world, but also, and given the special characteristics of their activities, for the Colombian society as a whole. That is, in addition to a rise in the stock of knowledge produced, the Colombian society could have more opportunities to benefit from the knowledge produced and diffused through scientific and technological publications, and through involving Colombia either as laboratory or object of research.

7.1.3 Working with Foreign Researchers Does Not Seem to Significantly Affect Team Performance

The study also shows that the effects of international collaboration on team performance depend on the type of collaboration chosen. Although working with foreign funded projects and co-authoring with partners located overseas positively affect team's S&T capabilities, hosting foreign researchers does not seem to significantly affect team performance. Arguably, high transaction costs are at the root of this issue. In fact, the possible occurrence of locking-in effects for the teams that host foreign researchers may be affecting their productivity. Since they are involved in collaboration activities, research teams do not have the same time to produce bibliographic products as non-collaborating teams do. Indeed, the net effect of collaborating internationally may consist of two opposite effects: first the increased productivity resulting from the increase creativity derived from the collaborative activity, and second the reduced productivity during the collaboration activity as a result of administration and transaction costs. Since both effects cannot be disentangled, we only observe the net effect and have to take this into account when interpreting results. Although we therefore should expect an initial negative effect from hosting foreign researchers (and in fact any kind of collaboration), a successful collaboration should overcompensate for this initial fall in the long term. So if we are able to observe the outcome of the teams for a reasonable time after begin or end of the collaboration activity, the occurrence of locking-in effects would pose fewer problems.

In addition, the interviews show that sometimes teams engage foreign researchers as a result of interinstitutional internship programs which usually result in a burden the teams are not prepared (or willing) to handle. In other cases, foreign researchers either do not spend enough time in the country and therefore are not able to produce new products, or they are so concentrated in dealing with their living and teaching experience that leave a small portion of their time for doing research. Not all the interviewees coincided with this view, however. In fact, many team leaders see foreign researchers as people much more organized and better prepared than Colombian researchers. According to an interviewee, “I prefer to work with foreign researchers rather than with local researchers because they are more respectful of the intellectual property of what is being produced in the team. Their contribution to the team may take some time to materialize but it is always positive[19]." In consequence governments, together with home institutions, should help in reducing such transaction costs by supporting the teams in the management of human capital coming from aboard.

Alternatively, and since hosting foreign researchers is not significantly associated with the teams' ability to contribute to local knowledge either, a better selection process of those foreign researchers willing to work in Colombia (which by the way are rare) could also be put in place. Possible explanation as to why foreign researchers are not significantly contributing to team research productivity or team research orientation in Colombia may be that either the teams are not being able to absorb the advantages foreign researchers can offer, or because their role in research teams are other than supporting the research endeavor. These are hypotheses that are worth exploring further.

7.1.4 Collaborating with Southern Countries Rises Team Productivity More Than Collaborating with Northern Countries

The study also found that the effects of international collaboration on team performance depend on the type of partner involved. Collaborating with partners from the South yields greater impact on team output than collaborating with partners from the North. From the policy point of view, this finding should be taken into account in systems like Colombia which has traditionally put more emphasis on North-South collaboration than on South-South collaboration. The reasons why horizontal collaboration shows greater impact on team productivity than vertical collaboration may be associated with

[19] Translation from Spanish by the author.

what in this dissertation is called the ‘epistemological similarity argument’, which states that teams working in countries with similar characteristics, types of problems and level of S&T development may be in a better position for overcoming the transaction costs typically present in a collaborative enterprise. Looking at the research issues from a similar perspectives and working with materials the partners are more familiar with may contribute to their productivity.

Another reason why working with the South yields greater productivity than working with the North is that, given the relatively large economic effort the southern country makes to collaborate, they are more interested than northern countries to assure verifiable results, that is, the accountability upon the investment done is stricter by southern partner countries than by northern partner countries, and reporting publications is usually the preferred way to show that the investment done (in money or human resources) did pay the effort. These effects deserve further investigation, however.

7.1.5 Collaborating with Northern Countries Rises Team’s Odds of Involving Colombia in their Research Process More Than Collaborating with Southern Countries

According to the study, collaborating with northern countries contributes more to teams’ ability to add to local knowledge than collaborating with southern countries. This finding is also a matter to be taken into account in public debates as the relationship between North and South has traditionally been seen as an imbalanced process, where, the argument goes, the South gives more than what they get. Although this finding does not reject nor support this claim, as we do not quantify what the local teams invest or, alternatively, how many the partners from the North gain, it seems fair to conclude that, at least in the Colombian case, collaborating with northern countries does pay: it increases team's odds of involving Colombia in their research process more than it does when collaborating with southern countries and even more than when the team does not collaborate internationally. Although the overall impact of collaborating with northern countries is relatively small, at least it is not negative, as one would have hypothesized based on the 'outsourcing argument' discussed earlier.

The reason why North-South collaboration positively affects team orientation may be the result of what we called in this study the 'diversity argument' where, in line with Granovetter's and Burt's claims, one have more to learn from our differents than from our peers (i.e. the 'strength-of-weak-ties' and the 'structural-holes' arguments) (Granovetter 1973; Granovetter 1983; Burt 2004). By studying scientific issues with different materials and from different perspectives one gets better ideas as to how to deal with local issues.

However, the policy challenge is not to prefer one type of partner over another, but to understand why the effects are different, and, as a consequence, to design and implement the mechanisms through which teams and society can benefit the most.

7.1.6 Working with Projects Funded by Southern Institutions Yields the Greatest Impact on Team Performance

According to the data, preferring a specific combination of collaboration activity (hosting foreign researchers, working with foreign funding, and co-authoring) and type of partner (North and South) also seems to matter.

In fact, although working with projects funded by foreign institutions yields the greatest impact on team performance, both in terms of productivity and of the probability of involving Colombia, it is working with projects funded by institutions from the South that shows the greatest positive impact on team's performance. Indeed, working with projects funded by institutions from the South contributes between 5 and 9 bibliographic products alone, that is, it more than doubles average team production. On the other hand, working with projects funded by institutions from the North has greater impact on teams' ability to contribute to local knowledge than on team's production. In fact, it has the greatest impact on team's orientation: it raises team's probability of involving Colombia in the research process by between 5% and 17%.

However, as the results show, the overall effects of working with projects funded by institutions from the South on team output is larger than the overall effects of working with projects funded by institutions from the North on team's probability of involving Colombia in their research process. To conclude this we assume that both effects are comparable, however. This is a strong assumption since it is hard to conclude that a 7% increase in the probability of involving Colombia in the team's research process represents a lower positive effect than an increase by three products. The debate is open, however.

The reasons why these findings result are that there is a combined effect of both collaborating with foreign funded projects and of working with specific types of partners which, as discussed earlier, may affect team output and team orientation differently. However, as discussed earlier, to better understand the reasons why a specific combination of collaboration activity and partner would be preferred over another, requires further investigation.

In sum, the patterns found here are illustrative of the challenges policy makers would face if they were interested on implementing tools strategically oriented at achieving the greatest benefits possible from their support to team performance and to their process of internationalization. The findings reported here are the first steps in that direction. In addition, a more illustrated decision-making process would help the teams themselves in gaining from collaborating internationally, and by that means, positively affecting the society as a whole.

7.1.7 Policy Recommendations

Despite the positive effects derived from collaborating internationally, it is a matter of concern to witness that few research teams actually collaborate internationally in Colombia. In fact, this clearly is an account of a market-failure situation that justifies government intervention. In this sense, it is important to note that governments may have different levels of influence to positively affect research team performance. For example, they cannot influence some of the characteristics discussed here such as team scientific specialization, their affiliation to a specific sector or institution, their location, or their age. Governments may or may not be able to influence team size, or their internal dynamism. By contrast, they may make substantial contributions to team output and ability to contribute to local knowledge by facilitating international collaboration or by encouraging higher standards of member quality. For these reasons, and based on the impacts these two variables have for explaining team bibliographic production and team orientation, there should be aggressive policies stimulating both international research collaboration and the strengthening of S&T human capital in developing countries.

The understanding of what explains international collaboration is an important input for the design of policies in S&T. In this sense, we found that the number of PhDs, the number of projects active, team age, and the characteristics of the team leader were the factors with the greatest impact on the decision of collaborating internationally. This leads to the conclusion that what Colombian teams need is more support to engage members with PhD, more funding for the performance of R&D activities, more stability for their members, and leaders able to write well in a second language and with foreign education.

More importantly, and according to the interviews done, the main reason why Colombian teams do not collaborate internationally is because they lack direct public support for such activities. Public policies could include tools to encourage physical interaction among scientists, network creation, network membership and operation, access to external information, and diplomatic support among other alternatives. Several ways governments can foster international research collaboration include:

1. Promoting the participation of local teams in international projects
2. Supporting workshop participation by local scientists, when they are held overseas
3. Supporting international scientific workshops organized by local research institutions
4. Funding international dissemination of information related to local scientific activities and communities (through the web, the internationalization of local scientific journals, or the countries' diplomatic representations overseas)

5. Funding local dissemination of information related to international scientific activities (including translation of relevant work into local language) and communities
6. Sponsoring courses of foreign languages for local researchers
7. Supporting the negotiation of collaborative agreements between institutions
8. Sponsoring international education at the graduate level
9. Training local researchers on international cooperation for the performance of science, technology and innovation activities
10. Supporting international research internships and networks
11. Sponsoring local access to international databases (both Journal and Patent databases)
12. Encouraging university-government-enterprise partnerships
13. Training local scientists and engineers in intellectual property rights issues
14. Supporting programs oriented at attracting foreign researchers and national researchers living overseas to work in or with local institutions
15. Promoting workshops where the international research collaboration is the object of study

For this goal, public funding for these activities should dramatically increase. There is no information on government expenditure on the internationalization of local S&T available, but it is easy to guess that the budget assigned to facilitate international research collaboration in Colombia is meager.

A second policy implication of the findings of this research refers to the strategy of supporting local research teams as the basic structural unit of the national S&T system. As discussed through the dissertation, research teams are an incontestable need for the advancement of science and technology as they are the building blocks of the national innovation system. For these reasons, they should be seen as the target of R&D policies oriented toward the development of local scientific and technological capabilities. For international research collaboration to positively impact the local system, it has to be mediated or channeled, by the local units of research and development. In this context, research teams act as the bridges between the society and the external world, which is full of opportunities somewhat unexploited by local communities in developing countries.

For this purpose, a strategic research team policy that takes into account their structure, dynamism and potential should be developed. If there are structural deficiencies that prevent teams from capturing the benefits of international research collaboration, a set of tools should be designed to increase their readiness and ability to become multipliers of such benefits.

For instance, research teams in developing countries need to increase their absorptive capacity (Cohen and Levinthal 1990) of international research collaboration. To do that, it is necessary to combine human capital policies with the support to local infrastructure for research and with policies oriented at encouraging the diffusion of ideas and skills in Colombia.

As for the policies oriented at increasing teams' quality is concerned, the analyses showed that a large portion of team productivity and of teams' ability to contribute to local knowledge is explained by the number of PhD holders a team has, and that these effects rise almost linearly. In some cases, the effects of having PhDs were larger than that of collaborating internationally.

For this reason, Colombian government should make an important improvement and raise its support to researchers' education and training. In fact, in Colombia less than 15% of the researchers are PhD holders, and less than 62% of the teams have a PhD graduate. Colombia has one of the lowest ratios of PhD graduates per million inhabitants in the region (less than 1.5 per million a year). Colombian expenditure in PhD education is one of the lowest in the region. Without such human capital, both team productivity and team ability to contribute to local knowledge would remain at the low levels they currently are.

However, the study also found that the number of doctorates a team has was negatively associated with teams' ability to contribute to local knowledge. This implies a trade-off that suggests the need to evaluate the pertinence of the graduate education received by Colombian students. More investigation in this respect is worth doing, however as the study also shows that such negative effects tends to reverse among teams with 9 members with PhD, when the effects become positive at an increasing rate. Similarly interesting, a matter that deserves further investigation refers to the curvilinear effects found in the size of teams, their age, and the number of R&D projects active they manage. These effects, implies the identification of specific characteristics a team should have in order to be productive and relevant.

In sum, the combination of the fact that few teams collaborate internationally; that there are high opportunity costs associated with not collaborating internationally; that there a potentially high transaction costs associated with managing international collaboration; that there might be low levels of team absorptive capacity; that there seems to be a lack of strategic selection processes in place; and that teams tend to overlook south-south collaboration results in a clear justification for government intervention in developing countries.

7.2 Generalizability of the Conclusions

The results of this study are potentially generalizable to the countries sharing several characteristics with Colombia both in terms of its national science, technology and innovation system (STIS)'s characteristics, dynamic and performance, and in terms of its overall social, political, historical, and macroeconomic conditions. More importantly, in the author's opinion the Colombian case is generalizable to those developing countries with a minimum level of absorptive capacity, that is, to those lacking the ability to take substantial advantage from the contributions made by their foreign scientific peers.

In this sense and as for the characteristics of the local STIS, the results can arguably be generalized to the following types of countries:

a) 'S&T-Developing Countries', defined by the InterAcademy Council (IAC) as those that have scientific and technological strength in one or more research areas but lack important aspects of S&T capacity in personnel, infrastructure, investment, institutions, and regulatory framework (IAC 2004)[20];

[20] The IAC classification also includes 'S&T-Lagging Countries', that is, those with little scientific or technological research strengths and no discernable overall S&T capacity in the terms defined; 'S&T-Proficient Countries; and 'S&T-Advanced Countries.' http://www.interacademycouncil.net/

b) Countries in the early stages of innovation system development, defined by the United Nations Industrial Development Organization (UNIDO) as those that have establishing threshold conditions for the emergency of innovation systems but still fail in promoting functional innovation systems for innovation-based growth (UNIDO 2005) [21];

c) Countries with relatively low levels of 'technological readiness' or belonging to the 'non-core countries', defined by the World Economic Forum (WEF) as those that show relatively low scientific and technological absorption, slow pace of technological innovation, low levels of expenditure in R&D, few collaborations between academy and the business community, few patent registrations, and export mainly primary goods (WEF 2005).

d) 'Latecomers', defined by Archibugi and Coco as those that "in one way or another, try to stimulate their technology growth parallel to their development efforts: technological infrastructure and formation of human skills (but fail to achieve large numbers of technological innovations)" (Archibugi and Coco 2004)[2223].

---

[21] The UNIDO identifies 3 phases of innovation system development. In terms of the respective strategic priorities, they consist of: first, establishing threshold conditions for the emergency of innovation systems; second, promoting functional innovation systems for innovation-based growth; and third, promoting the growth of differentiated and specialized innovation systems, which systematically generate innovative responses to emerging opportunities (UNIDO, 2005 p.73).

[22] The ArCo Index is also used to classify the countries studied as 'leaders,' 'potential leaders,' and 'marginalized'.

[23] Other efforts designed to group countries sharing similar characteristics and following similar patterns include the Science and Technology Capacity Index produced for the RAND Corporation (Wagner, C. S., I. Brahmakulam, et al. (2001). Science and Technology Collaboration: Building Capacities in Developing Countries? Santa Monica, CA, RAND.); the Technology Achievement Index produced by the UN Development Program (UNDP (2001). Human Development Report 2001: Making New Technologies Work for Human Development. New York, Oxford University Press.); the Industrial Development Scoreboard produced by the UN Industrial Development Organization (UNIDO (2002). Industrial Development Report 2002-2003. Competing through Innovation and Learning. Vienna, United Nations Industrial Development Organization.); and the S&T Capacity Index proposed by Sagasti (Sagasti, F.

Other relevant aspects to take into account to judge whether or not the conclusions of this study can be generalized to other countries relate to the process of institutionalization of the local science and technology system that characterizes not only Colombia but also many developing countries. In fact –although far from satisfactorily meeting the definition of 'National Innovation Systems' proposed by Freeman, Lundvall or Nelson to describe the process of creation, diffusion and use of knowledge and innovations- in some developing countries similar to Colombia there is actually a dynamic process of formalization of the production and support to S&T activities taking place. As Eduardo Martinez posits, Latin America seems to be experiencing at least 10 major changes having mixed effects on its transition toward knowledge-based societies. According to the author, these changes include: a) a transition from restrictive public policies to modernizing policies, where the focus on S&T activities is being replaced by a focus on innovation; b) a shift from an emphasis on the sustained supply of knowledge needed to meet long lasting social demands, to an emphasis on the short-term market demands of skills, techniques, and technologies; c) from traditional R&D management practices and resource allocation focused on control, to a more efficient management based on performance evaluation, and chain-link processes; d) from an intervening role of the Government in supporting R&D, to a role of Governments as facilitators of the creation of the so-called NISs; e) from a lack of quality control of the higher education system, to the demanding process of evaluation and accreditation now in place; f) a transition in progress toward smaller Governments; g) from formal guidance and

---

(2004). Knowledge and Innovation for Development: The Sisyphus Challenge of the 21st Century. Northampton, MA., Edward Elgar.)

regulation to institutional “laissez-faire”; h) an increased support to S&T policy design of multilateral organizations such as the Inter American Development Bank and the World Bank as well as a greater reliance on international cooperation; i) from ‘closed’ national systems to globalized R&D and localization strategies; and j) from nation-economies to region-economies and institutional networks of knowledge (Martinez 2005). According to the author these shifts are not necessarily leading to set the basis necessary for the purpose of the Latin American countries to become knowledge-based societies, and in most of the cases they seem to follow policies designed, followed and found successful in developed countries where the process of institutionalization is rather mature.

Similarly, aspects not directly associated with the local science, technology and innovation system’s characteristics, dynamics and performance that may also affect the applicability and generalizability of the conclusions of this study relate to a) the degree of openness to foreign science and technology; b) the levels of self confidence; c) the vulnerability to external and internal chocks including international conflicts and conjectural economic crises, among other. In many countries with similar STI systems to Colombia’s, these aspects may not mirror the Colombian case and therefore may not be comparable. However, to know how much these factors affect the generalizability of the conclusions of this study is hard to judge.

Finally, in the author’s opinion, the main differences between the Colombian case and other cases would be not so much on the direction and the characteristics of the impacts found but on the possibility for accounting for such impacts. As discussed earlier, to assess the effect of international research collaboration, it seems easier to use a developing country as a case study than a developed country, mostly because in the latter case distinguishing between domestic and foreign partners is much harder. In addition, very few countries have the type of information that was used in this study. This is an issue discussed in the next section.

7.3 Agenda for Future Research

To better account for the determinants of team performance and the ways international research collaboration affects such performance, the observation of other factors would be needed. These factors include individuals' characteristics such as **a)** researchers' *age* (Cole 1979; Diamond 1985; Levin and Stephan 1991; Stephan and Levin 1997; Dietz 2004); **b)** *sex* (Fox and Faver 1985; Long 1992; Long, Allison et al. 1993; Prpic 2002); **c)** *level of education* (Becker 1964; Barro and Lee 2001; Bozeman, Dietz et al. 2001; David and Goddard L 2001); **d)** *experience* (Dietz 2004; Melin 2004); **e)** *cosmopolitanism* (Lee 2004; Lee and Bozeman 2005). Motivations for collaboration (Melin 2000), collaboration strategies (Moed 2000), additional demographic and psychological characteristics of the team leader, public policies (Georghiou 1998; Georghiou 2001; Wagner, Brahmakulam et al. 2001; Smeby and Trondal 2005), and other types of collaborative activities should also be considered.

Alternatively, further research can be done regarding the effects and processes of research collaboration at the discipline or field level, by sector, by city or region of location, by institution of affiliation, by government program, and by partner country or region. Similarly, empirical analysis is needed regarding the effects of other type of international collaborations such as sharing equipment, hosting visitors in the framework of internship programs, and other more informal ways of collaboration. It would be interesting to compare the effects of collaborating internationally with studying abroad; to compare patterns of collaboration activities and impacts using other developing countries; and to compare the effects of international research collaboration in developing versus developed countries. Qualitative research is needed on the characteristics, processes, determinants, and impacts of teamworking both in developed and developing countries, and in 'best practices' resulting from case studies.

Finally, research is needed using other dependent variables such as research team's contribution to the creation of S&T human capital, research team's innovative capacity, and research team output quality. In sum, the research done here makes important contributions to the literature on sociology of science, research policy and research evaluation but the topic demands many more studies to be made in the future.

APPENDIX A

BIBLIOGRAPHIC PRODUCTS

Cod. in the DB	Product
11	Artículos publicados en revistas científicas
111	Completo
112	Corto (Resumen)
113	Revisión (Survey)
114	Caso clínico
12	Trabajos en eventos (Capítulos de memoria)
121	Completo
122	Resumen
13	Libros y capítulos de libros publicados
131	Libro publicado
132	Capítulo de libro publicado
133	Libro organizado o edición
134	Libro resultado de investigación
1C	Prefacio, epílogo
1C1	Prefacio
1C2	Epílogo
1C3	Presentación
1C4	Introducción
1Z	Otra producción bibliográfica
1Z1	Documento de trabajo (working paper)
1Z2	Otra produccion bibliografica
2I	Informes de investigación
2Z5	Base de datos de referencia para investigación
2Z6	Colección biológica de referencia con información sistematizada

APPENDIX B

CLASSIFICATION OF PARTNER COUNTRIES

Country	NORTH	SOUTH
Albania		X
Argelia		X
Argentina		X
Australia	X	
Austria	X	
Bahamas		X
Bangladesh		X
Belarus		X
Belgium	X	
Belize		X
Benin		X
Bermuda		X
Bolivia		X
Botswana		X
Brazil		X
Brunei		X
Bulgaria	X	
Burkina Faso		X
Cameroon		X
Canada	X	
Chad		X
Chile		X
China		X
Colombia		X
Costa Rica		X
Croatia		X
Cuba		X
Czech Republic	X	
Denmark	X	
Dominican Republic		X
Ecuador		X
Egypt		X
Ethiopia		X
Fiji		X
Finland	X	
France	X	
Gambia		X
Germany	X	
Ghana		X
Greece		X
Guadeloupe		X
Guatemala		X
Guyana		X
Honduras		X
Hong Kong		X
Hungary	X	
Iceland	X	
India		X
Indonesia		X
Iran		X
Iraq		X
Ireland	X	
Israel	X	
Italy	X	
Jamaica		X
Japan	X	
Kazakhstan		X
Kenya		X
Korea Dem. Rep.		X
Korea Rep.		X
Kuwait		X
Latvia		X
Lebanon		X

Country	NORTH	SOUTH
Lithuania	X	
Macedonia		X
Malagasy Republ		X
Malawi		X
Malaysia		X
Mali		X
Mauritius		X
Mexico		X
Moldova	X	
Mongolia		X
Morocco		X
Nepal		X
Netherlands	X	
New Caledonia		X
New Zealand	X	
Nicaragua		X
Nigeria		X
Norway	X	
Oman		X
Pakistan		X
Panama		X
Papua N Guinea		X
Paraguay		X
Peru		X
Philippines		X
Poland	X	
Portugal	X	
Puerto Rico		X
Rep of Georgia		X
Romania	X	
Russian Federation	X	
Saudi Arabia		X
Senegal		X
Serbia Montenegro		X
Singapore		X
Slovakia	X	
Slovenia		X
South Africa		X
Spain	X	
Sri Lanka		X
Sudan		X
Swaziland		X
Sweden	X	
Switzerland	X	
Syria		X
Taiwan		X
Tanzania		X
Thailand		X
Trinidad and Tobago		X
Tunisia		X
Turkey		X
Uganda		X
Ukraine	X	
United Kingdom	X	
United States	X	
Uruguay		X
Venda		X
Venezuela		X
Vietnam		X
W Ind Assoc St		X
Yugoslavia		X
Zambia		X
Zimbabwe		X

APPENDIX C

SAMPLING STRATEGY

Total Teams Registered in 2005 by Colciencias	3342
Teams Excluded (of which:)	1453
Teams created in 2005 (i)	6
Teams with less than 2 R/E active* by 2003 (ii)	919
Teams with no R&D projects active** btw 2003 and 2005 (iii)	1172
Teams used in the analysis of the larger sample	1889
Teams used in the analysis of the smaller sample***	672

Sums do not add up due to double counting of teams' attributes

(i) This is justified as many teams may form only to be registered as such by Colciencias during the registration process

(ii) This is justified as there is no "team" of only one member

(iii) This is justified as there is no "research team" without at least one R&D project acting as their main common activity

* An active R/E is a Researcher or Engineer that reports research activities done in 2003 or before

** An active R&D project refers to Research and Development work reported as being in progress between 2003 and 2005

** The smaller sample was created using the same criteria of the larger sample, includes teams created by March 2004 (to assure random selection), and excludes teams working in the social sciences or in the humanities

APPENDIX D

TEAM LOCATION AND CITY SIZE

1. SMALL	2. MEDIUM	3. LARGE
Arauca	Armenia	Barranquilla
Bello*	Barrancabermeja	Bogota
Bojaca*	Cartagena	Bucaramanga
Carepa (AN)*	Chinchina*	Cajica*
Cartago*	Chiquinquira	Cali
Cienaga	Cucuta	Chia*
Duitama*	Fusagasuga	Envigado*
Florencia	Manizales	Medellin
Florida Blanca*	Pasto	Mosquera*
Ibague	Pereira	Rio Negro (AN)*
Leticia	Popayan	
Monteria	Quibdo	
Neiva*	Rioacha	
Palmira*	San Andres	
Pamplona	Santa Marta	
Pie de Cuesta (SN)*	Sincewlejo	
	Sogamoso	
	Tulua*	
	Tumaco	
	Tunja	
	Ubate	
	Valledupar	
	Villavicencio	

* Towns near big cities or in a cluster of cities

APPENDIX E

EQUIVALENCES ISI-UNESCO-SILAC05

Equivalences: ISI-UNESCO.

(Some fields are classified more than once)

Natural Sciences	Medical Sciences
Acoustics	Allergy
Astronomy & Astrophysics	Anatomy & Morphology
Biochemical Research Methods	Andrology
Biochemistry & Molecular Biology	Anesthesiology
Biodiversity Conservation	CARDIAC & CARDIOVASCULAR SYSTEM
Biology	Cardiac & Cardiovascular Systems
Biophysics	Chemistry, Medicinal
Biotechnology & Applied Microbiology	Clinical Neurology
Cell Biology	Critical Care Medicine
Chemistry, Analytical	Dentistry, Oral Surgery & Medicine
Chemistry, Applied	Dermatology
Chemistry, Inorganic & Nuclear	ENGINEERING, BIOMEDICAL
Chemistry, Medicinal	Emergency Medicine
Chemistry, Multidisciplinary	Endocrinology & Metabolism
Chemistry, Organic	Ergonomics
Chemistry, Physical	GERONTOLOGY
Crystallography	Gastroenterology & Hepatology
Developmental Biology	Genetics & Heredity
ENGINEERING, OCEAN	Geriatrics & Gerontology
Ecology	Health Care Sciences & Services
Electrochemistry	Hematology
Engineering, Chemical	Immunology
Engineering, Environmental	Infectious Diseases
Engineering, Geological	Integrative & Complementary Medicine
Entomology	Medical Informatics
Environmental Sciences	Medical Laboratory Technology
Evolutionary Biology	Medicine, General & Internal
Geochemistry & Geophysics	Medicine, Legal
Geography	Medicine, Research & Experimental
Geography, Physical	NEUROIMAGING
Geology	Neurosciences
Geosciences, Multidisciplinary	Nutrition & Dietetics
Limnology	Obstetrics & Gynecology
MATERIALS SCIENCE, PAPER & WOOD	Oncology
MICROSCOPY	Ophthalmology
Marine & Freshwater Biology	Orthopedics
Materials Science, Biomaterials	Otorhinolaryngology
Materials Science, Ceramics	Parasitology
Materials Science, Characterization & Testing	Pathology
Materials Science, Coatings & Films	Pediatrics
Materials Science, Composites	Peripheral Vascular Disease
Materials Science, Multidisciplinary	Pharmacology & Pharmacy
Mathematics	Psychiatry
Mathematics, Applied	Psychology, Biological
Mathematics, Interdisciplinary Applications	Psychology, Clinical
Meteorology & Atmospheric Sciences	Public, Environmental & Occupational Health
Microbiology	Radiology, Nuclear Medicine & Medical Imaging
Mineralogy	Rehabilitation
Mycology	Respiratory System
Nanoscience & Nanotechnology	Rheumatology
Nuclear Science & Technology	Sport Sciences
Oceanography	Substance Abuse
Optics	Surgery
Ornithology	Toxicology
Physics, Applied	Transplantation
Physics, Atomic, Molecular & Chemical	Tropical Medicine
Physics, Condensed Matter	Urology & Nephrology
Physics, Fluids & Plasmas	Veterinary Sciences
Physics, Mathematical	Virology
Physics, Multidisciplinary	
Physics, Nuclear	
Physics, Particles & Fields	
Physiology	
Polymer Science	
Remote Sensing	
Reproductive Biology	
Spectroscopy	
Thermodynamics	
Veterinary Sciences	
Water Resources	
Zoology	

Author: Gonzalo Ordonez M., 2007.

APPENDIX E Cont'd:
Equivalences: ISI-UNESCO Contd.

AgroSciences
Agricultural Engineering
Agriculture, Dairy & Animal Science
Agriculture, Multidisciplinary
Agriculture, Soil Science
Agronomy
Entomology
Fisheries
Food Science & Technology
Forestry
Horticulture
Marine & Freshwater Biology
Ornithology
Plant Sciences
Veterinary Sciences
Water Resources

Social Sciences
AGRICULTURAL ECONOMICS & POLICY
Architecture
Business
Business, Finance
COMMUNICATION
Criminology & Penology
ECONOMICS
Health Care Sciences & Services
Health Policy & Services
Information Science & Library Science
Law
Management
Operations Research & Management Science
Planning & Development
Planning And Development
Psychology, Applied
Psychology, Experimental
Public Administration
SOCIAL SCIENCES, BIOMEDICAL
SOCIAL SCIENCES, MATHEMATICAL METHODS
Social Sciences, Interdisciplinary
Social Work
Statistics & Probability
Urban Studies

Humanities
ANTHROPOLOGY
Archaeology
Area Studies
Art
Behavioral Sciences
Education & Educational Research
Education, Scientific Disciplines
Education, Special
Environmental Studies
Ethics
Family Studies
History & Philosophy Of Science
History Of Social Sciences
History
Humanities, Multidisciplinary
International Relations
Language & Linguistics Theory
Literary Reviews
Literature, Romance
Music
Paleontology
Philosophy
Political Science
Psychology
Psychology, Developmental
Psychology, Educational
Psychology, Psychoanalysis
Psychology, Social
Sociology
Women's Studies

Engineering
Agricultural Engineering
Automation & Control Systems
Computer Science, Artificial Intelligence
Computer Science, Cybernetics
Computer Science, Information Systems
Computer Science, Interdisciplinary Applications
Computer Science, Software Engineering
Computer Science, Theory & Methods
Construction & Building Technology
ENGINEERING, BIOMEDICAL
ENGINEERING, OCEAN
Energy & Fuels
Engineering, Aerospace
Engineering, Chemical
Engineering, Civil
Engineering, Electrical & Electronic
Engineering, Environmental
Engineering, Geological
Engineering, Industrial
Engineering, Manufacturing
Engineering, Mechanical
Engineering, Multidisciplinary
Engineering, Petroleum
Ergonomics
Imaging Science & Photographic Technology
Instruments & Instrumentation
MINING & MINERAL PROCESSING
Mechanics
Metallurgy & Metallurgical Engineering
NEUROIMAGING
Nanoscience & Nanotechnology
Operations Research & Management Science
Robotics
Telecommunications
Thermodynamics
Transportation Science & Technology
Transportation
Water Resources

Other
Multidisciplinary Sciences
Psychology, Multidisciplinary

Author: Gonzalo Ordonez M., 2007.

APPENDIX E Cont'd

Equivalences: ISI-Silac05.

(Some fields are classified more than once)

Biologia
Anatomy & Morphology
Biology
Biophysics
Biotechnology & Applied Microbiology
Cell Biology
Developmental Biology
ENGINEERING, BIOMEDICAL
Evolutionary Biology
Genetics & Heredity
Microbiology
Microscopy
Neurosciences
Nutrition & Dietetics
Parasitology
Physiology
Reproductive Biology
Virology

Biologia Aplicada, Ecologia
Agricultural Engineering
Agriculture, Dairy & Animal Science
Agriculture, Multidisciplinary
Agriculture, Soil Science
Agronomy
Biodiversity Conservation
Botanics
Ecology
Entomology
Fisheries
Food Science & Technology
Forestry
Horticulture
Mycology
Ornithology
Plant Sciences
Zoology

Ciencias Medicas
Allergy
Andrology
Anesthesiology
CARDIAC & CARDIOVASCULAR SYSTEM
Cardiac & Cardiovascular Systems
Chemistry, Medicinal
Clinical Neurology
Critical Care Medicine
Dentistry, Oral Surgery & Medicine
Dermatology
Emergency Medicine
Endocrinology & Metabolism
Gastroenterology & Hepatology
Geriatrics & Gerontology
Gerontology
Hematology
Immunology
Infectious Diseases
Integrative & Complementary Medicine
Medical Informatics
Medical Laboratory Technology
Medicine, General & Internal
Medicine, Legal
Medicine, Research & Experimental
Neuroimaging
Nursing
Obstetrics & Gynecology
Oncology
Ophthalmology
Orthopedics
Otorhinolaryngology
Pathology
Pediatrics
Peripheral Vascular Disease
Psychiatry
Psychology, Biological
Psychology, Clinical
Public, Environmental & Occupational Health
Radiology, Nuclear Medicine & Medical Imaging
Rehabilitation
Respiratory System
Rheumatology
Sport Sciences
Substance Abuse
Surgery
Toxicology
Transplantation
Tropical Medicine
Urology & Nephrology
Veterinary Sciences

Author: Gonzalo Ordonez M., 2007.

APPENDIX E Cont'd :

Equivalences: ISI-Silac05 Contd.

Ciencias de la Tierra y el Universo
Astronomy & Astrophysics
Environmental Sciences
Geochemistry & Geophysics
Geography
Geography, Physical
Geology
Geosciences, Multidisciplinary
Limnology
Marine & Freshwater Biology
Meteorology & Atmospheric Sciences
Mineralogy
Mining & Mineral Processing
Oceanography
Paleontology
Remote Sensing
Water Resources

Fisica
Acoustics
Instruments & Instrumentation
Mechanics
Nanoscience & Nanotechnology
Optics
Physics
Physics, Applied
Physics, Atomic, Molecular & Chemical
Physics, Condensed Matter
Physics, Fluids & Plasmas
Physics, Mathematical
Physics, Multidisciplinary
Physics, Nuclear
Physics, Particles & Fields
Spectroscopy
Thermodynamics

Quimica
Biochemical Research Methods
Biochemistry & Molecular Biology
Chemistry
Chemistry, Analytical
Chemistry, Applied
Chemistry, Inorganic & Nuclear
Chemistry, Multidisciplinary
Chemistry, Organic
Chemistry, Physical
Crystallography
Electrochemistry
Materials Science, Biomaterials
Materials Science, Ceramics
Materials Science, Characterization & Testing
Materials Science, Coatings & Films
Materials Science, Composites
Materials Science, Multidisciplinary
Materials Science, Paper & Wood
Materials Science, Textiles
Pharmacology & Pharmacy
Polymer Science

Ingenieria y Tecnologia
Automation & Control Systems
Computer Science
Computer Science, Artificial Intelligence
Computer Science, Cybernetics
Computer Science, Information Systems
Computer Science, Interdisciplinary Applications
Computer Science, Software Engineering
Computer Science, Theory & Methods
Construction & Building Technology
Energy & Fuels
Engineering, Aerospace
Engineering, Chemical
Engineering, Civil
Engineering, Electrical & Electronic
Engineering, Environmental
Engineering, Geological
Engineering, Industrial
Engineering, Manufacturing
Engineering, Marine
Engineering, Mechanical
Engineering, Multidisciplinary
Engineering, Ocean
Engineering, Petroleum
Engineering, Sanitation
Engineering, Transportation
Ergonomics
Imaging Science & Photographic Technology
Metallurgy & Metallurgical Engineering
Nuclear Science & Technology
Operations Research & Management Science
Robotics
Telecommunications
Transportation Science & Technology
Transportation

Matematicas
Mathematics
Mathematics, Applied
Mathematics, Interdisciplinary Applications
SOCIAL SCIENCES, MATHEMATICAL METHODS
Statistics & Probability

Author: Gonzalo Ordonez M., 2007.

APPENDIX E Cont'd :

Equivalences: ISI-Silac05 Contd.

Ciencias Sociales y Humanidades
AGRICULTURAL ECONOMICS & POLICY
ANTHROPOLOGY
Archaeology
Architecture
Area Studies
Art
Behavioral Sciences
Business
Business, Finance
COMMUNICATION
Criminology & Penology
ECONOMICS
Education & Educational Research
Education, Scientific Disciplines
Education, Special
Environmental Studies
Ethics
Family Studies
Health Care Sciences & Services
Health Policy & Services
History & Philosophy Of Science
History Of Social Sciences
History
Humanities, Multidisciplinary
Information Science & Library Science
International Relations
Language & Linguistics Theory
Law
Literary Reviews
Literature, Romance
Management
Medical Ethics
Music
Philosophy
Planning & Development
Planning And Development
Political Science
Psychology
Psychology, Applied
Psychology, Developmental
Psychology, Educational
Psychology, Experimental
Psychology, Psychoanalysis
Psychology, Social
Public Administration
Social Sciences, Biomedical
Social Sciences, Interdisciplinary
Social Work
Sociology
Urban Studies
Women's Studies

Otra
Multidisciplinary Sciences
Psychology, Multidisciplinary

Author: Gonzalo Ordonez M., 2007.

APPENDIX F

INTERVIEW PROTOCOL

Introduction:

From the research policy perspective, research collaboration represents a useful strategy for increasing creativity, research productivity, output quality, and innovative capacity. It allows researchers to access valuable resources (material and cognitive ones) sometimes unavailable otherwise, create human capital for science and technology, and strengthen their research streams. In the case of developing countries, international research collaboration can, in addition, contribute to narrow the gaps between scientific and technological communities, allowing them to increase their competitiveness and reduce the social and environmental effects that results from their condition of underdevelopment.

A non-negligible number of empirical works argue that research collaboration can also have negative effects on the same indicators of relevance, however. High transaction costs and institutional constraints may affect research performance reducing its potential to produce promising results.

The purpose of this research is to assess the impacts of International Research Collaboration on the performance of Research Teams in Colombia. In this framework, research teams are seen as the unit of analysis and policy focus as a result of the strategic role they play as the building blocks of the Colombian Science and Technology System.

For this purpose, this research relies on a multimethods approach. Whereas the evaluation of some of these effects can efficiently be done using quantitative techniques, their validation for interpretation purposes and the assessment of causal relationships on the ground requires, in addition, the use of qualitative methods. In this case, the administration of a number of interviews to randomly selected research teams members include the list of questions annexed.

Questions

1. Name:
2. Occupation:
3. Sector:
4. City:
5. Years associated with the team:
6. Main discipline of expertise:
7. Main team's discipline of expertise:
8. Highest degree and granting university:
9. Have you participated in collaborative research with foreign partners in the last two years? in the last 4 years?
10. If no, why?

11. If yes, for each collaborative activity specify (no need to provide information about the partner's identification):
    - Type of Collaborative Activity:
    - Your main motivation for participating:
    - Your main role in this activity:
    - Country of origin of partner (the answer can also refer to North or South origin):
    - Duration of collaboration:
    - How important was (is) it for your work? For the team? (very important, somewhat important, not important) Explain:
    - Main results:
    - Main benefits:
    - Main costs:
    - Main enablers:
    - Main barriers:
    - What would you do differently?
    - Who initiated/invited the collaborative activity?
    - Would you collaborate with the same partner in the future? (yes/no)
    - Did the activity affect 'teamworking'? Explain
    - Did it affect team performance? (yes, a lot; somewhat; not really) Explain.
    - Team productivity?
    - Team output quality?
    - Team visibility?
    - How many people participated in the collaborative activity?
    - Did the collaborative activity involve students? How many? What were their roles?
    - What would you need to improve your collaboration experience in the future?

APPENDIX G

ANALYSIS OF THE POISSON DISTRIBUTION OF TEAM OUTPUT AND SELECTION OF THE MODEL

As Figure 4 reveals, our outcome variable, team output, has a frequency distribution highly skewed to the left, showing many teams reporting zero or small number of products during the period observed, and very few teams reporting large number of bibliographical products.

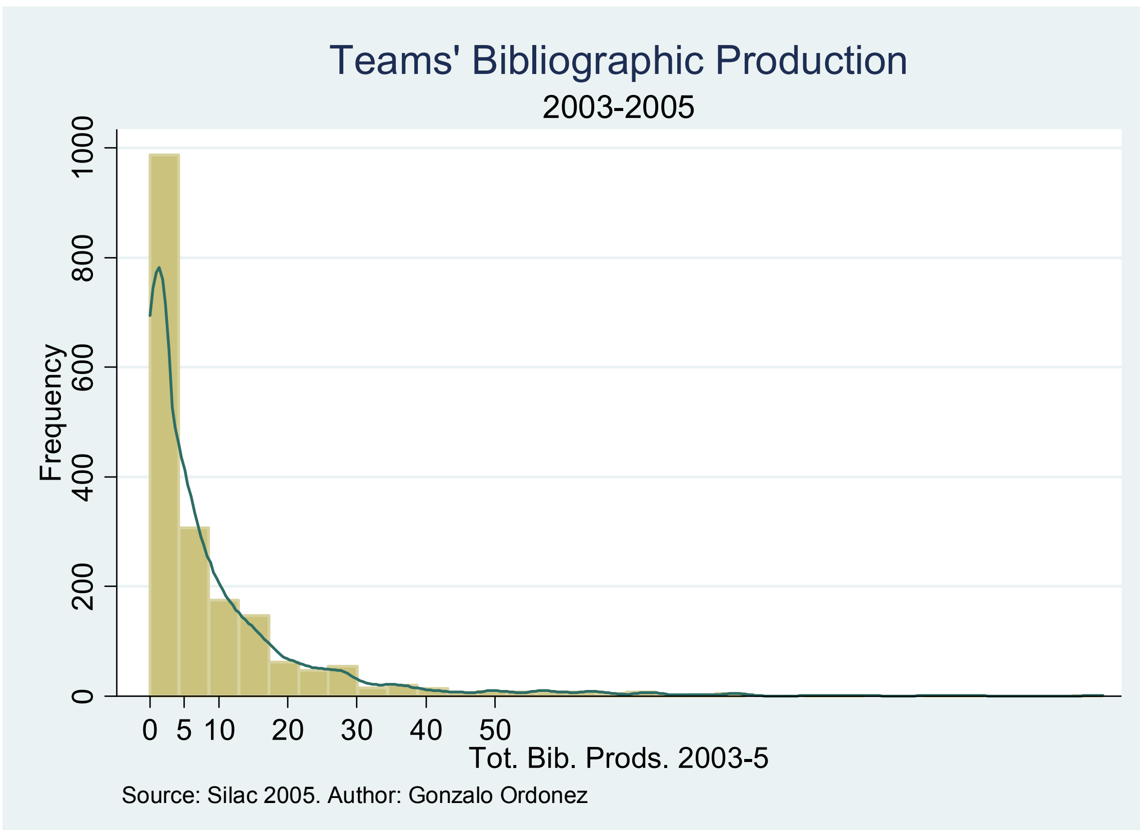

Figure 4: Frequency Distribution of Team Bibliographic Production

The boxplot for our outcome variable in Figure 5 shows positive skew of the distribution. The median (the line) is pulled to the low end of the box, and the 95th percentile is stretched out away from the box. If the number of bibliographic products had a normal distribution, the line would have been in the middle of the box (the 25th and 75th percentiles) and the ends of the whiskers (5th and 95th percentile) would have been equidistant from the box.

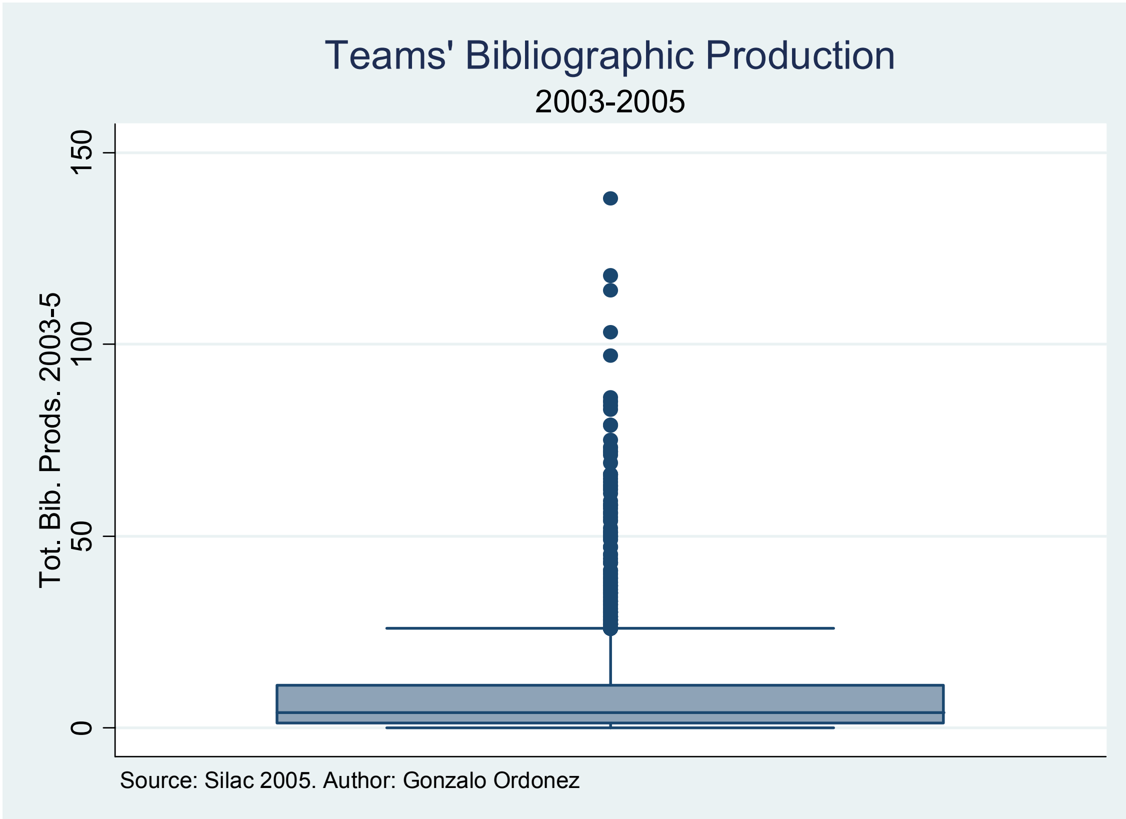

Figure 5: Box Plot of Team Bibliographic Production

To better understand regression models for count variables, a brief analysis of the univariate Poisson distribution is helpful.

Let y be a random variable indicating the number of bibliographic products made by the teams observed. If y has a Poisson distribution, then

$$Pr(y|\mu) = ((e\text{^}- \mu)\ \mu\ \text{^}y) / y! \qquad for\ y = 0,1,2\ldots$$

where $\mu > 0$ is the sole parameter defining the distribution. To get a sense of this distribution we simply compare the observed distribution to a Poisson distribution that has the same mean. To do that, we first estimate the Poisson regression model for team productivity and no independent variables in order to fit a univariate Poisson distribution with a mean equal to that of our outcome variable. That is, we estimate the model:

$$\mu = exp\ (\beta o)$$

And we get:

```
. poisson  totbibprod05

Iteration 0:   log likelihood = -16360.276
Iteration 1:   log likelihood = -16360.276

Poisson regression                                Number of obs   =       1889
                                                  LR chi2(0)      =       0.00
                                                  Prob > chi2     =          .
Log likelihood = -16360.276                       Pseudo R2       =     0.0000

------------------------------------------------------------------------------
totbibprod05 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   2.181634   .0077294   282.25   0.000     2.166485    2.196783
------------------------------------------------------------------------------
```

Since βo = 2.1815, μ = exp (2.1815) = 8.8608, which is the same as the estimated mean of totbibprod05 reported earlier.

As shown on Figure 6, the fitted Poisson distribution (represented by Δ's) under-predicts 0s, 1s, 2s, 3s, and 4s (Observed > Predicted) and over-predicts counts 5 and above (Observed < Predicted).

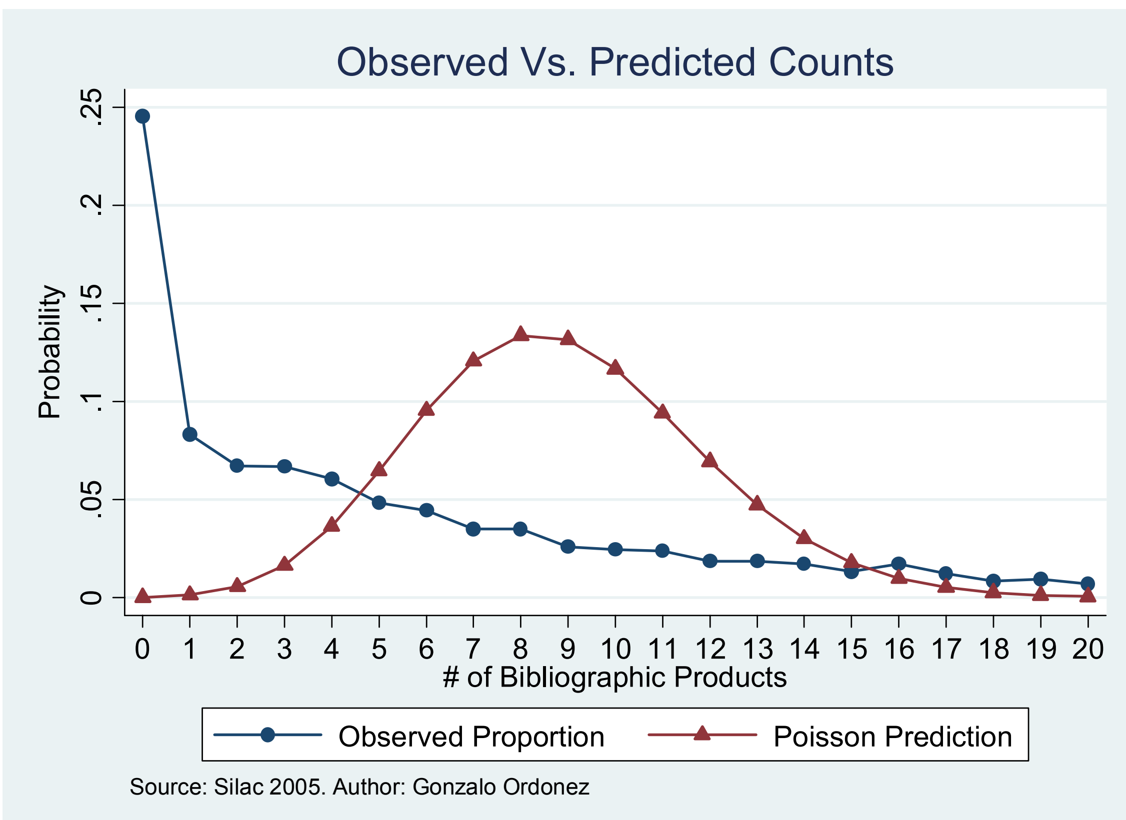

Figure 6: Comparison of Observed Counts Vs. Poisson Predictions

Both over- and under-prediction is characteristic of fitting a count model that does not take into account heterogeneity among sample members in their production rate μ. Since assuming that all teams have the same rate of bibliographic production is unrealistic, the next step is to incorporate observed heterogeneity (i.e., observed differences among sample members) in μ based on team characteristics and other independent variables.

In the framework of this dissertation, four alternative models were compared to find the one that fits best the data. These are: the Poisson Regression Model (PRM), the Negative Binomial Regression Model (NBRM), the Zero-Inflated Poisson (ZIP) Model, and the Zero-Inflated Negative Binomial (ZINB) Model.

The Poisson Regression Model (PRM) allows each team to have a different value of μ. It assumes that the observed count for team *i* is drawn from a Poisson distribution with mean μi, where μi is estimated from observed characteristics. This is:

$$\mu i = E(yi \mid xi) = exp\ (xi\beta)$$

Taking the exponential of $x\beta$ forces μ to be positive, which is necessary since counts can only be 0 or positive.

If teams that differ in their rates of production are combined, the univariate distribution of bibliographic products will be overdispersed, that is, with a variance greater than the mean. As we saw, heterogeneity among teams in their rate of production could be due to factors such as team size, team age, composition, dynamism, discipline, institution of affiliation, sector, location, and, of course, our variable of interest, collaboration status.

Figure 7 below shows that compared to the univariate Poisson model the multivariate model does improve prediction but is far from satisfying. According to the graph, even though many of the independent variables have significant effects on the number of bibliographic products done, there is a modest improvement in the predictions made by the PRM considered here over the univariate Poisson distribution, with a bit more 0s, more 1s to 6s, and fewer counts greater than 7.

Although the Poisson Regression Model (PRM) accounts for observed heterogeneity by specifying the rate μi as a function of observed xk's, in practice, L&F note, it rarely fits as it underestimates the amount of dispersion in the outcome. According to the authors, this failure is addressed by the Negative Binomial Regression Model (NBRM), which adds a parameter α that reflects unobserved heterogeneity among observations[24]. Put by the authors, if the assumptions of the NBRM are correct, the expected rate for a given level of the independent variables will be the same in the PRM and the NBRM as both have the same mean structure. However, the authors claim, even if the model includes the correct variables, and as a result of overdispersion, estimates from the PRM are inefficient and "the standard errors in the PRM will be biased downward, resulting in spuriously large z-values and spuriously small p-values."

[24] The demonstration can be found in L&F 2001, page. 243-244.

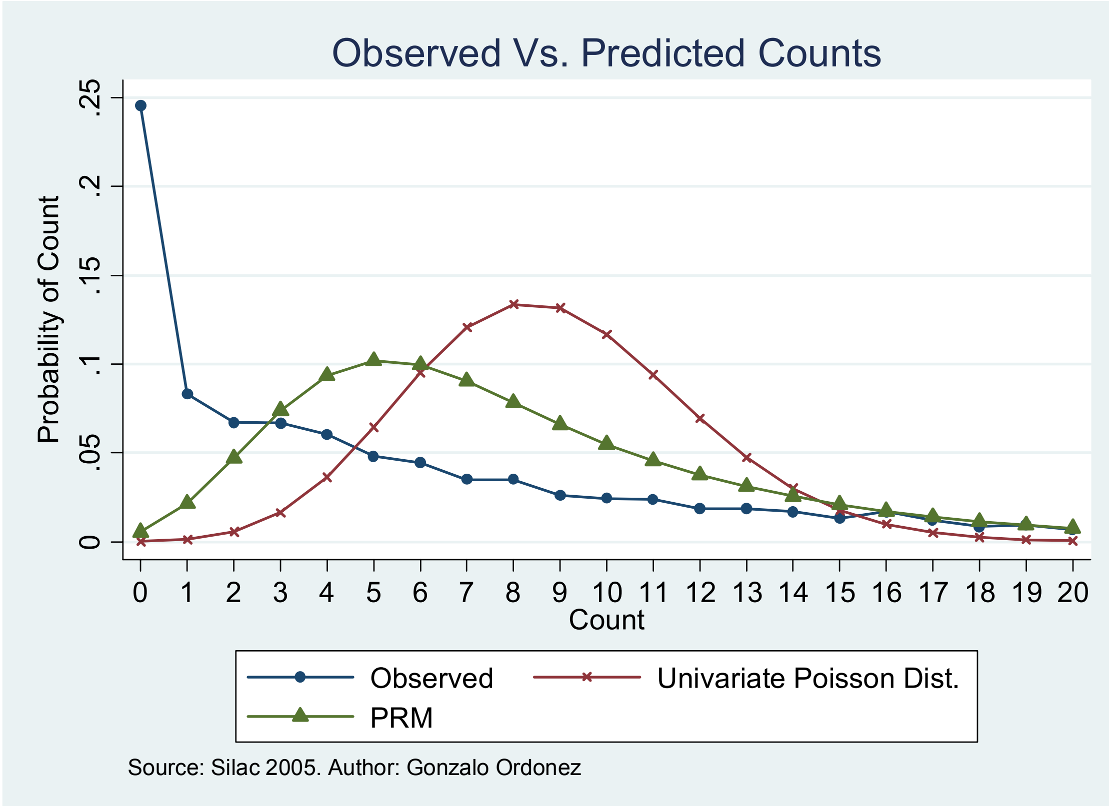

Figure 7: Comparison of Observed Vs. Poisson Regression Model Predictions

For these reasons, it is important to test for overdispersion, and this is usually done by testing the null hypothesis that $\alpha = 0$, as the NBRM reduces to the PRM when $\alpha = 0$. This is done by a LR test reported by Stata after the estimates of the parameters. The test statistic chibar2(01) is computed by the formula

$$
\begin{aligned}
G^2 &= 2(\ln L\ NBRM - \ln L\ PRM) \\
&= 2(-5688.8548 - -12950.263) \\
&= 14522.82\ \Xi\ 15000
\end{aligned}
$$

This results are very significant and provide strong evidence of overdispersion ($G^2 = 15000$, $p<.01$). Therefore, the negative binomial regression model is preferred to the Poisson regression model.

As shown in the outputs from the *prvalue* below, it seems that the NBRM is better than the PRM as it improves upon the underprediction of zeroes in the latter model by increasing the conditional variance without changing the conditional mean.

```
. quietly poisson totbibprod05 IRC05 Core03 age03 totphds03 totprojects03 agroscs
medscs social human engi othscs bussector govsector othsector medinst smallinst
smallcity medcity, nolog

        . prvalue, max(20)

        poisson: Predictions for totbibprod05

        Confidence intervals by delta method

                                      95% Conf. Interval
          Rate:               7.511   [ 7.3813,    7.6408]
          Pr(y=0|x):         0.0005   [ 0.0005,    0.0006]
          Pr(y=1|x):         0.0041   [ 0.0036,    0.0046]
          Pr(y=2|x):         0.0154   [ 0.0140,    0.0169]
          Pr(y=3|x):         0.0386   [ 0.0356,    0.0416]
          Pr(y=4|x):         0.0725   [ 0.0681,    0.0769]
          Pr(y=5|x):         0.1090   [ 0.1042,    0.1137]
          Pr(y=6|x):         0.1364   [ 0.1329,    0.1400]
          Pr(y=7|x):         0.1464   [ 0.1451,    0.1477]
          Pr(y=8|x):         0.1374   [ 0.1363,    0.1386]
          Pr(y=9|x):         0.1147   [ 0.1117,    0.1176]
          Pr(y=10|x):        0.0861   [ 0.0824,    0.0898]
          Pr(y=11|x):        0.0588   [ 0.0553,    0.0624]
          Pr(y=12|x):        0.0368   [ 0.0340,    0.0397]
          Pr(y=13|x):        0.0213   [ 0.0193,    0.0233]
          Pr(y=14|x):        0.0114   [ 0.0101,    0.0127]
          Pr(y=15|x):        0.0057   [ 0.0050,    0.0065]
          Pr(y=16|x):        0.0027   [ 0.0023,    0.0031]
          Pr(y=17|x):        0.0012   [ 0.0010,    0.0014]
          Pr(y=18|x):        0.0005   [ 0.0004,    0.0006]
          Pr(y=19|x):        0.0002   [ 0.0002,    0.0002]
          Pr(y=20|x):        0.0001   [ 0.0001,    0.0001]

        IRC05       Core03        age03    totphds03  totprojec~03      agroscs       medscs       social
x= .38962414    7.2583377    6.9047115    1.4552673     5.7517205   .06087877   .12652197   .16146109

        human         engi       othscs    bussector     govsector   othsector      medinst   smallinst
x= .24192695     .1339333   .04235045   .03388036     .03864479   .02064584   .37850715   .16781366

    smallcity      medcity
x= .01746956    .21916358

        . quietly nbreg totbibprod05 IRC05 Core03 age03 totphds03 totprojects03
agroscs medscs social human engi othscs bussector govsector othsector medinst
smallinst smallcity medcity, nolog

        . prvalue, max(20)

        nbreg: Predictions for totbibprod05

        Confidence intervals by delta method
                                      95% Conf. Interval
          Rate:              7.2085   [ 6.7792,    7.6378]
          Pr(y=0|x):         0.2040   [ 0.1969,    0.2110]
          Pr(y=1|x):         0.1183   [ 0.1148,    0.1219]
          Pr(y=2|x):         0.0887   [ 0.0865,    0.0909]
          Pr(y=3|x):         0.0716   [ 0.0701,    0.0730]
          Pr(y=4|x):         0.0597   [ 0.0588,    0.0606]
          Pr(y=5|x):         0.0509   [ 0.0503,    0.0514]
          Pr(y=6|x):         0.0439   [ 0.0436,    0.0441]
          Pr(y=7|x):         0.0382   [ 0.0382,    0.0383]
          Pr(y=8|x):         0.0335   [ 0.0334,    0.0337]
          Pr(y=9|x):         0.0296   [ 0.0293,    0.0298]
          Pr(y=10|x):        0.0262   [ 0.0258,    0.0265]
          Pr(y=11|x):        0.0233   [ 0.0228,    0.0237]
          Pr(y=12|x):        0.0207   [ 0.0203,    0.0212]
          Pr(y=13|x):        0.0185   [ 0.0180,    0.0190]
          Pr(y=14|x):        0.0166   [ 0.0160,    0.0171]
          Pr(y=15|x):        0.0149   [ 0.0143,    0.0154]
          Pr(y=16|x):        0.0134   [ 0.0128,    0.0139]
          Pr(y=17|x):        0.0120   [ 0.0115,    0.0126]
          Pr(y=18|x):        0.0108   [ 0.0103,    0.0114]
```

```
                Pr(y=19|x):          0.0098   [ 0.0092,    0.0103]
                Pr(y=20|x):          0.0088   [ 0.0083,    0.0094]

        IRC05       Core03        age03    totphds03  totprojec~03       agroscs       medscs       social
x= .38962414    7.2583377    6.9047115    1.4552673    5.7517205     .06087877    .12652197    .16146109

        human         engi       othscs    bussector    govsector   othsector      medinst    smallinst
x= .24192695     .1339333    .04235045    .03388036    .03864479    .02064584    .37850715    .16781366

    smallcity      medcity
x= .01746956    .21916358
```

The predicted rate is nearly identical for both models (7.511 versus 7.209) which shows that even with overdispersion the estimates from the PRM are consistent. But an exam of the predicted probabilities reveals substantial differences: Pr(y=0|x)= 0.0005 in the PRM, versus Pr(y=0|x)= 0.2040 in the NBRM. Also, as an illustration of the large dispersion in the NBRM compared to the PRM, the probabilities in the NBRM are higher than in the PRM for higher counts (e.g. Pr(y=15|x)= 0.0057 in the PRM versus, Pr(y=15|x)= 0.0149 in the NBRM).

Finally, as Figure 8 shows, the probability of having zero bibliographic products is higher in the NBRM than in the PRM. This is evident by plotting the probability of 0s as values of an independent variable change. In this case, this is computed when each variable except the number of PhDs is held at its mean. For both models, the probability of a zero decreases as the number of PhDs increases (hardly seen in the PRM due to scale), but the proportion of predicted zeroes is remarkably higher for the NBRM. Since both models have the same expected number of products, the higher proportion of predicted zeroes for the NBRM is offset by the higher proportion of larger counts that are also predicted by this model.

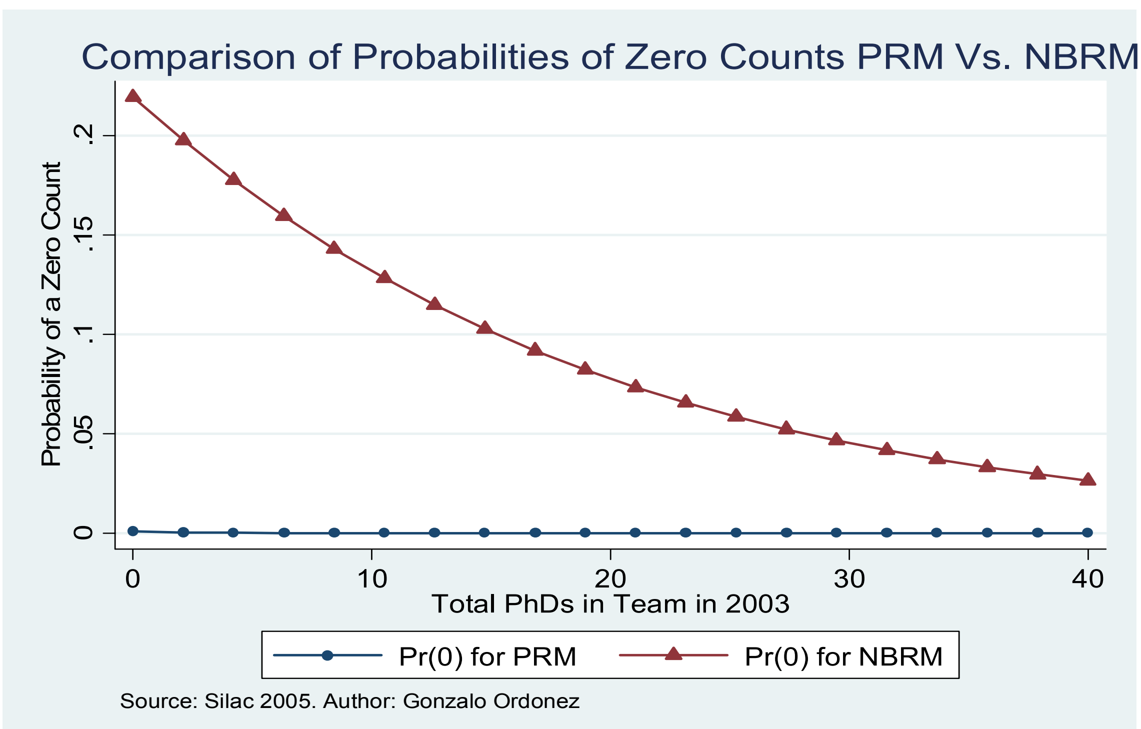

Figure 8: Prediction of Zero Counts by PRM and NBRM Compared

However, although the NBRM improves upon the underprediction of zeroes in the PRM by controlling for dispersion, it may still fail to satisfactorily account for excess of zeroes. In fact, both the PRM and the NBRM assume that every single team has a positive probability of producing any given number of bibliographic products. According to this model, the probability of producing research outputs differs across teams based to their characteristics, but all teams have some probability of producing bibliographic products.

In practice, this assumption may not be a valid one as not all research teams are potential producers of bibliographic products. In fact, teams affiliated to industries, or working under the aegis of a decision-making or policy unit may not be allowed to produce or report bibliographic products.

One thing is a team that did not produce or report bibliographic products during the period observed because it lacked the resources or the motivation necessary to do so, and another thing is a team that did not produce or report bibliographic products because it was not allowed to do so.

These two types of teams will look identical in the dependent variable: both report zero bibliographic products. But they in fact have arrived at the same outcome through two different processes. The first team could have produced bibliographic products during the period observed (had it been more active or got enough financial resources to perform their research), but did not[25]. The second team was certain to report zero products because it was prevented to do so.

Thus, the number of zeroes may be inflated and the number of 'unproductive' teams cannot be explained in the same manner as the number of teams that produced more than zero bibliographical products. Some teams reported zero products for the same reasons other teams reported one, two, or three products (resources, motivation) and while some teams did not report bibliographic products for a different set of reasons.

A standard Poisson Regression Model would not distinguish between the two processes causing an excessive number of zeroes, but a zero-inflated count model responds to this issue and allows for this possibility by increasing the conditional variance and the probability of zero counts.

According to Long and Freese 2001, one of the characteristics of the Zero-Inflated Count Models is that they assume that there are two latent or unobserved groups in which all teams fall depending on their inherent propensity to produce bibliographic products. Thus, there is an "Always -0 Group" and a "Not Always -0 Group." A team in the former

---

[25] The lack of time to produce bibliographic products in the case of the teams created one or two years before the time the data was gathered (September 2005), or the potential problems derived from the software used to capture the information, may be some of the other factors explaining the excess of zeroes. However, these factors are not important as, first, only 64 teams were created between 2003 and 2004 (3% of the total), of which only 25 (1.32%) reported zero bibliographic products. Second, technical problems were substantially reduced by the fact that Colciencias allowed the teams to revise the data they submitted in June 2005 by sending each of them a preliminary report, so that the teams could have four months to make changes and correct errors before the September deadline.

group, called Group A to simplify, has an outcome of 0 with a probability of 1. In contrast, a team in the latter group, called Group ~A, might have a zero count, but there is a nonzero probability that it has a positive count. To understand how this works, we need first to model membership into the latent groups; then we model counts for those in Group ~A; and finally, we need to compare observed probabilities as a mixture of the probabilities for the two groups.

Thus, following Long and Freese, let A = 1 if a team is in Group A, else A = 0. Group membership is a binary outcome that can be modeled using the logit or probit model

$$\Psi i = Pr(Ai = 1 / zi) = F(zi\gamma)$$

where Ψi is the probability of being in Group A for team i. The z-variables are referred to as inflation variables since they serve to inflate the number of 0s. If we had an observed variable indicating group membership, this would be a standard logit or probit model. But, since group membership is a latent variable, we do not know whether a team is in Group A or Group ~A.

On the other hand, the probability of each count (including zeroes) among those who are not always zero (Group ~A) is determined by either a PRM or a NBRM.

Hence, for the Zero-Inflated Poisson (ZIP) Model, we have

$$Pr(yi \mid xi, Ai = 0) = [(e\text{^-}\mu i) \cdot (\mu i\text{^ } yi)] / yi!$$

or, for the Zero Inflated Negative Binomial (ZINB) Model,

$$\begin{aligned} Pr(yi \mid xi, Ai = 0) = \{[&\Gamma(yi + \alpha\text{^-}1)] / [yi\, !\, \Gamma(\alpha\text{^-}1)] \cdot \\ &\{[(\alpha\text{^-}1) / (\alpha\text{^-}1) + \mu i]\}\text{^ } \alpha\text{^-}1 \cdot \\ &\{[\mu i / (\alpha\text{^-}1) + \mu i]\}\text{^ } yi \end{aligned}$$

Notice that in the equations for ZIP and ZINB, we are conditioning both on the xk's and on $A = 0$. Also, note that the xk's are not necessarily the same as the inflation variables zk explained above (although the two sets of variables can be the same).

In both equations, $\mu i = exp(xi\beta)$. If we knew which observations were in Group ~A, these equations would define the PRM and the NBRM. But, here the equations only apply to those observations in Group ~A, and we do not have an observed variable indicating group membership.

Finally, we need to combine Groups A and ~A according to their proportions in the population to determine the overall rate.

The proportion in each group is defined by

$$Pr(Ai = 1) = \Psi i$$
$$Pr(Ai = 0) = 1 - \Psi i$$

and the probabilities of a zero within each group are

$$Pr(yi = 0 \mid Ai = 1, xi, zi) = 1 \text{ by definition of the A Group}$$
$$Pr(yi = 0 \mid Ai = 0, xi, zi) = \text{outcome of PRM or NBRM}$$

Then, the overall probability of a 0 count is

$$\begin{aligned} Pr(yi = 0 \mid xi, zi) &= [\Psi i \cdot 1] + [(1 - \Psi i) \cdot Pr(yi = 0 \mid xi, Ai = 0)] \\ &= \Psi i + [(1 - \Psi i) \cdot Pr(yi = 0 \mid xi, Ai = 0)] \end{aligned}$$

Expected counts are computed in a similar fashion:

$$\begin{aligned} E(y \mid x, z) &= [0 \cdot \Psi] + [\mu \cdot (1 - \Psi)] \\ &= \mu(1 - \Psi) \end{aligned}$$

Since $0 \leq \Psi \leq 1$, the expected value will be smaller than μ, which shows that the mean structure in zero-inflated models differs from that in the PRM or NBRM.

In sum, we have seen so far that, based on the characteristics of our outcome variable we need to apply what is commonly called 'count' models for explaining the productive capacity of research teams. To do that, we first used a Poisson Regression Model to take into account the Poisson (skewed) distribution of the dependent variable, team production (totbibprod05). Then, we used a Negative Binomial Regression Model to account for the (over)dispersion found (the variance exceeds the mean by a great deal) and to control for the resulting underestimation of zero counts. Finally, we used a Zero-Inflated Negative Binomial Regression Model to account for the fact that some research teams may actually have zero probabilities of producing or reporting bibliographic products for the reasons explained earlier, while allowing the possibility that teams did not report products during the period observed due to chance or other unobserved factors. The question that begs an answer at this point is therefore, how does one know what model fits best based on the characteristics of the population and the outcome variable studied. Intuitively, and for the reasons considered, one would prefer the ZINB regression model over the other two models used as it seems to consider a more realistic situation. However, what does the empirical evidence says in support (o rejection) of that choice?

A comparison of models

The following figure shows a comparison of the predicted probabilities among the models used, and how they differ from the observed probabilities. That is, this plots the difference between the observed probabilities and the mean prediction for each count using each model (see Figure 9).

As the figure shows, points above the 0 on the y-axis indicate more observed counts than predicted; those below 0 indicate more predicted counts than observed. The figure shows that both the PRM and to some extent the NBRM have problem predicting the average number of 0s. Also, the PRM predicts too many 1s and 2s and too few larger counts. The NBRM does relatively well except that it predicts too few 1s and 2s. The ZIP model predicts too many 1s, 2s, 3s, and 4s and too few larger counts. The ZINB model fits almost perfectly among all counts. For this reason, the ZINB is preferred over the other models.

A more formal testing of model fit can be done with an LR test of overdispersion and a Vuong test to compare two models.

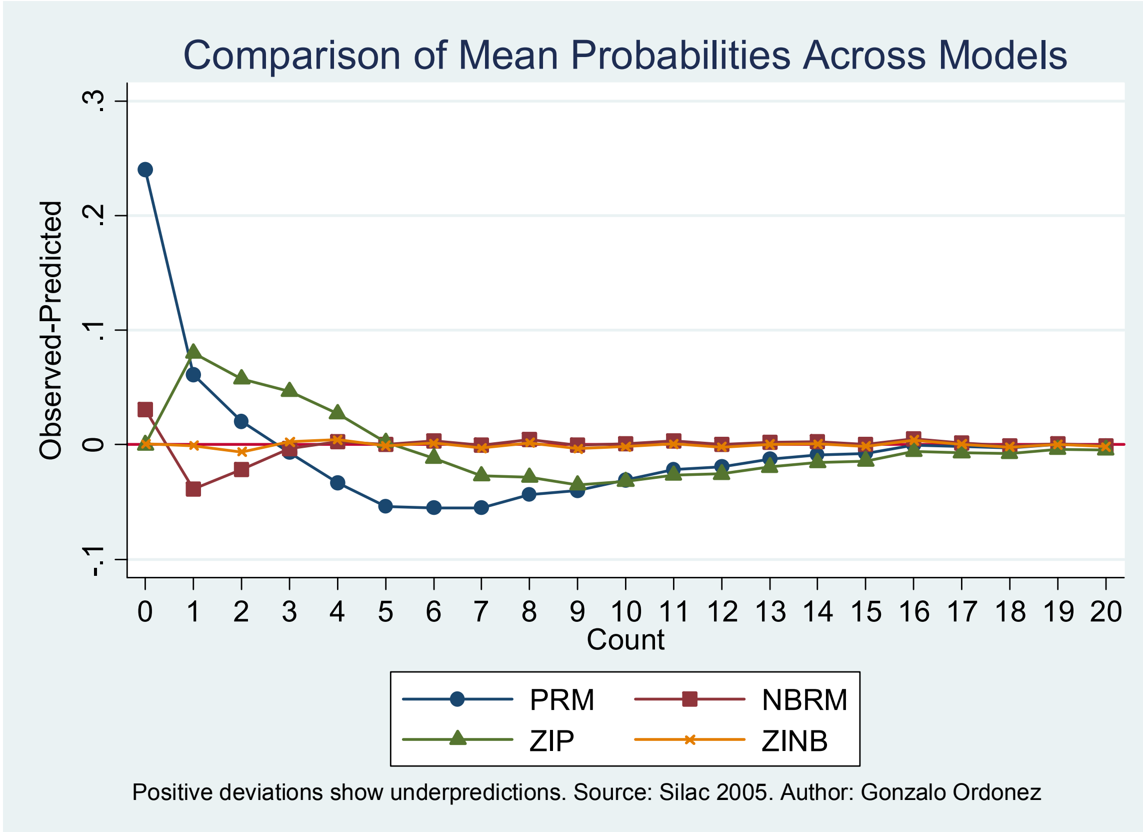

Figure 9: PRM, NBRM, ZIP and ZINB Compared

We showed earlier that a test for overdispersion using the null hypothesis that α = 0 (i.e. the NBRM reduces to the PRM) yielded strong evidence for preferring the NBRM over the PRM: *Likelihood-ratio test of alpha=0: chibar2(01) = 1.5e+04 Prob>=chibar2 = 0.000*

The comparison between ZIP and ZINB can be done applying the same LR test (G^2) discussed earlier as both models are nested. To do that, we first compute the difference between the two log likelihoods resulting from the estimation using each model. Then, we compute the p-value for a chi-squared test with one degree of freedom taking into account that α cannot be negative as discussed earlier (i.e. we need to divide by 2). Finally, we assign the estimated value of ln α to a scalar. If this value is very close to 0, we conclude that the p-value is 1.

Using Stata we get these results and we conclude the following:

```
. quietly zip totbibprod05 IRC05 Core03 age03 totphds03 totprojects03 agroscs
medscs social human engi othscs bussector govsector othsector medinst smallinst
smallcity medcity, inf(IRC05 Core03 age03 totphds03 totprojects03 agroscs medscs
social human engi othscs bussector govsector othsector medinst smallinst smallcity
medcity)

. scalar llzip = e(ll)

. quietly zinb totbibprod05 IRC05 Core03 age03 totphds03 totprojects03 agroscs
medscs social human engi othscs bussector govsector othsector medinst smallinst
smallcity medcity, inf(IRC05 Core03 age03 totphds03 totprojects03 agroscs medscs
```

```
social human engi othscs bussector govsector othsector medinst smallinst smallcity
medcity)

        . scalar llzinb = e(ll)

        . scalar lr = -2*(llzip-llzinb)

        . scalar pvalue = chiprob(1,lr)/2

        . scalar lnalpha = -.0203014

        . if (lnalpha <-20) scalar pvalue= 1

        . di as text "Likelihood ratio test comparing ZIP to ZINB: " as res %8.3f
lr as text " Prob>=" as res %5.3f pvalue

        Likelihood ratio test comparing ZIP to ZINB: 8954.753 Prob>=0.000
```

The ZINB model significantly improves the fit over the ZIP model.

Because, on the one hand, as Long and Freese note, the PRM and the ZIP are not nested, and the NBRM and the ZINB are not nested either, the Vuong test helps in deciding which model is best among each set of models.

Since *V* has an asymptotic normal distribution, if *V*> 1.96 the first model is favored, and if *V*< -1.96, the second model is favored:

V= (√N m-bar) / sm, where *m-bar* is the mean, and *sm* is the standard deviation of *mi*, which equals to *ln [Pr1 (yi | xi) / Pr2 (yi | xi)]*, and *Pr# (yi |xi)* is the predicted probability of observing *y* in each model.

While in Stata for ZIP the Vuong test computes the Vuong statistic comparing the ZIP model to the PRM; for ZINB it compares ZINB to NBRM:

```
. quietly zip totbibprod05 IRC05 Core03 age03 totphds03 totprojects03 agroscs
medscs social human engi othscs bussector govsector othsector medinst smallinst
smallcity medcity, inf(IRC05 Core03 age03 totphds03 totprojects03 agroscs medscs
social human engi othscs bussector govsector othsector medinst smallinst smallcity
medcity) vuong nolog

. listcoef, help

zip (N=1889): Factor Change in Expected Count

(output omitted)

Vuong Test = 15.33 (p=0.000) favoring ZIP over PRM.


. quietly zinb totbibprod05 IRC05 Core03 age03 totphds03 totprojects03 agroscs
medscs social human engi othscs bussector govsector othsector medinst smallinst
smallcity medcity, inf(IRC05 Core03 age03 totphds03 totprojects03 agroscs medscs
social human engi othscs bussector govsector othsector medinst smallinst smallcity
medcity) vuong nolog

. listcoef, help

zinb (N=1889): Factor Change in Expected Count

Vuong Test =  5.31 (p=0.000) favoring ZINB over NBRM.
```

To compare other pair of models, such as ZINB and PRM, the *countfit* command developed by L&F (Long and Freese 2006) yields the following results:

```
. countfit totbibprod05 IRC05 Core03 age03 totphds03 totprojects03 agroscs medscs
social human engi othscs bussector govsector othsector medinst smallinst smallcity
```

```
medcity, inf(IRC05 Core03 age03 totphds03 totprojects03 agroscs medscs social human
engi othscs bussector govsector othsector medinst smallinst smallcity medcity)
noestimates nograph maxcount(20)
```

```
        Comparison of Mean Observed and Predicted Count

                    Maximum       At      Mean
        Model     Difference    Value    |Diff|
        ---------------------------------------------
        PRM          0.240         0      0.034
        NBRM        -0.038         1      0.006
        ZIP          0.080         1      0.023
        ZINB        -0.006         2      0.002
```

Based on the above table, which lists the counts for which the deviation between the observed and average expected count is greatest, the biggest problem for the PRM is the prediction of zero counts, with a difference that is much larger than the maximum for the other models. Also, the average difference between observed and predicted is larger for the PRM (0.034) and smaller for the NBRM (0.006) and ZINB (0.002).

In these four tables, we are able to see, for counts 0 to 20, the actual proportion of our data records with the given count and the predicted proportion from each model. The absolute difference, the |Diff| columns of these tables, are the ones plotted in Figure 9 shown earlier. The given count's contribution to a Pearson Chi-Square statistic is also included. It compares the actual distribution of the data and the distribution proposed by the model. For a given row, the Pearson statistic can be calculated as N(|Diff|^2) / Predicted, where N is the number of observations in the dataset. Looking at the sum of the Pearson column gives us a sense of how close the predicted proportions were to the actual proportions. Using this method to compare, the Zero-Inflated Negative Binomial appears better than the other models.

A comparison of the fit of the four models by several standard criteria and tests, including BIC and AIC provides support for the ZINB model over the others. For each statistic comparing models, the last three columns indicate which model is preferred. Both the NBRM and ZINB consistently fit better than either the PRM or the ZIP.

Overall, this summary table shows that the ZINB model fits the data best, and this conclusion makes substantive sense. As discussed earlier, there are in fact teams who for structural reasons do not produce or cannot report bibliographic products, but for other teams the failure to report products in a given period is a matter of chance. This is what the zero-inflated models are all about. As for the NBRM is concerned, it is preferable over the PRM, as it corrects for overdispersion in the outcome. In sum, the ZINB makes sense and fits the data very well.

```
Tests and Fit Statistics

PRM            BIC= 11793.614  AIC=    13.731  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs NBRM      BIC= -2721.658  dif= 14515.272  NBRM    PRM   Very strong
               AIC=     6.044  dif=     7.687  NBRM    PRM
               LRX2=14522.816  prob=    0.000  NBRM    PRM   p=0.000
-------------------------------------------------------------------------
  vs ZIP       BIC=  6248.392  dif=  5545.222  ZIP     PRM   Very strong
               AIC=    10.740  dif=     2.991  ZIP     PRM
               Vuong=  15.329  prob=    0.000  ZIP     PRM   p=0.000
-------------------------------------------------------------------------
  vs ZINB      BIC= -2698.817  dif= 14492.431  ZINB    PRM   Very strong
               AIC=     6.001  dif=     7.731  ZINB    PRM
```

```
-------------------------------------------------------------------------
NBRM           BIC= -2721.658  AIC=     6.044  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs ZIP       BIC=  6248.392  dif= -8970.050  NBRM    ZIP   Very strong
               AIC=    10.740  dif=    -4.696  NBRM    ZIP
-------------------------------------------------------------------------
  vs ZINB      BIC= -2698.817  dif=   -22.841  NBRM    ZINB  Very strong
               AIC=     6.001  dif=     0.044  ZINB    NBRM
               Vuong=   5.306  prob=    0.000  ZINB    NBRM  p=0.000
-------------------------------------------------------------------------
ZIP            BIC=  6248.392  AIC=    10.740  Prefer  Over  Evidence
-------------------------------------------------------------------------
  vs ZINB      BIC= -2698.817  dif=  8947.209  ZINB    ZIP   Very strong
               AIC=     6.001  dif=     4.739  ZINB    ZIP
               LRX2= 8954.753  prob=    0.000  ZINB    ZIP   p=0.000
-------------------------------------------------------------------------
```

The following table shows the effects of each variable depending on the model one chooses to use.

Table: 34 Team Bibliographic Production: A Comparison of Models

	PRM	NBRM	ZIP	ZINB
IRC05	0.523**	0.425**	0.371**	0.305**
	(29.65)	(6.11)	(20.89)	(4.85)
Core03	-0.006**	0.016*	0.006**	0.024**
	(4.59)	(2.07)	(3.75)	(3.41)
age03	0.000	0.004	0.003*	0.011
	(0.41)	(0.63)	(2.28)	(1.89)
totphds03	0.057**	0.086**	0.038**	0.075**
	(20.79)	(4.22)	(13.46)	(4.12)
totprojects03	0.038**	0.052**	0.031**	0.044**
	(44.27)	(8.10)	(35.42)	(7.97)
agroscs	-0.156**	-0.247	-0.134**	-0.098
	(4.28)	(1.73)	(3.61)	(0.73)
medscs	-0.127**	-0.089	-0.024	0.028
	(4.92)	(0.80)	(0.93)	(0.27)
social	-0.104**	-0.079	-0.051	0.003
	(3.88)	(0.75)	(1.87)	(0.03)
human	-0.309**	-0.221*	-0.278**	-0.181*
	(12.89)	(2.37)	(11.48)	(2.13)
engi	0.036	-0.023	0.003	0.017
	(1.41)	(0.21)	(0.11)	(0.17)
othscs	-0.011	-0.202	-0.014	-0.169
	(0.30)	(1.23)	(0.39)	(1.15)
bussector	-0.518**	-0.612**	-0.245**	-0.383
	(9.21)	(3.05)	(4.34)	(1.90)
govsector	0.030	-0.087	0.192**	0.205
	(0.69)	(0.51)	(4.33)	(1.20)
othsector	0.431**	0.185	0.353**	0.300
	(7.98)	(0.78)	(6.48)	(1.39)
medinst	0.053**	0.091	-0.010	0.036
	(2.84)	(1.26)	(0.53)	(0.55)

smallinst	-0.307**	-0.221*	-0.306**	-0.268*
	(9.47)	(2.02)	(9.35)	(2.57)
smallcity	0.123	0.053	0.203**	0.131
	(1.83)	(0.22)	(2.99)	(0.54)
medcity	-0.157**	-0.085	-0.191**	-0.139
	(7.39)	(1.08)	(9.00)	(1.94)
lnalpha:Constant		0.460**		-0.020
		(12.03)		(0.33)
inflate:IRC05			-0.617**	-0.824**
			(4.64)	(3.43)
inflate:Core03			0.036**	0.052**
			(2.71)	(2.89)
inflate:age03			0.024*	0.041*
			(2.25)	(2.57)
inflate:totphds03			-0.098*	-0.075
			(2.45)	(1.38)
inflate:totprojects03			-0.069**	-0.064**
			(4.82)	(2.89)
inflate:agroscs			0.503*	0.786
			(2.02)	(1.93)
inflate:medscs			0.451*	0.822*
			(2.27)	(2.40)
inflate:social			0.225	0.559
			(1.19)	(1.56)
inflate:human			0.164	0.288
			(0.97)	(0.84)
inflate:engi			0.043	0.216
			(0.21)	(0.57)
inflate:othscs			0.085	0.178
			(0.28)	(0.32)
inflate:bussector			0.856**	1.198*
			(2.68)	(2.51)
inflate:govsector			0.743**	1.256**
			(2.67)	(3.22)
inflate:othsector			0.049	0.579
			(0.11)	(0.94)
inflate:medinst			-0.211	-0.373
			(1.60)	(1.58)
inflate:smallinst			-0.023	-0.328
			(0.12)	(0.91)
inflate:smallcity			0.290	0.521
			(0.75)	(0.89)
inflate:medcity			-0.144	-0.363
			(1.02)	(1.31)
inflate:Constant			-0.992**	-1.957**
			(5.50)	(5.49)
Constant	1.739**	1.385**	2.035**	1.521**
	(76.43)	(13.49)	(89.07)	(15.84)
Observations	1889	1889	1889	1889

Absolute value of z statistics in parentheses
* significant at 5%; ** significant at 1%

This table illustrates the differences in the size of the coefficients (not exponentiated), the direction of the effects, and the statistical significance of the results based on all models used here to explain team output.

Clearly the size of the effect or IRC tends to reduce as we approach to the zero-inflated models. However, it remains large and statistically significant at the 0.001 level.

The effect of number of PhDs, and that of the number of projects active are consistently in the hypothesized direction across the models, and in all cases they are statistically significant.

The same is true regarding the effect of working in the humanities as opposed to working in the natural sciences, and being affiliated to small institutions as opposed to being affiliated to big institutions, which was hypothesized in the same direction it was actually found, and which was statistically significant in all models.

By contrast, the effect of team size on team output is first portrayed as having a negative effect on team productivity based on the PRM while it appears having a positive effect in the other models.

The effect of team age is consistently shown in the hypothesized direction but it is statistically significant only in the ZIP model.

The effect of working in disciplines other than in the natural sciences is shown to be negatively associated to team output in most of the models, but it is systematically statistically significant regarding the difference between the teams working in the humanities versus teams working in the natural sciences. The differences between the teams working in the agro-sciences and the reference group for predicting team bibliographic productivity is statistically significant in the PRM and the ZIP models. The difference between the reference group and the teams working in the medical sciences or the social sciences in predicting team scientific capacity is statistically significant in the PRM only. The direction of the effects of working in either of these two fields changes in the ZINB model but it is not statistically significant in either case. The differences between the teams working in the engineering or in other sciences as opposed to working in the natural sciences are statistically insignificant in all four models.

The difference between teams affiliated to the business sector compared to those affiliated to the academic sector is systematically portrayed in the same direction as hypothesized but it is not statistically significant in the ZINB model. The difference between the teams affiliated to the government and those of the reference group is found to be in the opposite direction of the research hypothesis, but it is only statistically significant in the ZIP model. Similar situation happens regarding the teams affiliated to the NGOs, except that in this case the finding is significant in both the ZIP and the PRM models.

While the difference between the teams affiliated to small institutions and those affiliated to big institutions is systematically in the direction hypothesized and is statistically significant, that between the teams affiliated to mid- size institutions is found to be in the opposite direction to that hypothesized, but it is statistically significant only in the PRM.

Finally, the difference between the teams located in mid-size cities as opposed to those located in big cities is in the direction hypothesized in all four models. However, their differences are statistically significant in the PRM and ZIP models only. In contrast, the difference between the teams located in small cities compared to those of the reference group appears to be in the opposite direction to the one hypothesized, but this finding is statistically significant in the ZIP model only.

APPENDIX H

DATA DESCRIPTION AND SUMMARY STATISTICS –RESEARCH TEAM CHARACTERISTICS AND PERFORMANCE 2003-2005

```
. d totbibprod05 ppkeycol05 Core03 age03 totphds03 totprojects03 totbibprod03 leadewritesol05 leadstudover05 natscs agroscs medscs social human engi othscs educsector bussector govsector othsector smallinst medinst biginst smallcity medcity bigcity IRC05

              storage  display     value
variable name   type   format      label      variable label
-------------------------------------------------------------------------------
totbibprod05    int    %8.0g                  Tot. Bib. Prods. 2003-5
ppkeycol05      byte   %8.0g                  'Colombia' in Prod or Proj
Core03          float  %9.0g                  Team Size in 2003
age03           float  %9.0g                  Team Age in 2003
totphds03       byte   %8.0g                  Total PhDs in 2003
totprojects03   byte   %8.0g                  Tot. Proj. in 2003
totbibprod03    int    %8.0g                  Tot. Bib. Prods. by 2003
leadewritesol05 byte   %8.0g                  Leader Writes Oth Langua
leadstudover05  byte   %8.0g                  Leader Studied Overseas
natscs          float  %9.0g                  Natural Sciences
agroscs         float  %9.0g                  Agrosciences
medscs          float  %9.0g                  Medical Sciences
social          float  %9.0g                  Social Sciences
human           float  %9.0g                  Humanities
engi            float  %9.0g                  Engineering
othscs          float  %9.0g                  Other Sciences
educsector      float  %9.0g                  Education Sector
bussector       float  %9.0g                  Business Sector
govsector       float  %9.0g                  Government
othsector       float  %9.0g                  Other Sector
smallinst       float  %9.0g                  Small Home Inst.
medinst         float  %9.0g                  Mid. Home Inst.
biginst         float  %9.0g                  Big Home Inst.
smallcity       float  %9.0g                  Small City
medcity         float  %9.0g                  Midsize City
bigcity         float  %9.0g                  Big City
IRC05           float  %9.0g       IRC        Internat. Res. Coll.
```

```
. tabstat totbibprod05 ppkeycol05 Core03 age03 totphds03 totprojects03 totbibprod03 leadewritesol05 leadstudover05 natscs agroscs medscs social human engi othscs educsector bussector govsector othsector smallinst medinst biginst smallcity medcity bigcity IRC05, statistics( sum mean sd median min max ) columns(statistics)

    variable |       sum      mean        sd       p50       min       max
-------------+------------------------------------------------------------
totbibprod05 |     16738  8.860773  13.78552         4         0       138
  ppkeycol05 |       681  .3605082  .4802751         0         0         1
      Core03 |     13711  7.258338   5.71711         6         2        74
       age03 |     13043  6.904711  5.837255         5         0        68
   totphds03 |      2749  1.455267  2.210634         1         0        47
totprojec~03 |     10865   5.75172  6.603742         4         0        70
totbibprod03 |     57323  30.34569  43.87513        15         0       458
leadewrit~05 |      1034  .5473796  .4978819         1         0         1
leadstudo~05 |      1047  .5542615  .4971786         1         0         1
      natscs |       440  .2329275  .4228084         0         0         1
     agroscs |       115  .0608788  .2391711         0         0         1
      medscs |       239   .126522  .3325247         0         0         1
      social |       305  .1614611  .3680531         0         0         1
```

```
      human |       457  .2419269  .4283637          0          0          1
       engi |       253  .1339333  .3406708          0          0          1
     othscs |        80  .0423504  .2014407          0          0          1
 educsector |      1710  .9052409  .2929595          1          0          1
  bussector |        64  .0338804  .1809691          0          0          1
  govsector |        73  .0386448  .1927979          0          0          1
  othsector |        39  .0206458  .1422333          0          0          1
  smallinst |       317  .1678137  .3737997          0          0          1
    medinst |       715  .3785071  .4851434          0          0          1
    biginst |       857  .4536792  .4979816          0          0          1
  smallcity |        33  .0174696  .1310476          0          0          1
    medcity |       414  .2191636  .4137893          0          0          1
    bigcity |      1424   .753838  .4308881          1          0          1
      IRC05 |       736  .3896241  .4877941          0          0          1
--------------------------------------------------------------------------
N= 1,889 Research Teams


. tab UNESCO

 NombreUNESCO: UNESCO's |
              S&T Area |      Freq.     Percent        Cum.
------------------------+-----------------------------------
     Ciencias Agrícolas |        115        6.09        6.09
       Ciencias Médicas |        239       12.65       18.74
      Ciencias Sociales |        305       16.15       34.89
Cs. Naturales y Exactas |        440       23.29       58.18
            Humanidades |        457       24.19       82.37
Ingeniería y Tecnología |        253       13.39       95.76
                  Otros |         80        4.24      100.00
------------------------+-----------------------------------
                  Total |      1,889      100.00

. tab sectorplus

     Team's |
Inst-Sector |
  (Cleaned) |      Freq.     Percent        Cum.
------------+-----------------------------------
   Academic |      1,710       90.67       90.67
 Enterprise |         64        3.39       94.06
 Government |         73        3.87       97.93
      Other |         39        2.07      100.00
------------+-----------------------------------
      Total |      1,886      100.00


. tab rdsizeinsti

     Team's |
Institution |
  -R&D Size |      Freq.     Percent        Cum.
------------+-----------------------------------
  smallinst |        317       16.78       16.78
    medinst |        715       37.85       54.63
    biginst |        857       45.37      100.00
------------+-----------------------------------
      Total |      1,889      100.00


. tab citysize

Team's Home |
  City-Size |      Freq.     Percent        Cum.
------------+-----------------------------------
  smallcity |         33        1.76        1.76
    medcity |        414       22.13       23.89
    bigcity |      1,424       76.11      100.00
------------+-----------------------------------
      Total |      1,871      100.00
```

APPENDIX I

BOOTSTRAP TO TEST STATISTICAL SIGNIFICANCE OF TREATMENT EFFECTS –TEAM OUTPUT

```
. bootstrap r(att), reps(999): psmatch2 IRC05 Core03 age03 totphds03
       leadewritesol05 leadstudover05 totprojects03 totbibprod03 agroscs medscs
       social human engi othscs medinst smallinst smallcity medcity, kernel
       outcome(totbibprod05)kerneltype (normal) bwidth (0.01) common logit
       quietly
(running psmatch2 on estimation sample)
Note: S.E. for ATT does not take into account that the propensity score is
       estimated.

Bootstrap replications (999)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................    50
..................................................   100
..................................................   150
..................................................   200
..................................................   250
..................................................   300
..................................................   350
..................................................   400
..................................................   450
..................................................   500
..................................................   550
..................................................   600
..................................................   650
..................................................   700
..................................................   750
..................................................   800
..................................................   850
..................................................   900
..................................................   950
.................................................

Bootstrap results                               Number of obs      =      1889
                                                Replications       =       999

      command:     psmatch2 IRC05 Core03 age03 totphds03 leadewritesol05
       leadstudover05 totprojects03 totbibprod03 agroscs medscs social human engi
       othscs medinst smallinst smallcity medcity, kernel outcome(totbibprod05)
       kerneltype(normal) bwidth(0.01) common logit quietly
        _bs_1:  r(att)

------------------------------------------------------------------------------
             |   Observed   Bootstrap                         Normal-based
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |   2.079651   1.059307     1.96   0.050     .0034466    4.155855
------------------------------------------------------------------------------
```

According to Caliendo and Kopeining 2008 the distribution of these means approximate the sampling distribution (and thus the standard error) of the population mean.

APPENDIX J

DATA DESCRIPTION AND SUMMARY STATISTICS –

INTERNATIONAL RESEARCH COLLABORATION 2003-2005

```
. d ircpeople05 ircproj05 ircN05 ircS05 peopN05 peopS05 projN05 projS05

              storage  display     value
variable name   type   format      label      variable label
-------------------------------------------------------------------------------
ircpeople05     byte   %8.0g                  Team has Foreign Researchers
ircproj05       byte   %8.0g                  Does Team has Foreign Funding?
ircN05          float  %9.0g                  1 if peopN05 + projN05 >=1
ircS05          float  %9.0g                  1 if peopS05 + projS05 >=1
peopN05         byte   %8.0g                  People from North? 2003-5
peopS05         byte   %8.0g                  People from South? 2003-5
projN05         byte   %8.0g                  Funding from North? 2003-5
projS05         byte   %8.0g                  Funding from South? 2000-5

. sum ircpeople05 ircproj05 ircN05 ircS05 peopN05 peopS05 projN05 projS05

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
 ircpeople05 |      1889    .2149285    .4108815          0          1
   ircproj05 |      1889    .2588671    .4381285          0          1
-------------+--------------------------------------------------------
      ircN05 |      1889    .3128639    .4637822          0          1
      ircS05 |      1889    .1604023     .367076          0          1
-------------+--------------------------------------------------------
     peopN05 |      1889    .1344627    .3412391          0          1
     peopS05 |      1889    .1064055    .3084376          0          1
-------------+--------------------------------------------------------
     projN05 |      1889    .2382213     .426108          0          1
     projS05 |      1889    .0629963    .2430206          0          1
```

APPENDIX K

DATA DESCRIPTION AND SUMMARY STATISTICS –SAMPLE

```
. d artkeycol035 artirc012

              storage  display     value
variable name   type   format      label      variable label
-------------------------------------------------------------------------------
artkeycol035    byte   %8.0g                  Art. w Kword 'Colombia'? 2003-5
artirc012       byte   %8.0g                  Int. Co-Authorship in 2001-2
```

```
. sum totbibprod05 artkeycol035 Core03 age03 totphds03 totprojects03 totbibprod03
    natscs agroscs medscs engi othscs educsector bussector govsector othsector
    biginst medinst smallinst smallcity medcity bigcity artirc012

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
totbibprod05 |       672    9.133929    13.76046          0        138
artkeycol035 |       672    .1770833    .3820236          0          1
      Core03 |       672     6.59375    5.631783          0         60
       age03 |       672    7.537202    6.617399          0         68
   totphds03 |       672    1.611607    2.030186          0         12
-------------+--------------------------------------------------------
totprojec~03 |       672    6.080357    7.982438          0         70
totbibprod03 |       672    32.61161    45.16276          0        348
      natscs |       672    .4092262    .4920573          0          1
     agroscs |       672     .110119    .3132713          0          1
      medscs |       672    .2142857    .4106315          0          1
-------------+--------------------------------------------------------
        engi |       672    .2276786    .4196464          0          1
      othscs |       672    .0357143    .1857151          0          1
  educsector |       672    .8690476    .3375996          0          1
   bussector |       672    .0610119    .2395304          0          1
   govsector |       672    .0565476    .2311482          0          1
-------------+--------------------------------------------------------
   othsector |       672    .0119048    .1085383          0          1
     biginst |       672    .4806548    .4999978          0          1
     medinst |       672    .3497024    .4772311          0          1
   smallinst |       672    .1696429    .3755983          0          1
   smallcity |       672    .0267857    .1615769          0          1
-------------+--------------------------------------------------------
     medcity |       672    .2291667    .4206098          0          1
      bigcity |       672     .735119    .4415985          0          1
   artirc012 |       672     .235119    .4243891          0          1
```

APPENDIX L

RESEARCH TEAM OUTPUT: ZINB USING ALL TYPES OF COLLABORATION AND PARTNERS

	(1)	(2)	(3)	(4)
IRC05	0.305**			
	(4.85)			
ircpeople05		0.125		
		(1.82)		
ircproj05		0.343**		
		(5.02)		
ircN05			0.165*	
			(2.54)	
ircS05			0.376**	
			(5.05)	
peopN05				-0.067
				(0.82)
peopS05				0.273**
				(3.08)
projN05				0.248**
				(3.55)
projS05				0.427**
				(4.07)
Core03	0.024**	0.025**	0.023**	0.024**
	(3.41)	(3.54)	(3.30)	(3.47)
age03	0.011	0.011	0.009	0.008
	(1.89)	(1.95)	(1.64)	(1.45)
totphds03	0.075**	0.071**	0.076**	0.075**
	(4.12)	(3.88)	(4.20)	(4.10)
totprojects03	0.044**	0.042**	0.043**	0.041**
	(7.97)	(7.65)	(7.86)	(7.45)
agroscs	-0.098	-0.096	-0.110	-0.126
	(0.73)	(0.72)	(0.83)	(0.95)
medscs	0.028	0.020	0.008	-0.018
	(0.27)	(0.19)	(0.08)	(0.18)
social	0.003	-0.002	-0.024	-0.040
	(0.03)	(0.02)	(0.25)	(0.41)
human	-0.181*	-0.167*	-0.203*	-0.186*
	(2.13)	(1.96)	(2.39)	(2.21)
engi	0.017	0.018	0.014	-0.013
	(0.17)	(0.18)	(0.14)	(0.14)
othscs	-0.169	-0.213	-0.220	-0.255
	(1.15)	(1.45)	(1.51)	(1.74)
bussector	-0.383	-0.401*	-0.432*	-0.435*

	(1.90)	(1.99)	(2.16)	(2.18)
govsector	0.205	0.158	0.224	0.194
	(1.20)	(0.92)	(1.32)	(1.14)
othsector	0.300	0.255	0.267	0.189
	(1.39)	(1.18)	(1.25)	(0.89)
medinst	0.036	0.051	0.055	0.051
	(0.55)	(0.78)	(0.84)	(0.79)
smallinst	-0.268*	-0.239*	-0.258*	-0.248*
	(2.57)	(2.30)	(2.51)	(2.42)
smallcity	0.131	0.109	0.125	0.102
	(0.54)	(0.45)	(0.52)	(0.43)
medcity	-0.139	-0.146*	-0.134	-0.149*
	(1.94)	(2.04)	(1.88)	(2.11)
inflate:IRC05	-0.824**			
	(3.43)			
inflate:ircpeople05		-0.028		
		(0.12)		
inflate:ircproj05		-1.341**		
		(4.02)		
inflate:ircN05			-0.938**	
			(3.57)	
inflate:ircS05			-0.245	
			(0.90)	
inflate:peopN05				-0.291
				(0.96)
inflate:peopS05				0.252
				(0.93)
inflate:projN05				-0.988**
				(3.23)
inflate:projS05				-1.905
				(1.67)
inflate:Core03	0.052**	0.051**	0.053**	0.051**
	(2.89)	(2.88)	(3.00)	(2.88)
inflate:age03	0.041*	0.042**	0.037*	0.040*
	(2.57)	(2.72)	(2.38)	(2.56)
inflate:totphds03	-0.075	-0.070	-0.063	-0.082
	(1.38)	(1.22)	(1.19)	(1.38)
inflate:totprojects03	-0.064**	-0.054**	-0.062**	-0.055**
	(2.89)	(2.62)	(2.84)	(2.62)
inflate:agroscs	0.786	0.791	0.664	0.627
	(1.93)	(1.95)	(1.72)	(1.63)
inflate:medscs	0.822*	0.876*	0.683*	0.717*
	(2.40)	(2.53)	(2.12)	(2.27)
inflate:social	0.559	0.559	0.406	0.400
	(1.56)	(1.56)	(1.20)	(1.22)
inflate:human	0.288	0.232	0.144	0.112
	(0.84)	(0.68)	(0.44)	(0.36)
inflate:engi	0.216	0.228	0.077	0.042
	(0.57)	(0.60)	(0.21)	(0.12)
inflate:othscs	0.178	0.225	0.077	0.194
	(0.32)	(0.39)	(0.14)	(0.36)
inflate:bussector	1.198*	1.143*	1.130*	1.080*

	(2.51)	(2.42)	(2.44)	(2.31)
inflate:govsector	1.256**	1.333**	1.190**	1.261**
	(3.22)	(3.44)	(3.21)	(3.39)
inflate:othsector	0.579	0.834	0.569	0.646
	(0.94)	(1.38)	(0.94)	(1.03)
inflate:medinst	-0.373	-0.343	-0.277	-0.316
	(1.58)	(1.47)	(1.23)	(1.43)
inflate:smallinst	-0.328	-0.293	-0.233	-0.281
	(0.91)	(0.84)	(0.69)	(0.84)
inflate:smallcity	0.521	0.550	0.444	0.511
	(0.89)	(0.91)	(0.76)	(0.88)
inflate:medcity	-0.363	-0.339	-0.360	-0.340
	(1.31)	(1.22)	(1.36)	(1.30)
inflate:Constant	-1.957**	-2.061**	-1.857**	-1.865**
	(5.49)	(5.65)	(5.67)	(5.85)
lnalpha:Constant	-0.020	-0.026	-0.047	-0.060
	(0.33)	(0.41)	(0.76)	(0.98)
Constant	1.521**	1.525**	1.558**	1.592**
	(15.84)	(15.91)	(16.52)	(16.92)
Observations	1889	1889	1889	1889

Absolute value of z statistics in parentheses
* significant at 5%; ** significant at 1%

APPENDIX M

BOOTSTRAP TO TEST STATISTICAL SIGNIFICANCE OF TREATMENT EFFECTS –TEAM CONTRIBUTION TO LOCAL KNOWLEDGE

```
. bootstrap r(att), reps(999): psmatch2 IRC05 Core03 age03 totphds03
    leadewritesol05 leadstudover05 totprojects03 totbibprod03 agroscs medscs social
    human engi othscs medinst smallinst smallcity medcity, kernel
    outcome(ppkeycol05)kerneltype (normal) bwidth (0.01) common logit quietly
(running psmatch2 on estimation sample)
Note: S.E. for ATT does not take into account that the propensity score is
    estimated.

Bootstrap replications (999)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................    50
..................................................   100
..................................................   150
..................................................   200
..................................................   250
..................................................   300
..................................................   350
..................................................   400
..................................................   450
..................................................   500
..................................................   550
..................................................   600
..................................................   650
..................................................   700
..................................................   750
..................................................   800
..................................................   850
..................................................   900
..................................................   950
.................................................

Bootstrap results                               Number of obs      =      1889
                                                Replications       =       999

      command:  psmatch2 IRC05 Core03 age03 totphds03 leadewritesol05
    leadstudover05 totprojects03 totbibprod03 agroscs medscs social human engi
    othscs medinst smallinst smallcity medcity, kernel outcome(ppkeycol05)
    kerneltype(normal) bwidth(0.01) common logit quietly
        _bs_1:  r(att)

------------------------------------------------------------------------------
             |   Observed   Bootstrap                         Normal-based
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _bs_1 |    .078314   .0325944     2.40   0.016      .0144302    .1421978
------------------------------------------------------------------------------
```

REFERENCES

Acs, Z., J. de la Mothe, et al. (1996). Local Systems of Innovation: In Search of an Enabling Strategy. The Implications of Knowledge-Based Growth for Micro-Economic Policies. P. Howitt. Calgary, University of Calgary Press: 339-358.

Adams, J. D., G. C. Black, et al. (2005). "Scientific teams and institutional collaborations: Evidence from US universities, 1981-1999." Research Policy **34**(3): 259-285.

Allen, T. (1977). Managing the flow of technology, MIT Press.

Almus, M. and D. Czarnitzki (2003). "The effects of public R&D subsidies on firms' innovation activities: The case of Eastern Germany." Journal of Business & Economic Statistics **21**(2): 226-236.

Amsterdamska, O. (2008). Practices, People, and Places. The Handbook of Science and Technology Studies. E. Hackett, O. Amsterdamska, M. Lynch and J. Wajcman. Cambridge MA., The MIT Press: 205-209.

Andersson, T., S. S. Serger, et al. (2004). The Cluster Policies Whitebook. Stortorget, Sweden, Chapters 1-3. IKED, Holmbergs.

Andrews, F. M., Ed. (1979). Scientific Productivity: the effectiveness of research groups in six countries. Paris, Unesco.

Archibugi, D. and A. Coco (2004). "A new indicator of technological capabilities for developed and developing countries (ArCo)." World Development **32**(4): 629-654.

Babu, A. R. and Y. P. Singh (1998). "Determinants of research productivity." Scientometrics **43**(3): 309-329.

Barro, R. J. and J. W. Lee (2001). "International data on educational attainment: Updates and implications." Oxford Economic Papers-New Series **53**(3): 541-563.

Basu, A. and R. Aggarwal (2001). "International collaboration in science in India and its impact on institutional performance." **52**(3): 379-394.

Basu, A. and B. S. V. Kumar (2000). "International collaboration in Indian scientific papers." Scientometrics **48**(3): 381-402.

Bauer, H. H. (1990). "Barriers against Interdisciplinarity - Implications for Studies of Science, Technology, and Society (Sts)." Science Technology & Human Values **15**(1): 105-119.

Bayona, C., T. Garcia-Marco, et al. (2001). "Firms' motivations for cooperative R&D: an empirical analysis of Spanish firms." Research Policy **30**(8): 1289-1307.

Beaver, D. D. (2001). "Reflections on scientific collaboration, (and its study): past, present, and future." Scientometrics **52**(3): 365-377.

Beaver, D. D. (2004). "Does collaborative research have greater epistemic authority?" Scientometrics **60**(3): 399-408.

Beaver, D. D. and R. Rosen (1979). "Studies in Scientific Collaboration .2. Scientific Co-Authorship, Research Productivity and Visibility in the French Scientific Elite, 1799-1830." Scientometrics **1**(2): 133-149.

Beaver, D. D. and R. Rosen (1979). "Studies in Scientific Collaboration .3. Professionalization and the Natural-History of Modern Scientific Co-Authorship." Scientometrics **1**(3): 231-245.

Becher, T. (1981). "Towards a Definition of Disciplinary Cultures." Studies in Higher Education **6**(2): 109-122.

Becher, T. (1994). "The Significance of Disciplinary Differences." Studies in Higher Education **19**(2): 151-161.

Becker, G. (1964). Human capital: A theoretical and empirical analysis, with a special reference to education. Chicago, University of Chicago Press.

Behrens, T. R. and D. O. Gray (2001). "Unintended consequences of cooperative research: impact of industry sponsorship on climate for academic freedom and other graduate student outcome." Research Policy **30**(2): 179-199.

Belderbos, R., M. Carree, et al. (2004). "Cooperative R&D and firm performance." Research Policy **33**(10): 1477-1492.

Bonaccorsi, A. and C. Daraio (2005). "Exploring size and agglomeration effects on public research productivity." Scientometrics **63**(1): 87-120.

Bonaccorsi, A., C. Daraio, et al. (2006). "Advanced indicators of productivity of universities. An application of robust nonparametric methods to Italian data." Scientometrics **66**(2): 389-410.

Bordons, M. and I. Gomez (2000). Collaboration networks in science. Web of Knowledge - a Festschrift in Honor of Eugene Garfield. B. Cronin and H. B. Atkins. Medford, NJ., Information Today Inc**:** 197-213.

Bordons, M., I. Gomez, et al. (1996). "Local, domestic and international scientific collaboration in biomedical research." Scientometrics **37**(2): 279-295.

Bordons, M. and M. A. Zulueta (1997). "Comparison of research team activity in two biomedical fields." Scientometrics **40**(3): 423-436.

Bordons, M., M. A. Zulueta, et al. (1998). "Scientific activity of the most productive Spanish research teams in pharmacology and pharmacy during the period 1986-1993 as covered by the Science Citation Index (SCI)." Medicina Clinica **111**(13): 489-495.

Bozeman, B. and E. Corley (2004). "Scientists' collaboration strategies: implications for scientific and technical human capital." Research Policy **33**(4): 599-616.

Bozeman, B., J. S. Dietz, et al. (2001). "Scientific and technical human capital: an alternative model for research evaluation." International Journal of Technology Management **22**(7-8): 716-740.

Bozeman, B. and J. D. Rogers (2002). "A churn model of scientific knowledge value: Internet researchers as a knowledge value collective." Research Policy **31**(5): 769-794.

Burt, R. S. (2004). "Structural holes and good ideas." American Journal of Sociology **110**(2): 349-399.

Busom, I. (2000). "An Empirical Evaluation of the Effects of R&D Subsidies." Economics of Innovation & New Technology **9**(2): 111-148.

Calero, C., R. Buter, et al. (2006). "How to identify research groups using publication analysis: an example in the field of nanotechnology." Scientometrics **66**(2): 365-376.

Caliendo, M. and S. Kopeining (2008). "Some Practical Guidance for Implementation of Propensity Score Matching." Journal of Economic Surveys **22**(1): 31-72.

Callon, M. (1992). The Dynamics of Techno-economic Networks. Technological Change and Company Strategies. R. Coombs, P. Saviotti and V. Walsh. London, Academic Press.**:** 72-102.

Callon, M., J. P. Courtial, et al. (1991). "Co-Word Analysis as a Tool for Describing the Network of Interactions between Basic and Technological Research - the Case of Polymer Chemistry." Scientometrics **22**(1): 155-205.

Carayannis, E. G., J. Alexander, et al. (2000). "Leveraging knowledge, learning, and innovation in forming strategic government-university-industry (GUI) R&D partnerships in the US, Germany, and France." Technovation **20**(9): 477-488.

Carayannis, E. G. and P. Laget (2004). "Transatlantic innovation infrastructure networks: public-private, EU-US R&D partnerships." R & D Management **34**(1): 17-31.

Carayol, N. and M. Matt (2004a). "The exploitation of complementarities in scientific production process at the laboratory level." Technovation **24**(6): 455-465.

Carayol, N. and M. Matt (2004b). "Does research organization influence academic production? Laboratory level evidence from a large European university." Research Policy **33**(8): 1081-1102.

Carayol, N. and M. Matt (2006). "Individual and collective determinants of academic scientists' productivity." Information Economics and Policy **18**(1): 55-72.

Cardinal, L. B. and D. E. Hatfield (2000). "Internal knowledge generation: the research laboratory and innovative productivity in the pharmaceutical industry." Journal of Engineering and Technology Management **17**(3-4): 247-271.

Casper, S. and A. Karamanos (2003). "Commercializing science in Europe: The Cambridge biotechnology cluster." European Planning Studies **11**(7): 805-822.

Chaparro, F., H. Jaramillo, et al. (2004). The role of diaspora in facilitating participation in global knowledge networks: lessons of the Red Caldas in Colombia. Washington, D.C., Report prepared for the Knowledge for Development Program. The World Bank**:** 25.

Cohen, J. E. (1980). "Publication Rate as a Function of Laboratory Size in a Biomedical-Research Institution." Scientometrics **2**(1): 35-52.

Cohen, J. E. (1981). "Publication Rate as a Function of Laboratory Size in 3 Biomedical-Research Institutions." Scientometrics **3**(6): 467-487.

Cohen, J. E. (1991). "Size, Age and Productivity of Scientific and Technical Research Groups." Scientometrics **20**(3): 395-416.

Cohen, W. M. and D. A. Levinthal (1990). "Absorptive capacity: A new perspective on learning and innovation." Administrative Science Quarterly [Special Issue: Technology, organizations, and innovation] **35**(1): 128-152.

Colciencias (2000a). Módulo de Cálculo para Escalafonamiento de Grupos y Centros de Investigación Científica y Tecnológica -2000. Convocatoria Nacional para Grupos y Centros de Investigación Científica y Tecnológica 2000. Bogotá, D.C. Colombia**:** 20p.

Cole, J. R. (2000). A Short history of the use of citations as a measure of the impact of scientific and scholarly work. Web of Knowledge - a Festschrift in Honor of Eugene Garfield. B. Cronin and H. B. Atkins. Medford, NJ., Information Today Inc**:** 281-300.

Cole, S. (1979). "Age and Scientific Performance." American Journal of Sociology **84**(4): 958-977.

Coleman, J. (1988). "Social Capital in the Creation of Human Capital." American Journal of Sociology **94**(Supp.): S95-s120.

Cozzens, S., S. Gatchair, et al. (2008). Knowledge and Development. The Handbook of Science and Technology Studies. E. J. Hackett, O. Amsterdamska, M. Lynch and J. Wajcman. Cambridge, MA, MIT Press**:** 787-812.

Cozzens, S. E. (1999). "Are new accountability rules bad for science?" Issues in Science and Technology **15**(4): 59-66.

Cozzens, S. E. (1999). "Results and Responsibility: Science, Society, and GPRA." AAAS Science and Technology Policy Yearbook: 165-172.

Cozzens, S. E., K. Bobb, et al. (2005). "Distributional effects of science and technology-based economic development strategies at state level in the United States." Science & Public Policy **32**(1): 29-38.

Cozzens, S. E., S. Popper, et al. (1994). Methods for Evaluating Fundamental Science., Critical Technologies Institute, Rand Corporation.

Crane, D. (1972). Invisible Colleges: diffusion of knowledge in scientific communities. Chicago, IL, The University of Chicago Press.

Cronin, B. (1982). "Progress in documentation: invisible colleges and information transfer, a review and commentary with particular reference to the social sciences." Journal of Documentation(38): 212-236.

Cronin, B., D. Shaw, et al. (2004). "Visible, less visible, and invisible work: Patterns of collaboration in 20th century chemistry." Journal of the American Society for Information Science and Technology **55**(2): 160-168.

Crow, M. and B. Bozeman (1998). Limited by Design: R&D Laboratories in the U.S. National Innovation Systems. New York, Chapter 1. Columbia University Press.

Cummings, J. N. and S. Kiesler (2005). "Collaborative research across disciplinary and organizational boundaries." Social Studies of Science **35**(5): 703-722.

Dahl, M. S. and C. O. R. Pedersen (2004). "Knowledge flows through informal contacts in industrial clusters: myth or reality?" Research Policy **33**(10): 1673-1686.

DANE (2006). Innovacion y Desarrollo Tecnologico en la Industria Manufacturera. Colombia 2003-2004. Bogota, D.C., Departamento Administrativo Nacional de Estadistica -DANE, Departamento Nacional de Planeacion -DNP, Insituto Colombiano para el Desarrollo de la Ciencia y la Tecnologia -COLCIENCIAS.

Danilovic, M. and W. Mats (2005). "A tentative framework for analyzing integration in collaborative manufacturing network settings: a case study." Journal of Engineering and Technology Management **22**(1-2): 141-158.

David, P. A. and J. Goddard L (2001). Knowledge, Capabilities and Human Capital Formation in Economic Growth. New Zealand Treasury. Wellington, NZ**:** 155.

Diamond, A. M. (1985). "The Money Value of Citations to Single-Authored and Multiple-Authored Articles." Scientometrics **8**(5-6): 315-320.

Dietz, J. (2004). Scientists and Engineers in Academic Research Centers -An examination of career patterns and productivity. PhD Thesis. School of Public Policy. Atlanta, GA, Georgia Institute of Technology**:** 186.

Duque, R. B., M. Ynalvez, et al. (2005). "Collaboration paradox: Scientific productivity, the Internet, and problems of research in developing areas." Social Studies of Science **35**(5): 755-785.

Ettorre, E. (2000). "Recognizing diversity and group processes in international, collaborative research work: A case study." Social Policy & Administration **34**(4): 392-407.

Etzkowitz, H. and C. Kemelgor (1998). "The Role of Research Centers in the Collectivisation of Academic Science." Minerva **36**: 271-288.

Etzkowitz, H. and L. Leydesdorff (2000). "The dynamics of innovation: from National Systems and "Mode 2" to a Triple Helix of university-industry-government relations." Research Policy **29**(2): 109-123.

Farrell, M. P. (2001). Collaborative circles: friendship dynamics and creative work. Chicago, University of Chicago Press.

Fleming, L. (2001). "Recombinant uncertainty in technological search." Management Science **47**(1): 117-132.

Florida, R. (1999). "The Role of the University: Leveraging Talent, Not Technology." Issues in Science & Technology(Summer).

Forero-Pineda, C. and H. Jaramillo-Salazar (2002). "The access of researchers from developing countries to international science and technology." International Social Science Journal **54**(1): 129-+.

Fountain, J. (1998). Social Capital: a key enabler of innovation. Investing in Innovation. Creating a Research and Innovation Policy that Works. L. M. Branscomb and J. H. Keller. Cambridge, Massachusetts, The MIT Press.

Fox, M. F. and C. A. Faver (1984). "Independence and Cooperation in Research - the Motivations and Costs of Collaboration." Journal of Higher Education **55**(3): 347-359.

Fox, M. F. and C. A. Faver (1985). "Men, Women, and Publication Productivity: Patterns among Social Work Academics." Sociological Quarterly **26**(4): 537–49.

Frame, J. D. and M. P. Carpenter (1979). "International Research Collaboration." Social Studies of Science **9**(4): 481-497.

Frederiksen, L. F. (2004). "Disciplinary determinants of bibliometric impact in Danish industrial research: Collaboration and visibility." Scientometrics **61**(2): 253-270.

Frenken, K., W. Hölzl, et al. (2005). "The citation impact of research collaborations: the case of European biotechnology and applied microbiology (1988–2002)." Journal of Engineering and Technology Management **22**(1-2): 9-30.

Garay, L. J. (1998). Estructura y Dinamica Industrial. Colombia: Estructura Industrial e Internacionalizacion 1967-1996. L. J. Garay, L. F. Quintero, J. A. Villamil et al. Bogota, Colombia, Tercer Mundo Editores.

Gelijns, A. C. and S. O. Thier (2002). "Medical Innovation and Institutional Interdependence: Rethinking University-Industry Connections." JAMA. Journal of the American Medical Association **287**: 72-77.

George, G., S. A. Zahra, et al. (2002). "The effects of business-university alliances on innovative output and financial performance: a study of publicly traded biotechnology companies." Journal of Business Venturing **17**(6): 577-609.

Georghiou, L. (1998). "Global cooperation in research." Research Policy **27**(6): 611-626.

Georghiou, L. (2001). "Evolving frameworks for European collaboration in research and technology." Research Policy **30**(6): 891-903.

Gibbons, M., C. Limoges, et al. (1994). The New Production of Knowledge: The dynamics of science and research in contemporary societies., Chapters 1-2. Sage Publications.

Glanzel, W. and A. Schubert (2005). "Domesticity and internationality in co-authorship, references and citations." Scientometrics **65**(3): 323-342.

Glänzel, W. and A. Schubert (2004). Analyzing scientific networks through co-authorship. Handbook of Quantitative Science and Technology Research: The Use of Publication and Patent Statistics in studies on S&T Systems. H. Moed, W. Glänzel and U. Schmoch. Dordrecht, The Netherlands, Kluwer Academic Publishers.**:** 257–76.

Gmur, M. (2003). "Co-citation analysis and the search for invisible colleges: A methodological evaluation." Scientometrics **57**(1): 27-57.

Godin, B. and Y. Gingras (2000). "Impact of collaborative research on academic science." Science and Public Policy **27**(1): 9.

Granovetter, M. (1973). "The Strength of Weak Ties." The American Journal of Sociology **78**(6): 1380.

Granovetter, M. (1983). "The strength of weak ties: a network theory revisited." Sociological Theory **1**: 201-233.

Granovetter, M. (2005). "The impact of social structure on economic outcomes." Journal of Economic Perspectives **19**(1): 33-50.

Guan, J. C. and J. X. Wang (2004). "Evaluation and interpretation of knowledge production efficiency." Scientometrics **59**(1): 131-155.

Hagedoorn, J. (2002). "Inter-firm R&D partnerships: an overview of major trends and patterns since 1960." Research Policy **31**(4): 477-492.

Hagedoorn, J., A. N. Link, et al. (2000). "Research partnerships." Research Policy **29**(4-5): 567-586.

Harrison, D. A., K. H. Price, et al. (2002). "Time, teams, and task performance: Changing effects of surface- and deep-level diversity on group functioning." Academy of Management Journal **45**(5): 1029-1045.

Harsanyi, M. A. (1993). "Multiple Authors, Multiple Problems Bibliometrics and the Study of Scholarly Collaboration - a Literature-Review." Library & Information Science Research **15**(4): 325-354.

Heimeriks, G., M. Horlesberger, et al. (2003). "Mapping communication and collaboration in heterogeneous research networks." Scientometrics **58**(2): 391-413.

Helble, Y. and L. C. Chong (2004). "The importance of internal and external R&D network linkages for R&D organisations: evidence from Singapore." R & D Management **34**(5): 605-612.

Herbertz, H. (1995). "Does It Pay to Cooperate - a Bibliometric Case-Study in Molecular-Biology." Scientometrics **33**(1): 117-122.

Hicks, D. M., P. A. Isard, et al. (1996). "A morphology of Japanese and European corporate research networks." Research Policy **25**(3): 359-378.

Hicks, D. M. and J. S. Katz (1996). "Where is science going?" Science Technology & Human Values **21**(4): 379-406.

Hoegl, M. and L. Proserpio (2004). "Team member proximity and teamwork in innovative projects." Research Policy **33**(8): 1153-1165.

Holbrook, J. A. D. and M. Salazar (2004). "Regional Innovation Systems within a Federation: Do National Policies Affect all Regions Equally?" Innovation: Management, Policy & Practice **6**(1): 50 - 64.

Holbrook, J. A. D. and D. A. Wolfe, Eds. (2000). Innovation, Institutions and Territory: Regional Innovation Systems. Kingston, McGill-Queen's University Press.

Holbrook, J. A. D. and D. A. Wolfe, Eds. (2002). Knowledge, Clusters and Regional Innovation: Economic Development in Canada. Kingston, McGill-Queen's University Press.

IAC (2004). Inventing a better future: a strategy for building worldwide capacities in science and technology. Amsterdam, The Netherlands, InterAcademy Council.

Imbens, G. (2004). "Nonparametric Estimation of Average Treatment Effects under Exogeneity: A Review." The Review of Economics and Statistics **86**(1): 4-29.

Jaramillo, H. (2007). Colombia: evolucion, contexto y resultados de las politicas de ciencia, tecnologia e innovacion. Claves del desarrollo cientifico y tecnologico en America Latina. J. Sebastian. Madrid, Espana, Fundacion Carolina y Siglo XXI de Espana Editores**:** 458.

Kastelli, I., Y. Caloghirou, et al. (2004). "Cooperative R&D as a means for knowledge creation. Experience from European publicly funded partnerships." International Journal of Technology Management **27**(8): 712-730.

Katz, J. S. (1994). "Geographical Proximity and Scientific Collaboration." Scientometrics **31**(1): 31-43.

Katz, J. S. and D. Hicks (1997). "How much is a collaboration worth? A calibrated bibliometric model." Scientometrics **40**(3): 541-554.

Katz, J. S. and B. R. Martin (1997). "What is research collaboration?" Research Policy **26**(1): 1-18.

Kleinman, D. (1998). "Pervasive Influence: Intellectual Property, Industrial History, and University Science." Science and Public Policy: 95-102.

Klette, T. J., J. Moen, et al. (2000). "Do subsidies to commercial R&D reduce market failures? Microeconometric evaluation studies." Research Policy **29**(4-5): 471-495.

Kretschmer, H. (1985). "Cooperation Structure, Group-Size and Productivity in Research Groups." Scientometrics **7**(1-2): 39-53.

Kuhn, T. S. (1966). The Structure of Scientific Revolutions. Chicago, The University of Chicago Press.

Kyvik, S. (1995). "Are Big University Departments Better Than Small Ones." Higher Education **30**(3): 295-304.

Landry, R. and N. Amara (1998). "The impact of transaction costs on the institutional structuration of collaborative academic research." Research Policy **27**(9): 901-913.

Landry, R., N. Amara, et al. (2002). "Does social capital determine innovation? To what extent?" Technological Forecasting and Social Change **69**(7): 681-701.

Landry, R., N. Traore, et al. (1996). "An econometric analysis of the effect of collaboration on academic research productivity." Higher Education **32**(3): 283-301.

Laredo, P. (2003). "University Research Activities: On-going Transformations and New Challenges." Higher Education Management and Policy **15**(1): 105-123.

Laredo, P. and P. Mustar, Eds. (2001). Research and Innovation Policies in the New Global Economy. An International Comparative Analysis., Edward Elgar.

Laudel, G. (2001). "Collaboration, creativity and rewards: why and how scientists collaborate." International Journal of Technology Management **22**(7-8): 762-781.

Laudel, G. (2002). "What do we measure by co-authorships?" Research Evaluation **11**(1): 3-15.

Le Bas, C., F. Picard, et al. (1998). "Innovation technologique, comportement de reseaux et performances: Une analyse sur donnees individuelles." Revue D'Economie Politique **108**(5): 625–644.

Leclerc, M. and J. Gagne (1994). "International Scientific Cooperation - the Continentalization of Science." Scientometrics **31**(3): 261-292.

Lee, S. (2004). Forein-born scientists in the United States -Do they performe differently than native-born scientists? School of Public Policy. Atlanta, GA, Georgia Institute of Technology. PhD Dissertation**:** 228.

Lee, S. and B. Bozeman (2005). "The impact of research collaboration on scientific productivity." Social Studies of Science **35**(5): 673-702.

Leuven, E. and B. Sianesi (2003). PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing, Boston College Department of Economics.

Levin, S. and P. Stephan (1991). "Research Productivity over the Life Cycle: evidence from academic research." American Economic Review **81**(1): 114-132.

Levine, J. M. and R. L. Moreland (2004). "Collaboration: The social context of theory development." Personality and Social Psychology Review **8**(2): 164-172.

Lewis, G. (2002). Advanced Techniques. Lecture Notes 06. PAUS 9111. Advance Research Methods I. Georgia State University. Atlanta, GA.

Lewis, G. (2003). Dichotomous Dependent Variables. Lecture Notes 01. PAUS 9121. Advance Research Methods II. Georgia State University. Atlanta, GA.

Leydesdorff, L. and H. Etzkowitz (1998). "The Triple Helix as a model for innovation studies." Science and Public Policy **25**(3): 195-203.

Leydesdorff, L. and M. Meyer (2003). "The Triple Helix of university-industry-government relations." Scientometrics **58**(2): 191-203.

Liang, L. M. and L. Zhu (2002). "Major factors affecting China's inter-regional research collaboration: Regional scientific productivity and geographical proximity." Scientometrics **55**(2): 287-316.

Lima, M., S. Liberman, et al. (2005). "Scientific group cohesiveness at the National University of Mexico." Scientometrics **64**(1): 55-66.

Link, A. N., D. Paton, et al. (2002). "An analysis of policy initiatives to promote strategic research partnerships." Research Policy **31**(8-9): 1459-1466.

Long, J. S. (1992). " Measure of Sex Differences in Scientific Productivity." Social Forces **71**(1): 159–78.

Long, J. S., P. Allison, et al. (1993). "Rank advancement in academic careers: sex differences and the effects of productivity." American Sociological Review **58**: 703-722.

Long, J. S. and J. Freese (2001). Regression Models for Categorical Dependent Variables Using Stata. College Station, Texas, Stata Corporation.

Long, J. S. and J. Freese (2006). Regression Models for Categorical Dependent Variables Using Stata. College Station, Texas, Stata Press.

Lundvall, B. A., Ed. (1992). National Systems of Innovation. London, Printer.

Luukkonen, T., O. Persson;, et al. (1992). "Understanding Patterns of International Scientific Collaboration." Science, Technology, & Human Values **17**(1): 101-126.

Malo, S. and A. Geuna (2000). "Science-technology linkages in an emerging research platform: The case of combinatorial chemistry and biology." Scientometrics **47**(2): 303-321.

Martinez, E. (2005). Institucionalizacion de la CyT, Copperacion Internacional y Redes de Conocimiento en A.L. Lecture at the International Seminar "Redes de conocimiento como nueva forma de creacion colaborativa: su construccion, dinamica y gestion". Buenos Aires, Argentina. Nov. 24-25, RICYT, CYTED, UNESCO.

Martin-Sempere, M. J., J. Rey-Rocha, et al. (2002). "The effect of team consolidation on research collaboration and performance of scientists. Case study of Spanish university researchers in Geology." Scientometrics **55**(3): 377-394.

McDowell, J. M. and J. K. Smith (1992). "The Effect of Gender-Sorting on Propensity to Coauthor - Implications for Academic Promotion." Economic Inquiry **30**(1): 68-82.

McKelvey, M., H. Alm, et al. (2003). "Does co-location matter for formal knowledge collaboration in the Swedish biotechnology-pharmaceutical sector?" Research Policy **32**(3): 483-501.

Melin, G. (1996). "The networking university - A study of a Swedish university using institutional co-authorships as an indicator." Scientometrics **35**(1): 15-31.

Melin, G. (2000). "Pragmatism and self-organization - Research collaboration on the individual level." Research Policy **29**(1): 31-40.

Melin, G. (2004). "Postdoc abroad: inherited scientific contacts or establishment of new networks?" Research Evaluation **13**(2): 95-102.

Mills, J. S. (1848). Principles of Political Economy. 1987. Fairchild, N.J.:, Augustus M. Kelley.

Mirowski, P. and E.-M. Sent (2008). The Commercialization of Science and the Response of STS. The Handbook of Science and Technology Studies. E. Hackett, O. Amsterdamska, M. Lynch and J. Wajcman. Cambridge MA., The MIT Press**:** 635-689.

Moed, H. F. (2000). "Bibliometric indicators reflect publication and management strategies." Scientometrics **47**(2): 323-346.

Mora-Valentin, E. M., A. Montoro-Sanchez, et al. (2004). "Determining factors in the success of R&D cooperative agreements between firms and research organizations." Research Policy **33**(1): 17-40.

Nahapiet, J. and S. Ghoshal (1998). "Social Capital, Intellectual Capital, and the Organizational Advantage." Academy of Management Review **23**(2): 242-266.

Narin, F., K. Stevens, et al. (1991). "Scientific Cooperation in Europe and the Citation of Multinationally Authored Papers." Scientometrics **21**(3): 313-323.

Nelson, R. R., Ed. (1993). National Innovation Systems: a comparative study. New York, Oxford University Press.

Newman, M. E. J. (2001). "Scientific collaboration networks. I. Network construction and fundamental results." Physical Review E **6401**(1).

Newman, M. E. J. (2001). "The structure of scientific collaboration networks." Proceedings of the National Academy of Sciences of the United States of America **98**(2): 404-409.

Newman, M. E. J. (2004). Coauthorship networks and patterns of scientific collaboration. Colloquium Proceedings "Mapping Knowledge Domains", Arnold and Mabel Beckman Center of the National Academies of Sciences and Engineering in Irvine, CA., The National Academy of Sciences of the United States -PNAS.

Noltingk, B. E. (1985). "A Note on Effective Laboratory Size." R & D Management **15**(1): 65-69.

North, D. C. (1990). Institutions, Institutional Change and Economic Performance. New York, Cambridge University Press.

NSF-NSB (2006). Science and Engineering Indicators 2006. National Science Board. Arlington, VA, National Science Foundation, Division of Science Resources Statistics.

NSF-NSB (2008). Science and Engineering Indicators 2008. National Science Board. Arlington, VA, National Science Foundation, Division of Science Resources Statistics.

OCyT (2007). Indicadores de Ciencia y Tecnologia. Colombia 2007. Bogota, D.C. Colombia, Observatorio Colombiano de Ciencia y Tecnologia y Panamericana.

OECD (1997). National Innovation Systems. Paris, France, OECD. Online.

OECD (1999). Boosting innovation: the cluster approach. Paris.

Okubo, Y., J. C. Dore, et al. (1998). "A multivariate analysis of publication trends in the 1980s with special reference to south-east Asia." Scientometrics **41**(3): 273-289.

Ordonez, G. (2005). The Impact of Research Collaboration on the Quality of the Research Outputs in Colombia. Lecture at the Colloquium Harvard-MIT. Ciencia, Tecnología e Innovación en Colombia 2005, Nov. 18-19. Cambridge, MA.

Penner-Hahn, J. and J. M. Shaver (2005). "Does international research and development increase patent output? An analysis of Japanese pharmaceutical firms." Strategic Management Journal **26**(2): 121-140.

Polany, M. (1962). "The Republic of Science." Minerva **1**: 54-73 http://www.mwsc.edu/orgs/polanyi/mp-repsc.htm.

Porac, J. F., J. B. Wade, et al. (2004). "Human capital heterogeneity, collaborative relationships, and publication patterns in a multidisciplinary scientific alliance: a comparative case study of two scientific teams." Research Policy **33**(4): 661-678.

Porter, M. E. (2001). Clusters of Innovation: Regional foundations of U.S. Competitiveness. Washington, D.C., Concil of Competitiveness.

Price, D. J. D. and D. D. Beaver (1966). "Collaboration in an Invisible College." American Psychologist **21**(11): 1011-&.

Prpic, K. (2002). "Gender and productivity differentials in science." Scientometrics **55**(1): 27-58.

Qin, J., F. W. Lancaster, et al. (1997). "Types and levels of collaboration in interdisciplinary research in the sciences." Journal of the American Society for Information Science **48**(10): 893-916.

Qurashi, M. M. (1984). "Publication Rate as a Function of the Laboratory Group-Size." Scientometrics **6**(1): 19-26.

Qurashi, M. M. (1991). "Publication-Rate and Size of Two Prolific Research Groups in Departments of Inorganic-Chemistry at Dacca University (1944-1965) and Zoology at Karachi University (1966-84)." Scientometrics **20**(1): 79-92.

Qurashi, M. M. (1993). "Dependence of Publication-Rate on Size of Some University Groups and Departments in Uk and Greece in Comparison with Nci, USA." Scientometrics **27**(1): 19-38.

Rey-Rocha, J., M. J. Martin-Sempere, et al. (2002). "Research productivity of scientists in consolidated vs. non-consolidated teams: The case of Spanish university geologists." Scientometrics **55**(1): 137-156.

RICYT (2004). Indicadores de Ciencia y Tecnología actualizados al año 2003. Buenos Aires, Argentina, Red Iberoamericana de Indicadores de Ciencia y Tecnologia www.ricyt.edu.ar.

Rigby, J. and J. Edler (2005). "Peering inside research networks: Some observations on the effect of the intensity of collaboration on the variability of research quality." Research Policy **34**(6): 784-794.

Rinia, E. J., T. N. Van Leeuwen, et al. (2002). "Impact measures of interdisciplinary research in physics." Scientometrics **53**(2): 241-248.

Rogers, J. (2001). Theoretical Considerations of Collaboration in Scientific Research, AAA.

Rogers, J. D. and B. Bozeman (2001). ""Knowledge value alliances": An alternative to the R&D project focus in evaluation." Science, Technology & Human Values **26**(1): 23.

Rosenbaum, P. R. and D. B. Rubin (1983). "The Central Role of the Propensity Score in Observational Studies for Causal Effects." Biometrika **70**(1): 41-55.

Sabato, J. (1975). El pensamiento latinoamericano en la problematica ciencia-tecnologia-desarrollo-dependencia. Buenos Aires, Paidos.

Sabato, J. and M. Mackenzi (1982). La produccion de Tecnologia autonoma o transnacional. Mexico, Nueva Imagen.

Sagasti, F. (2004). Knowledge and Innovation for Development: The Sisyphus Challenge of the 21st Century. Northampton, MA., Edward Elgar.

Saxenian, A. (1994). Regional Advantage, Culture and Competition in Silicon Valley and Route 128. Cambridge, MA., Harvard University Press.

Schummer, J. (2004). "Multidisciplinarity, interdisciplinarity, and patterns of research collaboration in nanoscience and nanotechnology." Scientometrics **59**(3): 425-465.

Scott, A. J., Ed. (2001). Global City-Regions. Oxford, Oxford University Press.

Seglen, P. O. and D. W. Aksnes (2000). "Scientific productivity and group size: A bibliometric analysis of Norwegian microbiological research." Scientometrics **49**(1): 125-143.

Seibert, S. E., M. L. Kraimer, et al. (2001). "A social capital theory of career success." Academy of Management Journal **44**(2): 219-237.

Sen, A. (2000). Development as Freedom. New York, Anchor Books.

Shrum, W. (2005). "Reagency of the Internet, or, how I became a guest for science." Social Studies of Science **35**(5): 723-754.

Slaughter, S., T. Campbell, et al. (2002). "The "traffic" in graduate students: Graduate students as tokens of exchange between academe and industry." Science Technology & Human Values **27**(2): 282-312.

Smeby, J. C. and J. Trondal (2005). "Globalisation or europeanisation? International contact among university staff." Higher Education **49**(4): 449-466.

Smeby, J. C. and S. Try (2005). "Departmental contexts and faculty research activity in Norway." Research in Higher Education **46**(6): 593-619.

Smith, B. (1990). American Science Policy since World War II. Washington, D.C., The Brookings Institution.

Smith, J. and P. Todd (2005). "Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators?" Journal of Econometrics **125**(1-2): 305-353.

Stankiewicz, R. (1979). The size and age of Swedish academic research groups and their scientific performance. Scientific Productivity: the effectiveness of research groups in six countries. F. M. Andrews. Paris, Cambridge University Press and UNESCO**:** 191-222.

Stephan, P. E. (2001). "Educational Implications of University-Industry Technology Transfer." Journal of Technology Transfer **26**: 199-205.

Stephan, P. E. and S. G. Levin (1997). "The critical importance of careers in collaborative scientific research." Revue d'Economie Industrielle **79**: 45-61.

Stolpe, M. (2002). "Determinants of knowledge diffusion as evidenced in patent data: the case of liquid crystal display technology." Research Policy **31**(7): 1181-1198.

Thorpe, P. and P. G. Pardey (1990). "The Generation and Transfer of Agricultural Knowledge - a Bibliometric Study of a Research Network." Journal of Information Science **16**(3): 183-194.

Tsai, W. P. and S. Ghoshal (1998). "Social capital and value creation: The role of intrafirm networks." Academy of Management Journal **41**(4): 464-476.

Turner, L. and J. Mairesse (2000). Mesure de l'Intensité de Collaboration dans la Recherche Scientifique et Evaluation du Rôle de la Distance Géographique. Unpublished manuscript. CREST-INSEE. Paris**:** 21.

Turner, L. and J. Mairesse (2005). Individual Productivity Differences in Public Research: how important are non-individual determinants? An economic study of French physicists' publications and citations (1986-1997)**:** Unpublished Manuscript. 35 p.

UNDP (2001). Human Development Report 2001: Making New Technologies Work for Human Development. New York, Oxford University Press.

UNIDO (2002). Industrial Development Report 2002-2003. Competing through Innovation and Learning. Vienna, United Nations Industrial Development Organization.

UNIDO (2005). Industrial Development Report 2005. Capability Building for Catching-Up. Historical, empirical and policy dimensions. Vienna, United Nations Industrial Development Organization.

Uzzi, B. and J. Spiro (2005). "Collaboration and creativity: The small world problem." American Journal of Sociology **111**(2): 447-504.

Van Raan, A. F. J. (2000). The pandora's box of citation analysis: Measuring scientific excellence - The last evil? Web of Knowledge - a Festschrift in Honor of Eugene Garfield. B. Cronin and H. B. Atkins. Medford, NJ., Asist Monograph Series. Information Today Inc**:** 301-319.

Wagner, C. S. (2005). "Six case studies of international collaboration in science." Scientometrics **62**(1): 3-26.

Wagner, C. S., I. Brahmakulam, et al. (2001). Science and Technology Collaboration: Building Capacities in Developing Countries? Santa Monica, CA, RAND.

Wagner, C. S. and L. Leydesdorff (2004). Network Structures, Self-Organization and the Growth of International Collaboration in Science. Amsterdam School of Communications Research (ASCoR). Amsterdam**:** 33.

Wagner, C. S. and L. Leydesdorff (2005). "Mapping the network of global science: comparing international co-authorships from 1990 to 2000." International Journal of Technology and Globalization **1**(2): 185-207.

Wagner, C. S. and L. Leydesdorff (2006). Measuring the Globalization of Knowledge Networks. Blue Sky. Manuscript submitted in September 2006**:** 12.

Waguespack, D. M. and J. K. Birnir (2005). "Foreignness and the diffusion of ideas." Journal of Engineering and Technology Management **22**(1-2): 31-50.

Wang, J. X. and J. C. Guan (2005). "The analysis and evaluation of knowledge efficiency in research groups." Journal of the American Society for Information Science and Technology **56**(11): 1217-1226.

WEF (2005). The Global Competitiveness Report. World Economic Forum. New York, Oxford University Press.

Whitley, R. (2000). The intellectual and social organization of the sciences. Oxford, UK, Oxford University Press.

Williamson, O. E. (1985). The Economic Institutions of Capitalism. New York, Free Press.

Wray, K. B. (2002). "The epistemic significance of collaborative research." Philosophy of Science **69**(1): 150-168.

Yoshikane, F. and K. Kageura (2004). "Comparative analysis of coauthorship networks of different domains: The growth and change of networks." Scientometrics **60**(3): 433-444.

Zhao, Z. (2004). "Using Matching to Estimate Treatment Effects: Data Requirements, Matching Metrics, and Monte Carlo Evidence." The Review of Economics and Statistics **80**(1): 91-107.

Zitt, M., S. Ramanana-Rahary, et al. (2003). "Potential science-technology spillovers in regions: An insight on geographic co-location of knowledge activities in the EU." Scientometrics **57**(2): 295-320.

VITA

GONZALO ORDÓÑEZ-MATAMOROS

Gonzalo Ordóñez-Matamoros holds a B.A. in Finance and International Relations from the Universidad Externado de Colombia in Bogotá, an M.A. (D.E.A.) in International Economics and Economic Development from the Université Paris I Panthéon-Sorbonne, France, and an M.A. (D.E.A.) in International Economic Relations Law from the Université Paris X- Nanterre, France. Previous to joining the Doctoral Program in 2002, Mr. Ordóñez served as executive director of the Colombian Observatory of Science and Technology –OCyT and as a staff member at COLCIENCIAS. Mr. Ordóñez has worked as consultant to the Interamerican Development Bank (IDB), The Organization of American States (OAS), the United Nations Development Program (UNDP), USAID, the Iberoamerican S&T Indicators Network (RICYT), the government of Panamá, and several Colombian government agencies. Mr. Ordóñez is currently a faculty member and researcher at the Observatory of Public Policies of the Universidad Externado de Colombia in Bogota, and a research associate at the Technology Policy and Assessment Center –TPAC at the Georgia Institute of Technology. His main research interests focus on Science and Technology Policy in Developing Countries.

Permanent e-mail address: gonzaloord@hotmail.com